

Writing Analytics for Epistemic Features of Student Writing

Citation for published version (APA):

Knight, S., Allen, L., Littleton, K., Rienties, B., & Tempelaar, D. (2016). Writing Analytics for Epistemic Features of Student Writing. In C.-K. Looi, J. Polman, U. Cress, & P. Reimann (Eds.), *Transforming Learning, Empowering Learners: Conference Proceedings Volume 1* (pp. 194-201). International Society of the Learning Sciences. https://www.isls.org/icls/2016/docs/ICLS2016_Volume_1_30June2016.pdf

Document status and date:

Published: 01/01/2016

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

12th INTERNATIONAL CONFERENCE OF THE LEARNING SCIENCES



ICLS 2016
SINGAPORE
20-24 June

ISBN: 978-0-9903550-9-0

Edited by
Chee-Kit Looi,
Joseph Polman,
Ulrike Cress, and
Peter Reimann

Transforming Learning, Empowering Learners Conference Proceedings Volume I

Organiser



International Society of
the Learning Sciences

Host



An Institute of Nanyang
Technological University

Co-Host



Centre for Research and Development in Learning

Transforming Learning, Empowering Learners:
The International Conference of the Learning
Sciences (ICLS) 2016

Volume 1

12th International Conference
of the Learning Sciences

June 20-24, 2016, Singapore
National Institute of Education,
Nanyang Technological University,
Singapore

Editors:

Chee-Kit Looi, Joseph Polman, Ulrike Cress, and Peter Reimann

Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016

June 20-24, 2014, Singapore

National Institute of Education, Nanyang Technological University, Singapore.

© 2016 International Society of the Learning Sciences, Inc. [ISLS]. Rights reserved.



www.isls.org

ISBN: 978-0-9903550-9-0 (Volume 1, PDF Version)

ISSN: 1814-9316

Cite as: Looi, C. K., Polman, J. L., Cress, U., and Reimann, P. (Eds.). (2016). *Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016, Volume 1*. Singapore: International Society of the Learning Sciences.

All rights reserved. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear the copyright notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior written permission of the International Society of the Learning Sciences. The International Society of the Learning Sciences is not responsible for the use which might be made of the information contained in this book.

Conference logo and cover design by Li Li Ng

Conference Organisers and Committees

Conference Co-Chairs

Wenli Chen, *Nanyang Technological University, Singapore*
Seng Chee Tan, *Nanyang Technological University, Singapore*
Choon Lang Quek, *Nanyang Technological University, Singapore*

Programme Co-Chairs

Chee-Kit Looi, *Nanyang Technological University, Singapore*
Ulrike Cress, *University of Tuebingen, Germany*
Joe Polman, *University of Colorado Boulder, USA*
Peter Reimann, *University of Sydney, Australia*

Advisory Committee

Pierre Dillenbourg, *École Polytechnique Federale de Lausanne, Switzerland*
Frank Fischer, *University of Munich, Germany*
Susan Goldman, *University of Illinois at Chicago, USA*
Sten Runar Ludvigsen, *University of Oslo, Norway*
William Penuel, *University of Colorado Boulder, USA*
Miguel Nussbaum, *Pontificia Universidad Católica de Chile, Chile*
Jim Pellegrino, *University of Illinois at Chicago, USA*
Anna Sfard, *University of Haifa, Israel*
Dan Suthers, *University of Hawaii, USA*
Oon Seng Tan, *Nanyang Technological University, Singapore*

Workshop Co-Chairs

Cindy Hmelo-Silver, *Indiana University, USA*
Armin Weinberger, *Saarland University, Germany*
Carolyn Rosé, *Carnegie Mellon University, USA*

DSC Co-Chairs

Sadhana Puntambaer, *University of Wisconsin at Madison, USA*
Jan van Aalst, *University of Hong Kong, Hong Kong SAR*
Kristine Lund, *University of Lyon, France*

Early Career Workshop Co-Chairs

Julia Eberle, *Ruhr-Universität Bochum, Germany*
Manu Kapur, *Education University of Hong Kong, Hong Kong SAR*
Nikol Rummel, *Ruhr-Universität Bochum, Germany*
Paul Kirschner, *Open University of The Netherlands, Netherlands*

Mid-Career Workshop Co-Chairs

Susan Goldman, *University of Illinois at Chicago, USA*
Marcia Linn, *University of California, Berkeley, USA*
Marcela Borge, *The Pennsylvania State University, USA*
Wenli Chen, *Nanyang Technological University, Singapore*

Practitioners' Track Co-Chairs

Jaime Koh, *Nanyang Technological University, Singapore*
Choon Lang Quek, *Nanyang Technological University, Singapore*

Conference Local Organising Committee

Practitioners' Track

Jaime Koh, *Nanyang Technological University, Singapore*
Choon Lang Quek, *Nanyang Technological University, Singapore*
Yancy Toh, *Nanyang Technological University, Singapore*
Betsy Ng, *Nanyang Technological University, Singapore*
Kian Keong Aloysius Ong, *Nanyang Technological University, Singapore*
Qiyun Wang, *Nanyang Technological University, Singapore*
Kevin Hartman, *Nanyang Technological University, Singapore*
Huaqing Hong, *Nanyang Technological University, Singapore*

Conference Secretariat

Vivienne Woon Yuen Chia, *Nanyang Technological University, Singapore*

Financial

Patricia Hui Hian Ann, *Nanyang Technological University, Singapore*
Kok Hua Cheng, *Nanyang Technological University, Singapore*

Accommodation, Social Events and Hospitality

Doris Choy, *Nanyang Technological University, Singapore*
Angela Wong, *Nanyang Technological University, Singapore*
Joyce Koh, *Nanyang Technological University, Singapore*
Joyce Chou, *Nanyang Technological University, Singapore*

Venue and Logistics

Doris Choy, *Nanyang Technological University, Singapore*
Manjulah K Anand, *Nanyang Technological University, Singapore*
Long Kai Wu, *Nanyang Technological University, Singapore*
Connie Ng, *Nanyang Technological University, Singapore*

Sponsorship and Commercial Exhibition

Seng Chee Tan, *Nanyang Technological University, Singapore*
Clement Ng, *Nanyang Technological University, Singapore*

Students/RA Volunteer Coordination

Elizabeth Koh, *Nanyang Technological University, Singapore*
Ching Sing Chai, *Nanyang Technological University, Singapore*

Showcases and Posters

Qiyun Wang, *Nanyang Technological University, Singapore*
Wing Sum Cheung, *Nanyang Technological University, Singapore*

Technology, Website and Community Memory

Tsering Wangyal, *Nanyang Technological University, Singapore*
Hong Poh Liang, *Nanyang Technological University, Singapore*
Abdul Kamal Bin Ahmed, *Nanyang Technological University, Singapore*
Li Li Ng, *Nanyang Technological University, Singapore*
Kok Hua Cheng, *Nanyang Technological University, Singapore*
Dave Kian Seng Lim, *Nanyang Technological University, Singapore*
Kenny Keng Ngee Ng, *Nanyang Technological University, Singapore*

Proceedings and Programme Booklet

Lung Hsiang Wong, *Nanyang Technological University, Singapore*

Li Li Ng, *Nanyang Technological University, Singapore*

Rachel Lam, *Nanyang Technological University, Singapore*

Dragan Trninic, *Nanyang Technological University, Singapore*

Public Relations

David Huang, *Nanyang Technological University, Singapore*

Clement Ming Fu Lim, *Nanyang Technological University, Singapore*

Michael Tan, *Nanyang Technological University, Singapore*

Valerie Sim, *Nanyang Technological University, Singapore*

Communications Chair

Matthew Gaydos, *Nanyang Technological University, Singapore*

Opening and Closing Ceremonies, and School Related Matters

Lay Hoon Seah, *Nanyang Technological University, Singapore*

Peter Seow, *Nanyang Technological University, Singapore*

Azilawati Jamaludin, *Nanyang Technological University, Singapore*

Betsy Ng, *Nanyang Technological University, Singapore*

Gloria Li Ya Ho, *Nanyang Technological University, Singapore*

Senior Reviewers

Shaaron Ainsworth
Nancy Ares
Flavio Azevedo
Ryan Baker
Sasha Barab
Philip Bell
Yifat Ben-David Kolikant
Daniel Bodemer
Leah Bricker
Karin Brodie
Jürgen Buder
Angela Calabrese Barton
Britte Cheng
John Cherniavsky
Cynthia Carter Ching
Clark Chinn
Doug Clark
Ulrike Cress
Ton De Jong
Angela DeBarger
Yannis Dimitriadis
Ann Edwards
Jan Elen
Frank Fischer
Ellice Forman
Janice Gobert
Kimberley Gomez
Melissa Gresalfi
Päivi Häkkinen
Rogers Hall
Tony Hall
Erica Halverson
Richard Halverson
Neil Heffernan
Friedrich Hesse
Daniel Hickey
Ilana Horn
Heisawn Jeong
A. Susan Jurow
Manu Kapur
Eric Klopfer
Timothy Koschmann
Kristiina Kumpulainen
Eleni Kyza
Ard Lazonder
Rich Lehrer
Robb Lindgren
Chee-Kit Looi
Rose Luckin
Sten Ludvigsen
Kristine Lund
Jasmine Ma
Alecia Marie Magnifico
Lina Markauskaite
Scott McDonald
Steven McGee
Chrystalla Mouza
Angela O'Donnell
Jun Oshima
Annemarie Palincsar
James Pellegrino
William Penuel
Joe Polman
Chris Quintana
Jrene Rahm
Mimi Recker
Peter Reimann
K. Ann Renninger
Ann Rivet
Ido Roll
Nikol Rummel
William Sandoval
Keith Sawyer
R. Benjamin Shapiro
Bruce Sherin
Wesley Shumar
Hyo-Jeong So
Hans Spada
Gerry Stahl
Mike Stieff
Tamara Sumner
Pierre Tchounikine
Ley Tobias
Jennifer Vadeboncoeur
Phil Vahey
Wouter van Joolingen
Noreen Webb
Armin Weinberger
Alyssa Wise
Susan Yoon
Jianwei Zhan

Reviewers

Anthony Aakre
Dor Abrahamson
Andres Acher
Karlyn Adams-Wiggins
June Ahn
Sahar Alameh
Laura Allen
Alicia Alonzo
Jessica J. Andrews
Christa Asterhan
Ari Bader-Natal
Elizabeth Bagley
Lauren Barth-Cohen
Antonia Baumeister
Brian Belland
Ruchi Bhanot
Heather Birch
Lori Bland
Adélaïde Blavier
Phillip Boda
Michael Bolz
Rebecca Boncodd
Elizabeth Bonsignore
Marcela Borge
Ivica Boticki
Jonathan Boxerman
Tharrenos Bratitsis
Stein Brunvand
Qijie Cai
Murat Perit Cakir
Mia Čarapina
Adam Carberry
Monica Cardella
Aprea Carmela
Collazos Cesar
Margaret Chan
Elizabeth Charles
Gaowei Chen
Mark Chen
Ying-Chih Chen
Wenli Chen
Ellina Chernobilsky
Thomas K.F. Chiu
Gi Woong Choi
Stefanie Chye
Sherice Clarke
Tamara Clegg
Hunter Close
Andras Csanadi
Joe Curnow
Crina Damsa
Srikanth Dandotkar
Joshua Danish
Cesar Delgado
David DeLiema
Nicoletta Di Blas
Michael Dianovsky
Ning Ding
Neven Drljević
Tama Duangnamol
Ravit Duncan
Sean Duncan
Gregory Dyke
Julia Eberle
Hebbah El-Moslimany
Katharina Engelmann
Judith Enriquez-Gibson
Bernhard Ertl
Howard Everson
Cameron Fadjo
Nafsaniath Fathema
Leila Ferguson
Georgios Fessakis
Deborah Fields
Virginia J. Flood
Paul Flynn
Cory Forbes
Ina Fourie
Erin Marie Furtak
Judi Fusco
Brian Gane
Alejandra Garcia-Franco
Iolanda Garcia
Maria Teresa Gastardo
Matthew Gaydos
Xun Ge
Andreas Gegenfurtner
Libby Gerard
David Gerritsen
Aristotelis Gkiolmas
Hannah Gogel
Sao-Ee Goh
Monica Gordon Pershey
Amelia Gotwals
Jeffrey Greene
Xiaoqing Gu
José Gutiérrez
Seungyon Ha
Alan Hackbarth
Allison H. Hall
Jiangang Hao
Elisabeth Hayes
Joseph Henderson
Arnon Hershkovitz
Yotam Hod
Hayat Hokayem
Sameer Honwad
Michael Horn
Pei-Hsuan Hsieh
Joey Huang
Roland Hubscher
Sarah Hug
Chia Yuan Hung
Rebecca Itow
Kara Jackson
Ellen Jameson
Mingfong Jan
Sanna Järvelä
Allan Jeong
Osvaldo Jimenez
Aditya Johri
Zywica Jolene
R. Seth Jones
Michelle Jordan
Alfredo Jornet Gil
Yih-Ruey Juang
Seokmin Kang
Dominic Kao
Shulamit Kapon
Danielle Keifert
Molly Kelton
Lisa Kenyon
Ahmad khanlari
Derya Kici
Khusro Kidwai
Dongsik Kim
Kibum Kim
Mi Song Kim
Beaunie Kim
Hee-jeong Kim
Joachim Kimmerle
Jennifer King Chen
Jayne Klenner-Moore
Laure Kloetzer
Simon Knight
Mon-Lin Monica Ko
Marta Kobiela
Alaric Kohler
Ingo Kollar
Moshe Krakowski
Christina Krist
Beth Kubitskey
Eric Kuo
Samuel Kwon
Minna Lakkala
Rachel Lam
Mike Lawson
Heather Leary
Tiffany Lee

Kewman Lee	Lih-Bin Oh	Bernadette Sibuma
Kyungmee Lee	Jennifer Olsen	Amanda Siebert-Evenstone
Yew Jin Lee	Yann Shiou Ong	Tiffanyrose Sikorski
Jason Lee	Chandra Orrill	Glauco Silva
KoSze Lee	Hiroki Oura	Tuya Siqin
Victor Lee	Saliha Ozkan-Bekiroglu	Irene Skuballa
Dalit Levy	Prajakt Pande	Benzi Slakmon
Armanda Lewis	Frederick Peck	Erica Snow
Colleen Lewis	Tana Peterman	Yanjie Song
Xiao Li	Rachel Phillips	Minyoung Song
Chang-Yen Liao	Elizabeth Pier	Sarah Stager
Feng Lin	Torsten Porsch	Trausan-Matu Stefan
Oskar Lindwall	Aaron Price	Karsten Stegmann
Matthew Lira	Burke Quinn	Susanne Strachota
Breanne Litts	Rebecca Quintana	Jan-Willem Strijbos
Shiyu Liu	Antti Rajala	Shelley Stromholt
Stanley Lo	Aria Razfar	Sarah Sullivan
Chee Kit Looi	Stephen Reed	Jingjing Sun
Suzanna Loper	Daniel Rees Lewis	Daner Sun
Maritza Lozano	Richard Reeve	Patrick Sunnen
Margaret Lucero	Martin Rehm	Vanessa Svihla
Michelle Lui	Suzanne Rhodes	Richard Swan
Guanzhong Ma	Gabriela Richard	Michael Swart
Augusto Macalalag Jr	Christoph Richter	Lori Takeuchi
Kati Maekitalo-Siegl	Antti Rissanen	Miwa Takeuchi
Moseli Mafa	Giuseppe Ritella	Seng Chee Tan
Alejandra Magana	Kelsey Rodgers	Ling Tan
Stefania Manca	Margarida Romero	Ying Tang
Nicole Mancevice	Ricarose Roque	Kok-Sing Tang
Andrew Manches	Jeremy Roschelle	Dan Tao
Specht Marcus	Elizabeth Rowe	Wan Ying Tay
Lauren Margulieux	elsia rubegni	Meng Yew Tee
Thomas Martens	Donna Russell	Jakob Tholander
Fred Martin	Minjung Ryu	Kate Thompson
John Martin	Ornit Sagy	Mike Tissenbaum
Crystle Martin	Asmalina Saleh	Yancy Toh
Roberto Martinez Maldonado	Rafi Santo	Hakan Tuzun
Lucy McClain	Eleanor Sayre	Kersti Tyson
Veronica Cassone McGowan	Giovanna Scalone	Dora Tzialli
Katherine McNeill	Andrea Scheer	Nan Uhlik
Richard Medina	Bertrand Schneider	Josh Underwood
José Melendez	Kathy Schuh	Keri Duncan Valentine
Alysson Mendes de Oliveira	Stephan Schwan	Katie Van Horne
Muhsin Menekse	Neil Schwartz	Herre van Oostendorp
Joi Merritt	Jennifer Schwarz	Charles Vanover
Georgia Michael	Layhoon Seah	Keisha Varma
Toshio Mochizuki	Colleen Seifert	Dana Vedder-Weiss
Gaelle Molinari	Asli Sezen-Barrie	Jonathan Vitale
Anders Morch	Li Sha	Chieu Vu Minh
Hanni Muukkonen	David Shaenfield	Robert Wallon
Denise Nacu	Mamta Shah	Elizabeth Walsh
Iolie Nicolaidou	Niral Shah	X. Christine Wang
Kraemer Nicole	Kshitij Sharma	Jeremy Wang
Nicolae Nistor	Priya Sharma	Muffie Waterman
Michele Notari	Nicole Shea	Anne Weibert
Claire O'Malley	Ji Shen	Jennifer Weible
Amanda Ochsner	Mayumi Shinohara	Joanna Weidler-Lewis

Michelle Wilkerson
Caroline Williams-Pierce
Robert Williams
Nicholas Wilson
Kerri Wingert
Phil Winne
Lung-Hsiang Wong
Iwan Wopereis

Kenneth Wright
Bian Wu
Longkai Wu
qiong wu
Ruth Wylie
Sihan Xiao
Ying Xie
Senom Yalcin

Yau-yuen Yeung
Yawen Yu
Yonghe Zhang
Lin Zhang
Heather Toomey Zimmerman
Robert Zisk

Reviewers for Practitioners' Track

Azilawati Bte Jamaludin
Wing Sum Cheung
Jeanette Lyn Fung Choy
Kevin Hartman
Khe Foon Hew
Mi Song Kim
Elizabeth Koh
Chwee Beng Lee
Jason Lee
Vwen Yen Alwyn Lee
Hong Poh Liang
Kenneth Yang Teck Lim
Betsy Ng
Kian Keong Aloysius Ong
Hyo-Jeong So
Liang See Tan
Seng Chee Tan
Yancy Toh
Wangyal Tsering
Loke Heng Wang
Qiyun Wang
Angela Wong

Preface

The National Institute of Education, Nanyang Technological University of Singapore, is hosting the 12th International Conference of the Learning Sciences (ICLS), 2016, from 20 to 24 June, 2016. The international and interdisciplinary field of the Learning Sciences brings together researchers from the fields of cognitive science, educational research, psychology, computer science, artificial intelligence, information sciences, anthropology, sociology, neurosciences, and other fields to study learning in a wide variety of formal and informal contexts (see www.isls.org). The field emerged in the late 1980s and early 1990s, with the first ICLS held in 1991 at Northwestern University in Evanston, Illinois, USA. Subsequent meetings of ICLS were held again in Evanston, USA (1996), Atlanta, GA, USA (1998), Ann Arbor, MI, USA (2000), Seattle, WA, USA (2002), Santa Monica, CA, USA (2004), Bloomington, IN, USA (2006), Utrecht, the Netherlands (2008), Chicago, IL, USA (2010), Sydney, NSW, Australia (2012), and Boulder, CO, USA (2014). ICLS 2016 is the first time that the conference is being held in Asia.

Submissions for ICLS 2016 were received in November 2015, and then went through a process of peer review. An impressive number of submissions were received (571). The paper review process was very competitive. The overall acceptance rate for submissions was 37%. We accepted 65% of symposium submissions, 31% of full papers, 34% of short papers, and 43% of posters. The program reflects broad geographic representation, with contributions from 31 countries and regions. We are especially grateful to those who performed reviews. 342 experts completed over 1,600 reviews of the submissions. As in recent years, for each symposium, full paper and short paper, we assigned a senior reviewer who examined all reviews and made a recommendation regarding acceptance in the category submitted, acceptance in an alternate category, or rejection. These senior reviewers greatly helped us make the decisions for each submission.

The conference theme of ICLS 2016 is “Transforming Learning, Empowering Learners.” It directs our attention to a key commitment of the Learning Sciences: providing an insightful understanding of how people learn. As we trace the genesis of the Learning Sciences, we are reminded of the main goal of this field of research, that is, to gain a deep understanding of the conditions and processes that lead to effective learning, and to use the research findings to redesign learning environments to bring about deep learning. Learning Sciences is concerned with transforming learning and empowering learners.

This long-standing commitment extends our research focus to the design of pedagogical interventions and learning environments to foster among participants a kind of learning that is transformative and empowering. This requires challenging established beliefs about learning, teaching, and the design of learning environments. The theme “Transforming Learning, Empowering Learners” aims at reaffirming the key thrust of Learning Sciences research, discussing advances in the field, and strategizing future directions to enhance our impact on educational practice. ICLS 2016 aims to bring together learning scientists to adjudicate various academic renditions of how people learn, and to institute further inquiry that encourages deep and probing examination of the nexus of instruction and learning, as well as the roles of technology. To address the conference theme, we articulated the following strands:

Deep learning in effective learning environments

The field of Learning Sciences is committed to advancing research for explaining how people learn and what can be done to support deep learning. Towards these goals, different studies have examined learners’ prior knowledge and preconceptions, knowledge representation, and knowledge construction. Research provides insights into the underlying cognitive bases of problem-solving, knowledge construction, reasoning, reflection and deep learning. Such understanding is necessary for designing effective learning environments needed to transform and enhance learning. In continuing this tradition of the Learning Sciences, perspectives that demonstrate scholarly depth on cognitive, social, psychological and/or cultural aspects of learning are valued. Relevant topics include learning theories, pedagogies, individual and group learning, learner agency and identity in learning, cognition and instruction, learning in all areas of the curriculum, conceptual change, and scaffolding.

A number of studies in Learning Sciences have paid attention to technology use and its affordances for deep learning. Computer-mediated learning remains key in Learning Sciences as research on new media, e-learning, adaptive and intelligent systems continues to proliferate. The conference program includes work that contributes

to the ongoing dialogues about the use of adaptive systems and intelligent tutoring in innovative learning situations, computer-supported collaborative learning, use of computers for group and distributed cognition, computational models of how people learn, use of computers for assessment in virtual classrooms, and scholarly work on technology use to support learning communities. Sessions on emerging topics like MOOCs, big data, and the relevance of neuroscience should also stimulate thought and debate about the future of the field.

By emphasizing a strand on deep learning in effective learning environments, we hope the community can assess the extent to which educational institutions have shifted towards deep learning in their pedagogical approaches. As a research community, we can question how well the Learning Sciences have contributed to educational practice, and suggest what else can be done to design our social future of learning as we collectively expand our vision of transforming learning.

Digital epistemologies and the situated nature of learning

One key research thrust in Learning Sciences is aimed at understanding the situated nature of learning in diverse sociocultural practices. Empirical accounts of such work serve to extend our knowledge of how people learn by informing us about how young people learn informally in out-of-school practices. Sociocultural studies on situated cognition, the roles of context in cognition, collaborative discourse, self-directed learning in the online world, and participatory culture in cyberspace are topics important to advancing the field of Learning Sciences.

A number of studies have emphasized how young people have challenged our understanding of what is learnt and not learnt in schools, and how schools can appropriately respond to changing digital epistemologies. By digital epistemologies, we are referring to ways of knowing embraced by participants in myriad digital literacy practices. Young people's digital epistemologies and learning in the new media age open vistas on a whole range of complex issues. At this conference, we welcome debates on the roles of out-of-school literacies in school literacies and studies that illuminate how formal and informal learning can be synergized. It is worth rethinking ways of knowing in young people's digital literacy practices, and we hope ICLS 2016 can create the opportunities for rich dialogues that address issues of learner identities, new cultures of learning, and their implications to the design of learning environments as we seek to understand more of how young people take ownership of their learning in and out of school in their digital culture.

Teacher knowledge and professional development

To transform learning, we cannot ignore the critical roles played by teachers in enacting and innovating classroom practices. Traditionally, off-site workshops or courses are linked to programs offered by institutions of higher learning, often associated with the formal granting of a higher degree. While off-site learning experiences provide teachers with opportunities to interact with researchers and explore ideas based on research findings, they are often criticized for being too removed from authentic classroom contexts, thus lacking the transformative power to change classroom practices. Some studies have been conducted on teachers' development of pedagogical content knowledge, and some on meta-strategic knowledge. A number of studies in the Learning Sciences have focused on strategies for teacher professional development and learning, including the use of blogs, problem-based learning, animated classroom stories, immersive virtual reality, and small group reflection. The conference program includes research and discussion on new modes of teacher professional development and learning, and ways to enhance knowledge sharing among teachers, including the investigation of mechanisms that enable knowledge codification, validation, and dissemination of expert teacher knowledge. At ICLS 2016, we hope to engage in dialogues on how Learning Sciences research can further our investigation of ways to bring about changes in teacher beliefs—about epistemology, ontology, and practice—necessary to change classroom practices on a sustainable basis.

Reflexive relations between methods and theories

To realize the goals of "Transforming Learning, Empowering Learners," Learning Sciences researchers have developed new research methodologies and methods. One distinctive example is the development of the design experiment (or design research) methodology. At ICLS 2016, we welcome discussion and sharing of advancements on design research. The design experiment has the distinctive feature of advancing theories that can guide the design of learning activities. It involves theory-informed pedagogical interventions that are both reflective and prospective in nature. It has the dual goal of evaluating how well the theory-informed intervention works and generating ideas for further experiments. Many researchers have also engaged practitioners in the process of design experiment, for example, by involving teachers in the design of classroom interventions. This

method is an approach that can potentially lead to empowerment of practitioners while giving them voice and agency in the research process. From another perspective, it is an approach to develop practitioners' professional learning.

We also welcome experimental studies that analyzed the processes and outcomes of learning in detail. We hope the conference stimulates dialogues on the usefulness of big data and learning analytics. Several studies related to learning analytics highlight evolving lines of inquiry that are of interest to the Learning Sciences, including use of learning analytics for blended learning, learner assessment, new models of learning, and the roles of pedagogy in learning analytics. Our purpose in foregrounding learning analytics in this conference is to instill greater interest in studying how learning analytics can have a greater impact on educational practice, particularly in the area of using technology for assessment.

In these proceedings volumes, you will find a wide variety of perspectives and research findings concerning the above issues and questions, and we hope that you will have insightful and productive conversations during as well as after the conference.

Finally, we express our deepest gratitude to the many people who made the conference possible: the organizing committee, the advisory committee, the program committee, the co-chairs of workshops, doctoral student consortium, early career and mid-career workshops, reviewers, sponsors, volunteers, staff, and all conference presenters and participants. Your contributions make the learning sciences a thriving field, striving to transform learning and empowering learners.

Program Committee Chairs

Chee-Kit Looi, Joseph Polman, Ulrike Cress, and Peter Reimann

Conference Chairs

Wenli Chen, Seng-Chee Tan, and Choon Lang Quek

Table of Contents

Volume 1

Keynotes

- Educational Neuroscience: A Field Between False Hopes and Realistic Expectations3
Elsbeth Stern
- Transformative Learning in Design Research: The Story Behind the Scenes4
Yael Kali
- The Diffusion of Inquiry Based Practices in the Singapore Education System:
Navigating Eddies of the 21st Century5
David Hung, Azilawati Bte Jamaludin, Yancy Toh

Invited Symposia

- Beyond Tried and True: The Challenge of Education for Innovation9
Carl Bereiter, Marlene Scardamalia, Thérèse Laferrière, Linda Massey, Bruce W. Shaw, Shirleen Chee, Seng Chee Tan, Chew Lee Teo, David Istance
- Future of the CSCL Community.....16
Sten Ludvigsen, Heisawn Jeong, Cindy E. Hmelo-Silver, Nancy Law, Ulrike Cress, Peter Reiman, Manu Kapur, Nikol Rummel
- Analytics of Social Processes in Learning Contexts: A Multi-Level Perspective24
Carolyn P. Rosé, Dragan Gaesevic, Pierre Dillenbourg, Yohan Jo, Gaurav Tomar, Oliver Ferschke, Gijbert Erkens, Anouschka van Leeuwen, Jeroen Janssen, Mieke Brekelmans, Jennifer Tan, Elizabeth Koh, Imelda Caleon, Christin Jonathan, Simon Yang

Full Papers

- Scaling Studio-Based Learning Through Social Innovation Networks35
Natalia Smirnov, Matthew W. Easterday, Elizabeth M. Gerber
- Multiple Legitimate Language Games in Family Serendipitous Science Engagement43
Dana Vedder-Weiss
- Scientific Reasoning and Problem Solving in a Practical Domain: Are Two Heads Better Than One?.....50
Andras Csanadi, Ingo Kollar, Frank Fischer
- Combining Exploratory Learning With Structured Practice to Foster Conceptual and Procedural Fractions Knowledge.....58
Nikol Rummel, Manolis Mavrikis, Michael Wiedmann, Katharina Loibl, Claudia Mazziotti, Wayne Holmes, Alice Hansen
- The Effects of Self-Regulated Learning on Students' Performance Trajectory in the Flipped Math Classroom.....66
Zhiru Sun, Lin Lu, Kui Xie
- Supporting Inquiry Learning as a Practice: A Practice Perspective on the Challenges of IBL Design, Implementation and Research Methodology74
Fleur R. Prinsen
- Introducing Academically Low-Performing Young Science Students to Practices of Science82
Toi Sin Arvidsson, Deanna Kuhn
- Idea Identification and Analysis (I2A): A Search for Sustainable Promising Ideas Within Knowledge-Building Discourse.....90
Vwen Yen Alwyn Lee, Seng Chee Tan, Joon Kit Kelvin Chee

Training Learners to Self-Explain: Designing Instructions and Examples to Improve Problem Solving	98
<i>Lauren E. Margulieux, Briana B. Morrison, Mark Guzdial, Richard Catrambone</i>	
Validating a Model for Assessing Teacher’s Adaptive Expertise With Computer Supported Complex Systems Curricula and Its Relationship to Student Learning Outcomes.....	106
<i>Susan A. Yoon, Jessica Koehler-Yom, Emma Anderson, Chad Evans</i>	
What Do Learning Scientists Do? A Survey of the ISLS Membership	114
<i>Susan A. Yoon, Cindy E. Hmelo-Silver</i>	
The Interactional Work of Configuring a Mathematical Object in a Technology-Enabled Embodied Learning Environment.....	122
<i>Virginia J. Flood, Benedikt W. Harrer, Dor Abrahamson</i>	
Developing Pre-Service Teachers’ Professional Vision Through Collaborative Multimedia Artifacts	130
<i>Andi M. Rehak, Andrea S. Gomoll, Cindy E. Hmelo-Silver, Joshua A. Danish</i>	
Talking Back to the Future: Anatomy of Reflection as Collective Practice	138
<i>Alfredo Jornet</i>	
Becoming an “Expert”: Gendered Positioning, Praise, and Participation in an Activist Community	146
<i>Joe Curnow, Jody R. Chan</i>	
Prior Knowledge and Mathematics Different Order Thinking Skills in Multimedia Learning	154
<i>Thomas K.F. Chiu</i>	
Prior Knowledge for the Construction of a Scientific Model of Equilibration	162
<i>Hillary Swanson</i>	
A Knowledge Analytic Comparison of Cued Primitives When Students Are Explaining Predicted and Enacted Motions	170
<i>Victor R. Lee</i>	
Teacher Learning in a Professional Learning Community: Potential for Dual-layer Knowledge Building	178
<i>Seng Chee Tan, Shien Chu, Chew Lee Teo</i>	
No Lives Left: How Common Game Features Could Undermine Persistence, Challenge-Seeking and Learning to Program	186
<i>Laura J. Malkiewich, Alison Lee, Stefan Slater, Chenmu Xing, Catherine C. Chase</i>	
Writing Analytics for Epistemic Features of Student Writing.....	194
<i>Simon Knight, Laura Allen, Karen Littleton, Bart Rienties, Dirk Tempelaar</i>	
Visual Augmentation of Deictic Gestures in MOOC Videos	202
<i>Kshitij Sharma, Sarah D'Angelo, Darren Gergle, Pierre Dillenbourg</i>	
Designing the Idea Manager to Integrate STEM Content and Practices During a Technology-Based Inquiry Investigation.....	210
<i>Erika D. Tate, Mingyu Feng, Kevin W. McElhaney</i>	
Situated Learning, Situated Knowledge: Situating Racialization, Colonialism, and Patriarchy Within Communities of Practice	218
<i>Joe Curnow</i>	
Opportunities to Learn Through Design: Mapping Design Experiences to Teacher Learning.....	226
<i>Emily Horton, Jahneille Cunningham, Louis M. Gomez, Kimberley Gomez, Katherine Rodela</i>	
No One Ever Steps in the Same Discussion Twice: The Relationship Between Identities and Meaning.....	234
<i>Murat Öztok, Maarit Arvaja</i>	
“Hearts Pump and Hearts Beat”: Engineering Estimation as a Form of Model-Based Reasoning.....	242
<i>Aditi Kothiyal, Sahana Murthy, Sanjay Chandrasekharan</i>	

“Doing Double Dutch”: Becoming Attuned to Rhythms of Pathways In and Through Community Spaces	250
<i>Joyce M. Duckles, George Moses, Ryan Van Alstyne, Brandon Stroud</i>	
Towards a Framework of Pedagogical Paradoxes: A Phenomenographic Study of Teachers Designing Learning Experiences and Environments With ICT in Singapore Classrooms	258
<i>Wan Ying Tay, Boon Yen Chai, Nur Azarina Khamis, Samuel Tan, Chew Lee Teo</i>	
Joint Idea-Building in Online Collaborative Group Discussions.....	266
<i>Yann Shiou Ong, Marcela Borge</i>	
How Socio-Cognitive Information Affects Individual Study Decisions	274
<i>Lenka Schnaubert, Daniel Bodemer</i>	
Development of Disciplined Interpretations Using Computational Modeling in the Elementary Science Classroom.....	282
<i>Amy Voss Farris, Amanda Catherine Dickes, Pratim Sengupta</i>	
Mobilities of Criticality: Space-Making, Identity and Agency in a Youth-Centered Makerspace	290
<i>Angela Calabrese Barton, Edna Tan, Myunghwan Shin</i>	
Supporting Teachers in Navigating Change Towards Science Practices Focus in the Classroom: Investigating Current Teacher Support for Science Practices	298
<i>Nicole D. Martin, Sadhana Puntambekar</i>	
How a 6th Grade Classroom Develops Epistemologies for Building Scientific Knowledge	306
<i>Christina Krist</i>	
Metaphors Are Projected Constraints on Action: An Ecological Dynamics View on Learning Across the Disciplines	314
<i>Dor Abrahamson, Raúl Sánchez-García, Cliff Smyth</i>	
Examining Tensions Among Youth, Adults, and Curriculum as Co-Designers in 4-H STEM Learning Through Design Programs.....	322
<i>Steven M. Worker, Cynthia Carter Ching</i>	
Expanding Outcomes: Exploring Varied Forms of Teacher Learning in an Online Professional Development Experience.....	330
<i>Maxwell Yurkofsky, Sarah Blum-Smith, Karen Brennan</i>	
Investigating Effects of Embedding Collaboration in an Intelligent Tutoring System for Elementary School Students	338
<i>Jennifer K. Olsen, Nikol Rummel, Vincent Alevan</i>	
Integrating Science and Writing in Multimedia Science Fictions: Investigating Student Interactions in Role-taking	346
<i>Shiyan Jiang, Ji Shen, Blaine Elizabeth Smith</i>	
Process and Output: Relation Between Transactivity, Temporal Synchronicity, and Quality of Group Work During CSCL.....	354
<i>Vitaliy Popov, Anouschka van Leeuwen, Stan C. A. Buis</i>	
The Learning Experiences of Youth Online Information Brokers.....	362
<i>Jason C. Yip, Carmen Gonzalez, Vikki Katz</i>	
How Do We Assess Equity in Programming Pairs?.....	370
<i>Elise Deitrick, R. Benjamin Shapiro, Brian Gravel</i>	
The Effect of Concrete Materials on Children’s Subsequent Numerical Explanations: Metaphorical Priming	378
<i>Andrew Manches, Mihaela Dragomir</i>	

Expanding Coordination Class Theory to Capture Conceptual Learning in a Classroom Environment	386
<i>Lauren A. Barth-Cohen, Michael C. Wittmann</i>	
Examining the Impacts of Annotation and Automated Guidance on Essay Revision and Science Learning.....	394
<i>Libby Gerard, Marcia C. Linn, Jacqueline Madhok</i>	
Understanding Middle School Teachers' Processing of Student-Generated Resources in Science Classrooms	402
<i>Xinyang Liu, Shu Chen, Yinglian Jin, Xinning Pei, Tainian Zheng, Sihan Xiao</i>	
Analyzing Patterns of Emerging Understanding and Misunderstanding in Collaborative Science Learning: A Method for Unpacking Critical Turning Points.....	410
<i>Alisha Portolese, Lina Markauskaite, Polly K. Lai, Michael J. Jacobson</i>	
Hacking a Path In and Through STEM: Unpacking the STEM Identity Work of Historically Underrepresented Youth in STEM	418
<i>Edna Tan, Angela Calabrese Barton</i>	
Local Versus Global Connection Making in Discourse.....	426
<i>Wesley Collier, Andrew R. Ruis, David Williamson Shaffer</i>	
Advancing Learning Visualizations: Situated Action Networks as Scalable Representations of Learning in Social Settings.....	434
<i>Alejandro Andrade, Rafi Santo</i>	
Fighting for Desired Versions of a Future Self: Young African American Women's STEM-related Identity Negotiations in High School	442
<i>Carrie D. Allen, Margaret Eisenhart</i>	
Blending Play and Inquiry in Augmented Reality: A Comparison of Playing a Video Game to Playing Within a Participatory Model	450
<i>David DeLiema, Asmalina Saleh, Christine Lee, Noel Enyedy, Joshua Danish, Randy Illum, Maggie Dahn, Megan Humburg, Charlie Mahoney</i>	
Qualitative and Quantitative Information in Cognitive Group Awareness Tools: Impact on Collaborative Learning	458
<i>Melanie Erkens, Patrick Schlottbom, Daniel Bodemer</i>	
Exposing Piaget's Scheme: Empirical Evidence for the Ontogenesis of Coordination in Learning a Mathematical Concept.....	466
<i>Dor Abrahamson, Shakila Shayan, Arthur Bakker, Marieke F. van der Schaaf</i>	
Teacher Epistemic Learning in the Innovation Diffusion.....	474
<i>Longkai Wu, David Hung</i>	
Disruptions to Practice: Understanding Suspensions of Youths' Interest-related Activities	482
<i>Katie Van Horne, Erica Van Steenis, William R. Penuel, Daniela DiGiacomo</i>	
Toward an Integrated Framework of Scientific Reasoning	490
<i>Shiyu Liu</i>	
Perceptions of Productive Failure in Design Projects: High School Students' Challenges in Making Electronic Textiles.....	498
<i>Breanne K. Litts, Yasmin B. Kafai, Kristin A. Searle, Emily Dieckmeyer</i>	
The Effects of Coaching on the Teaching and Learning of English in Indian Government Schools	506
<i>Harini Raval, Chittaranjan Kaul, Susan McKenney</i>	
Listening Versus Looking: Learning About Dynamic Complex Systems in Science Among Blind and Sighted Students.....	514
<i>Ran Peleg, Sharona T Levy, Orly Lahav, Noha Chagab, Vadim Talis</i>	

Detecting Collaborative Dynamics Using Mobile Eye-Trackers.....	522
<i>Bertrand Schneider, Kshitij Sharma, Sebastien Cuendet, Guillaume Zufferey, Pierre Dillenbourg, Roy Pea</i>	
Communication Patterns and Their Role for Conceptual Knowledge Acquisition From Productive Failure.....	530
<i>Christian Hartmann, Nikol Rummel, Katharina Loibl</i>	
Competency-Based Digital Badges and Credentials: Cautions and Potential Solutions From the Field.....	538
<i>Daniel T. Hickey</i>	
Scaling Up Productive Disciplinary Engagement With Participatory Learning and Assessment	546
<i>Daniel T. Hickey, Suraj Uttamchandani, Joshua D. Quick</i>	
Providing Adaptive Scaffolds and Measuring Their Effectiveness in Open Ended Learning Environments.....	554
<i>Satabdi Basu, Gautam Biswas</i>	
Negotiation Towards Intersubjectivity and Impacts on Conceptual Outcomes	562
<i>Catherine L. Dornfeld, Sadhana Puntambekar</i>	
Supporting Elementary Students' Science Learning Through Data Modeling and Interactive Mapping in Local Spaces.....	570
<i>Kathryn A. Lanouette, Sarah Van Wart, Tapan S. Parikh</i>	
The Epistemology of Science and the Epistemology of Science Teaching	578
<i>Carl Bereiter</i>	
Embodied Cognition in Observational Amateur Astronomy.....	585
<i>Flávio S. Azevedo, Michele J. Mann</i>	
Designing a Data-Centered Approach to Inquiry Practices With Virtual Models of Density	591
<i>Jonathan M. Vitale, Jacqueline Madhok, Marcia C. Linn</i>	
Fostering University Freshmen's Mathematical Argumentation Skills With Collaboration Scripts	599
<i>Freydis Vogel, Ingo Kollar, Stefan Ufer, Kristina Reiss, Frank Fischer</i>	
Coordinating Collaborative Chat in Massive Open Online Courses	607
<i>Gaurav Singh Tomar, Sreecharan Sankaranarayanan, Xu Wang, Carolyn P. Rosé</i>	
Community Knowledge, Collective Responsibility: The Emergence of Rotating Leadership in Three Knowledge Building Communities	615
<i>Leanne Ma, Yoshiaki Matsuzawa, Bodong Chen, Marlene Scardamalia</i>	
Stories as Prototypes for Interdisciplinary Learning.....	623
<i>Vanessa Svihla, Richard Reeve</i>	
In Search of Conversational Grain Size: Modeling Semantic Structure Using Moving Stanza Windows	631
<i>Amanda L. Siebert-Evenstone, Golnaz Arastoopour, Wesley Collier, Zachari Swiecki, Andrew R. Ruis, David Williamson Shaffer</i>	
Realizing Research-Practice Connections: Three Cases From the Learning Sciences	639
<i>Susan McKenney, Jan van Aalst, Cory Forbes</i>	
Design Collaborative Formative Assessment for Sustained Knowledge Building Using Idea Thread Mapper	647
<i>Jingping Chen, Jianwei Zhang</i>	
"I Think We Were Pretty Powerful This Summer as Scientists": Generating New Possibilities for Youth of Color in Science	655
<i>Tammie Visintainer</i>	

“I’m Not Just a Mom”: Parents Developing Multiple Roles in Creative Computing.....	663
<i>Ricarose Roque, Karina Lin, Richard Liuzzi</i>	
Scaffolding Into Ambitious Teaching: Representations of Practice in Teacher Workgroups	671
<i>Jason Brasel, Brette Garner, Ilana Seidel Horn</i>	
Self-Directed Learning in Science Education: Explicating the Enabling Factors.....	679
<i>Khe Foon Hew, Nancy Law, Jarrad Wan, Yeung Lee, Amy Kwok</i>	
The Obj–Subj Dialectic and the Co–Construction of Hierarchical Positional Identities During a Collaborative Generalization Task	687
<i>José Francisco Gutiérrez</i>	
Designing a Blended, Middle School Computer Science Course for Deeper Learning: A Design-Based Research Approach.....	695
<i>Shuchi Grover, Roy Pea</i>	

Volume 2

Full Papers (continued)

Bringing Computational Thinking Into High School Mathematics and Science Classrooms	705
<i>Kai Orton, David Weintrop, Elham Beheshti, Michael Horn, Kemi Jona, Uri Wilensky</i>	
Making Sense of Making Waves: Co-constructing Knowledge and Group Understanding Without Conceptual Convergence	713
<i>Lisa Hardy, Tobin White</i>	
Students Using Graphs to Understand the Process of Cancer Treatment	721
<i>Irina Uk, Camillia Matuk, Marcia C. Linn</i>	
Developing a Geography Game for Singapore Classrooms	729
<i>Matthew Gaydos</i>	
Secondary Teachers’ Emergent Understanding of Teaching Science Practices	737
<i>William A. Sandoval, Jarod Kawasaki, Nathan Cournoyer, Lilia Rodriguez</i>	
Students’ Use of Knowledge Resources in Environmental Interaction on an Outdoor Learning Trail	745
<i>Esther Tan, Hyo-Jeong So</i>	
It Ain't What You Do, It's The Way That You Do It: Investigating the Effect of Students' Active and Constructive Interactions With Fractions Representations	753
<i>Claudia Mazziotti, Alice Hansen, Beate Grawemeyer</i>	
What Happens to the Innovation When Project Funding Ends? Learning Architecture Matters!	761
<i>Nancy Law, Leming Liang, Yeung Lee</i>	

Short Papers

Technology-Supported Dialog as a Bridge to Developing Individual Argumentive Thinking and Writing	771
<i>Deanna Kuhn, Wendy Goh</i>	
The Influence of Question Wording on Children’s Tendencies to Provide Teleological Explanations for Natural Phenomena	775
<i>Jonathan Halls, Shaaron Ainsworth, Mary Oliver</i>	
“That’s Your Heart!”: Live Physiological Sensing and Visualization Tools for Life-Relevant and Collaborative STEM Learning	779
<i>Leyla Norooz, Tamara L. Clegg, Seokbin Kang, Angelisa C. Plane, Vanessa Oguamanam, Jon E. Froehlich</i>	

Managing Threats to Teacher Face in Discussions of Video-Recorded Lessons	783
<i>Dana Vedder-Weiss, Aliza Segal, Adam Lefstein</i>	
Fostering More Informed Epistemic Views Among Students Through Knowledge Building	787
<i>Huang-Yao Hong, Bodong Chen, Chin-Chung Tsai, Chiu Pin Lin, Ying-Tien Wu</i>	
Crossing Boundaries: Reflexive Analysis of Collaborative Learning in Research Institutions.....	791
<i>Kevin Crouse, Jeanette Joyce, Veronica L. Cavera</i>	
Using Differentiated Feedback Messages to Promote Student Learning in an Introductory Statistics Course.....	795
<i>Qijie Cai, Han Wu, Bodong Chen</i>	
Use of Interactive Computer Models to Promote Integration of Science Concepts Through the Engineering Design Process.....	799
<i>Elizabeth A. McBride, Jonathan M. Vitale, Lauren Applebaum, Marcia C.Linn</i>	
Examining the Influences of Teacher's Framing of Modeling Practices on Elementary Students' Engagement in Modeling	803
<i>Li Ke, Christina V. Schwarz</i>	
Learning the Learning Sciences: An Investigation of Newcomers' Sociocultural Ideas.....	807
<i>Yotam Hod, Ornit Sagy,</i>	
Mathematical Argumentation and Proof – Supporting a Complex Cognitive Skill.....	811
<i>Daniel Sommerhoff, Stefan Ufer, Ingo Kollar</i>	
Modelling Authenticity in Teaching and Learning Contexts: A Contribution to Theory Development and Empirical Investigation of the Construct.....	815
<i>Anica Betz, Sabrina Flake, Marcel Mierwald, Marie Vanderbeke</i>	
Fostering Collaborative Learning Through Knowledge Building Among Students With Low Academic Achievement	819
<i>Yuqin Yang, Jan van Aalst, Carol K. K. Chan</i>	
Supporting Planning and Conducting Experiments.....	823
<i>Siswa A. N. van Riesen, Hannie Gijlers, Anjo Anjewierden, Ton de Jong</i>	
Designing Technology for Learning: How to Get from Disenfranchisement to Disinheritance and Why We Need to Go There	827
<i>Zaza Kabayadondo</i>	
Learning Environments to Facilitate Students' Regulation in Knowledge Building	831
<i>Jin Michael Splichal, Jun Oshima, Ritsuko Oshima</i>	
Resolving Disagreements in Evaluating Epistemic and Disciplinary Claims in Middle School Science.....	835
<i>Sihan Xiao, William A. Sandoval</i>	
Tracking Student Teachers' Technology-Enhanced Collaborative Problem Solving: Combining Objective Assessment Data With Subjective Verbal Reporting	839
<i>Johanna Pöysä-Tarhonen, Esther Care, Nafisa Awwal, Päivi Häkkinen, Arto K. Ahonen</i>	
Conceptualizing Authenticity and Relevance of Science Education in Interactional Terms.....	843
<i>Shulamit Kapon, Antti Laherto, Olivia Levrini</i>	
Developing Argument Skills Through the SOCRATES Learning Environment.....	847
<i>Kalypto Jordanou, Wendy Goh</i>	
Environmental Learning Through the Lens of Affinity Spaces: Transforming Community Members Into a Community Force.....	851
<i>Tamara Clegg, Jennifer Preece, Daniel Pauw, Elizabeth Warrick, Carol Boston</i>	

Students' Responses to Curricular Activities as Indicator of Coherence in Project-Based Science	855
<i>William R. Penuel, Katie Van Horne, Sam Severance, David Quigley, Tamara Sumner</i>	
A Qualitative Exploration of Self- and Socially Shared Regulation in Online Collaborative Learning.....	859
<i>Lauren Hensley, Jessica Cutshall, Victor Law, Kui Xie, Lin Lu</i>	
Exploring the Composition Process of Peer Feedback	862
<i>Maryam Alqassab, Jan-Willem Strijbos, Stefan Ufer</i>	
Let Your Data Tell a Story: Disciplinary Expert Feedback Locates Engaging in Argumentation in a Holistic System of Practices	866
<i>Elizabeth M. Walsh, Veronica C. McGowan</i>	
A Design Approach to Understanding the Activity of Learners in Interdisciplinary Settings: Environment and Diversity	870
<i>Kate Thompson, Julia Svoboda Gouvea, Geoffrey Habron</i>	
Teamwork in the Balance: Exploratory Findings of Teamwork Competency Patterns in Effective Learning Teams.....	874
<i>Elizabeth Koh, Antonette Shibani, Helen Hong</i>	
Problems With Different Interests of Learners in an Informal CSCL Setting	878
<i>Yong Ju Jung, Shulong Yan, Marcela Borge</i>	
Educational Affordances of Tablet-Mediated Collaboration to Support Distributed Leadership in Small Group Outdoor Activities	882
<i>Gi Woong Choi, Susan M. Land, Heather Toomey Zimmerman</i>	
Designing Science Curriculum for Implementation at Scale: Considerations for Diverse and Resource-Limited Settings.....	886
<i>Debra Bernstein, Brian Drayton, Susan McKenney, Christian Schunn</i>	
Interdisciplinary Computing and the Emergence of Boundary Objects: A Case-Study of Dance and Technology	890
<i>Kayla DesPortes, Monet Spells, Betsy DiSalvo</i>	
Unpacking Social Factors in Mechanistic Reasoning (Or, Why a Wealthy Person is Not Exactly Like a Grey Squirrel).....	894
<i>Arthur Hjorth, Christina Krist</i>	
Conceptual Fluency: Switching Between Pre- and Post-Threshold Assumptions of Molecular Dynamics.....	898
<i>Prajakt Pande, Hannah Sevian</i>	
Let Kids Solve Wicked Problems... Why Not?!	902
<i>Rachel Lam, Michelle Low</i>	
Effects of Implicit Guidance on Contribution Quality in a Wiki-Based Learning Environment.....	906
<i>Sven Heimbuch, Daniel Bodemer</i>	
Learning to Argue: The Role of Peer Assessment.....	910
<i>Shiyu Liu</i>	
Newcomer Integration in Online Knowledge Communities: Exploring the Role of Dialogic Textual Complexity	914
<i>Nicolae Nistor, Mihai Dascălu, Ștefan Trăușan-Matu</i>	
Is Small Group Collaboration Beneficial in Large Scale Online Courses? An Investigation of Factors Influencing Satisfaction and Performance in GroupMOOCs	918
<i>Elias Kyewski, Nicole C. Krämer, Nina Christmann, Malte Elson, Julia Erdmann, Tobias Hecking, Thomas Herrmann, H. Ulrich Hoppe, Nikol Rummel, Astrid Wichmann</i>	

Embodied Search Processes in Creative Problem Solving: How Do People Learn in Makerspaces?	920
<i>Michael Tan</i>	
Comparing Students' Solutions When Learning Collaboratively or Individually Within Productive Failure.....	926
<i>Claudia Mazziotti, Nikol Rummel, Anne Deiglmayr</i>	
Student Conceptions of Expertise.....	930
<i>Charles Bertram, Anne Leak, Eleanor C. Sayre, Mary Bridget Kustus, Scott V. Franklin</i>	
Distributions, Trends, and Contradictions: A Case Study in Sensemaking With Interactive Data Visualizations	934
<i>Vasiliki Laina, Michelle Wilkerson</i>	
A User Interface for the Exploration of Manually and Automatically Coded Scientific Reasoning and Argumentation	938
<i>Patrick Lerner, Andras Csanadi, Johannes Daxenberger, Lucie Flekova, Christian Ghanem, Ingo Kollar, Frank Fischer, Iryna Gurevych</i>	
Designing Outdoor Learning Spaces With iBeacons: Combining Place-Based Learning With the Internet of Learning Things	942
<i>Heather Toomey Zimmerman, Susan M. Land, Chrystal Maggiore, Robert W. Ashley, Chris Millet</i>	
Agentic Trajectories: Development and Learning in a Project-Based High School for Marginalized Students	946
<i>Vanessa Svihla, Liza Kittinger</i>	
Exploring Middle School Students' Science Learning and Discourse in Physical and Virtual Labs	950
<i>Dana Gnesdilow, Nafsaniath Fathema, Feng Lin, Seokmin Kang, Catherine L. Dornfeld, Sadhana Puntambekar</i>	
Iterative Curricular Design of Collaborative Infographics for Science Literacy in Informal Learning Spaces.....	954
<i>Stephen Sommer, Cynthia Graville-Smith, Joseph Polman, Leighanna Hinojosa</i>	
Exploring African-American Middle-School Girls' Perceptions of Themselves as Game Designers	960
<i>Jakita O. Thomas, Rachelle Minor, O. Carlette Odenwingie</i>	
Mauá Project: Citizenship and Environment Educational as Pathway to Critical Thinking and Students' Empowerment	962
<i>Glauco S. F. da Silva, Marcília E. Barcellos, Elisabeth G. de Souza</i>	
Exploring Visualization and Tagging to Manage Big Datasets for DBR: A Modest Proposal With Significant Implications.....	966
<i>Kelly J. Barber-Lester, Sharon Derry, Lana M. Minshew, Janice Anderson</i>	
What Are Crosscutting Concepts in Science? Four Metaphorical Perspectives	970
<i>Ann E. Rivet, Gary Weiser, Xiaoxin Lyu, Yi Li, Diego Rojas-Perilla</i>	
Situating Deep <i>Multimodal</i> Data on Game-Based STEM Learning.....	974
<i>Craig G. Anderson, John V. Binzak, Jennifer Dalsen, Jenny Saucerman, Anna Jordan-Douglass, Vishesh Kumar, Aybuke Turker, Matthew Berland, Kurt Squire, Constance Steinkuehler</i>	
Designing for Effective Collaborative Learning in High-Needs Rural Classrooms	978
<i>Lana M. Minshew, Sharon Derry, Janice Anderson, Kelly Barber-Lester</i>	
Supporting Calculus Learning Through “Smooth” Covariation.....	982
<i>Susanne M. Strachota</i>	
Teacher Noticing Associated With Responsive Support of Knowledge Building.....	986
<i>Darlene Judson</i>	

The Study of Cognitive Development in the Structured Collaborative Learning Task Mediated by Semantic Diagram Tools	990
<i>Huiying Cai, Xiaoqing Gu</i>	
Uncovering Teachers' Pedagogical Reasoning in Science Discussions	994
<i>Sherice Clarke, David Gerritsen, Rebecca Grainger, Amy Ogan</i>	
Investigating Analogical Problem Posing as the Generative Task in the <i>Productive Failure</i> Design	998
<i>Jun Song Huang, Rachel Lam, Manu Kapur</i>	
The Effect of Scaffolding on the Immediate Transfer of Students' Data Interpretation Skills Within Science Topics	1002
<i>Raha Moussavi, Janice Gobert, Michael Sao Pedro</i>	
Exploring the Relationship Between Gesture and Student Reasoning Regarding Linear and Exponential Growth.....	1006
<i>Sahar Alameh, Jason Morpew, Nitasha Mathayas, Robb Lindgren</i>	
Sketching a Pathway Through Design Worlds: Multimodal Communication in a Fifth-Grade Collaborative Engineering Project.....	1010
<i>Michelle E. Jordan, Jamie M. Collins</i>	
“Show Me” What You Mean: Learning and Design Implications of Eliciting Gesture in Student Explanations	1014
<i>Robb Lindgren, Robert C. Wallon, David E. Brown, Nitasha Mathayas, Nathan Kimball</i>	
The Function of Epistemic Emotions for Complex Reasoning in Mathematics	1018
<i>Sandra Becker, Reinhard Pekrun, Stefan Ufer, Elisabeth Meier</i>	
Symposia	
FUSE: An Alternative Infrastructure for Empowering Learners in Schools	1025
<i>Reed Stevens, Kemi Jona, Lauren Penney, Dionne N. Champion, Kay E. Ramey, Jaakko Hilppö, Ruben Echevarria, William R. Penuel</i>	
Qualitative Analysis of Video Data: Standards and Heuristics	1033
<i>Kay E. Ramey, Dionne N. Champion, Elizabeth B. Dyer, Danielle T. Keifert, Christina Krist, Peter Meyerhoff, Krystal Villanosa, Jaakko Hilppö</i>	
Connected Making: Designing for Youth Learning in Online Maker Communities In and Out of Schools	1041
<i>Breanne K. Litts, Yasmin B. Kafai, Deborah A. Fields, Erica R. Halverson, Kylie Peppler, Anna Keune, Mike Tissenbaum, Sara M. Grimes, Stephanie Chang, Lisa Regalla, Orkan Telhan, Michael Tan</i>	
Agentive Learning for Sustainability and Equity: Communities, Cooperatives and Social Movements as Emerging Foci of the Learning Sciences	1048
<i>Yrjö Engeström, Annalisa Sannino, Aydin Bal, Heila Lotz-Sisitka, Tichaona Pesanayi, Charles Chikunda, Manoel Flores Lesama, Antonio Carlos Picinatto, Marco Pereira Querol, Yew Jin Lee</i>	
Negotiating Academic Communities: Narratives of Life-long Learners	1055
<i>Sally Fincher, Sebastian Dziallas, Ofra Brandes, Yifat Ben-David Kolikant, R. Benjamin Shapiro</i>	
Future Learning Spaces for Learning Communities: New Directions and Conceptual Frameworks	1063
<i>Yotam Hod, Elizabeth S. Charles, Alisa Acosta, Dani Ben-Zvi, Mei-Hwa Chen, Koun Choi, Michael Dugdale, Yael Kali, Kevin Lenton, Scott P. McDonald, Tom Moher, Rebecca M. Quintana, Michael M. Rook, James D. Slotta, Phil Tietjen, Patrice L. Tamar Weiss, Chris Whittaker, Jianwei Zhang, Katerine Bielaczyc, Manu Kapur</i>	
Beyond Just Getting Our Word Out: Creating Pipelines From Learning Sciences Research to Educational Practices	1071
<i>Michael J. Jacobson, Kristine Lund, Christopher Hoadley, Ravi Vatrapu, Janet L. Kolodner, Peter Reimann</i>	

Building on Cultural Capacity for Innovation Through International Collaboration: In Memory of Naomi Miyake.....	1074
<i>Hajime Shirouzu, Marlene Scardamalia, Moegi Saito, Sonoko Ogawa, Shinya Iikubo, Naoto Hori, Carolyn Rosé</i>	
Exploring the Value of Drawing in Learning and Assessment.....	1082
<i>Shaaron Ainsworth, Mike Stieff, Dane DeSutter, Russell Tytler, Vaughan Prain, Dimitrios Panagiotopoulos, Peter Wigmore, Wouter van Joolingen, Dewi Heijnes, Frank Leenaars, Sadhana Puntambekar</i>	
Designing Learning Contexts Using Student-Generated Ideas.....	1090
<i>Rachel Lam, Lung-Hsiang Wong, Matthew Gaydos, Jun Song Huang, Manu Kapur, Lay Hoon Seah, Michael Tan, Katerine Bielaczyc, William A. Sandoval</i>	
Moving Ahead in the Study of STEM Interests and Interest Development: A New Research Agenda	1098
<i>Flavio S. Azevedo, June Ahn, Michele J. Mann, Rena Dorph, Matthew A. Cannady, Victor R. Lee, Ryan Cain, Philip Bell</i>	
Fostering Deliberative Discourse in Schools Towards the Constitution of a Deliberative Democracy.....	1106
<i>Baruch Schwarz, Antti Rajala, Carolyn P. Rosé, Sherice Clarke, Elizabeth Fynes-Clinton, Peter Renshaw, Anne Solli, Thomas Hillman, Åsa Mäkitalo, Tsafirir Goldberg, Tuure Tammi, Rupert Wegerif</i>	
The Learning Sciences @ Scale: Current Developments in Open Online Learning	1114
<i>James D. Slotta, Daniel Hickey, Carolyn P. Rosé, Pierre Dillenbourg, Hedieh Najafi, Stian Håkleiv, Suraj Uttamchandani, Joshua D. Quick</i>	
Real-Time Visualization of Student Activities to Support Classroom Orchestration.....	1120
<i>Mike Tissenbaum, Camillia Matuk, Matthew Berland, Leila Lyons, Felipe Cocco, Marcia C. Linn, Jan L. Plass, Nik Hajny, Al Olsen, Beat Schwendimann, Mina Shirvani Boroujeni, James D. Slotta, Jonathan M. Vitale, Libby Gerard, Pierre Dillenbourg</i>	
Researchers and Practitioners Co-Designing for Expansive Science Learning and Educational Equity.....	1128
<i>Philip Bell, Samuel Severance, William R. Penuel, Tamara Sumner, Wagma Mommandi, David Quigley, Katie Van Horne, Raymond Johnson, Shelley Stromholt, Heena Lakhani, Katie Davis, Adam Bell, Megan Bang</i>	
Teachers and Professional Development: New Contexts, Modes, and Concerns in the Age of Social Media.....	1136
<i>Christine Greenhow, Arnon HersHKovitz, Alona Forkosh Baruch, Emilia Askari, Dimitra Tsovaltzi, Christa Asterhan, Thomas Puhl, Armin Weinberger, Edith Bouton, Joseph Polman</i>	
Posters	
Enjoyment and Satisfaction as Details for Kinetic Learning.....	1147
<i>WeiQuan Lu, Mandi Jieying Lee, Chun Kit Lee, Linh-Chi Nguyen, Ellen Yi-Luen Do</i>	
Reforming the Undergraduate STEM Classroom Experience	1149
<i>Mike Stieff, Alison Castro Superfine</i>	
Design of Automated Guidance to Support Effortful Revisions of Science Essays	1152
<i>Charissa Tansomboon, Libby F. Gerard, Marcia C. Linn</i>	
Engineering Real-time Indicators for Targeted Supports: Using Online Platform Data to Accelerate Student Learning.....	1153
<i>HiroYuki Yamada, Ouajdi Manai, Christopher Thorn</i>	
A Design Research to Support Elementary Students' Epistemic Understanding of Their Scientific Argument Construction	1155
<i>Miki Sakamoto, Etsuji Yamaguchi</i>	

Journalistic Sources: Evaluation in Third Space	1157
<i>Gulnaz Saiyed, Wan Shun Eva Lam, Matthew Easterday</i>	
Using Video Recording Glasses to Get a First Person Perspective	1159
<i>Nina Bonderup Dohn, Niels Bonderup Dohn</i>	
Designing Side-By-Side and In-The-Moment: On the Participation of Youth in Professional Learning Settings.....	1161
<i>Meixi</i>	
The AFS Educational Approach in Developing and Assessing Intercultural Competence	1163
<i>Jason Wen Yau Lee, Melissa Liles, Hazar Yildirim, Frances Baxter</i>	
Improving Self-Detection of Confusion: Is Metacognitive Monitoring a Key?	1165
<i>Mariya Pachman, Lori Lockyer</i>	
How Teachers Can Boost Conceptual Understanding in Physics Classes.....	1167
<i>Ralph Schumacher, Sarah Hofer, Herbert Rubin, Elsbeth Stern</i>	
How to Support Senior Citizens' Media Literacies: A Review of Existing Research Literature	1169
<i>Päivi Rasi, Pirkko Hyvönen</i>	
Learning With Peers: Problem Solving Through Pair Collaboration	1171
<i>Allison Ritchie</i>	
Sequencing Physical Representations With Human Tutors and Virtual Representations With a Computer Tutor in Chemistry.....	1173
<i>Martina A. Rau, Sally P. Wu, Jamie Schuberth</i>	
From Belief Mode to Design Mode: Report on a Chinese 6th Graders' Interdiscipline on Traffic Control.....	1175
<i>Yibing Zhang, Yao Liu</i>	
Organizing to Cultivate Personal Relevance, Science Literacy, and Equity Through Data Journalism.....	1177
<i>Joanna Weidler-Lewis, Leighanna Hinojosa, Stephen Sommer, Joseph L. Polman</i>	
Developing a Language-neutral Instrument to Assess Fifth Graders' Computational Thinking	1179
<i>Ji Shen, Guanhua Chen, Lauren A. Barth-Cohen, Shiyan Jiang, Moataz Eltoukhy</i>	
New Creativity Examined With E-Textiles: Bridging Arts Craft and Programming	1181
<i>Hyunshin Choi, Juyeon Park</i>	
A Cross-Cultural Study of the Effect of a Graph-Oriented Computer-Assisted Project-Based Learning Environment on Students' Argumentation Skill.....	1183
<i>Pi-Sui Hsu, Margot Van Dyke, Eric Monsu Lee, Thomas J. Smith</i>	
Becoming a Teacher Who Takes an Inquiry Stance	1185
<i>Nastasia Lawton-Sticklor, Katerine Bielaczyc</i>	
Creating a Safe Learning Environment for Peer Assessment: Exploring Students' Conceptions Towards Fading Anonymity Over Time.....	1187
<i>Tijs Rotsaert, Tammy Schellens</i>	
The Lonesome Penguin: Voice Created Through Language, Identity, and Engagement in Climate Science	1189
<i>Kristen Dominguez, Elizabeth M. Walsh</i>	
Refinement of Semantic Network Analysis for Epistemic Agency in Collaboration.....	1191
<i>Jun Oshima, Ritsuko Oshima, Wataru Fujita</i>	
Designing Extensible and Flexible Augmented Mobile Learning Digital Lessons	1193
<i>Neven Drljević, Mirna Domančić, Ivica Botički, Manuela Kajkara</i>	

Knowledge Building and Conversational Functions in Online Interactions: A Coding Scheme.....	1195
<i>Stefano Cacciamani, Vittore Perrucci, Ahmad Khanlari</i>	
Robotics Programming in Support of Computational Thinking: Scaffolding From Teacher, Peer, and Robotics Agents.....	1197
<i>Bian Wu</i>	
Developing Literary Reasoning Practices in Class Discussions	1199
<i>Allison H. Hall</i>	
Using a Systematic Review for Cross-Theory Comparisons.....	1201
<i>Medha Dalal, Ruth Wylie, Erin Walker</i>	
Effects of Knowledge Building Pedagogy on High School Students' Learning Perception and English Composition Performance	1203
<i>I-Ting Yang, Hsien-Ta Lin, Ching-Hua Wang, Pei-Jung Li, Huang-Yao Hong</i>	
Theorizing Learning Spaces: A Sociomaterial Approach.....	1205
<i>Phil Tietjen, Scott McDonald, Michael M. Rook</i>	
Multi-Layered Online Workshop: Promoting Both Collaborative and Instructional Interactions at Medium Scale	1207
<i>Yuki Anzai, Hiroki Oura, Ryohei Ikejiri, Wakako Fushikida, Yuhei Yamauchi</i>	
From Knowing to Doing: Increasing the Confidence of Communities to Take Action and Save Lives.....	1209
<i>Grant Chartrand</i>	
Using Learning Performances to Design Science Assessments That Measure Knowledge-in-Use	1211
<i>Kevin W. McElhaney, Gauri Vaishampayan, Cynthia M. D'Angelo, Christopher J. Harris, James W. Pellegrino, Joseph S. Krajcik</i>	
Effects of Knowledge Building Activities on Student' Writing Performance.....	1213
<i>Ching Hua Wang, Hsien-Ta Lin, Pei-Jung Li, I-Ting Yang, Huang-Yao Hong</i>	
Communicating Culture: Building a Network of Participatory Teacher Meetups to Support K-12 Computing Education.....	1215
<i>Karen Brennan, Raquel Jimenez, Wilhelmina Peragine</i>	
Do Collaborative Learning Elements in Online Courses Improve Conceptual Knowledge Acquisition?.....	1217
<i>Julia Erdmann, Nikol Rummel</i>	
“That’s What Teachers Do!” “—Not <u>This</u> Teacher”: Tensions in Reform Science Classrooms.....	1219
<i>Aliza Stein, Brian J. Reiser</i>	
Affordances and Challenges of Using K-12 Classrooms to Support Preservice Teachers’ Learning to Teach Climate Change	1221
<i>Asli Sezen-Barrie</i>	
Finding Patterns in and Refining Characterizations of Students’ Epistemic Cognition: A Computational Approach.....	1223
<i>Christina Krist, Joshua Rosenberg</i>	
Algorithmic, Conceptual, and Physical Thinking: A Framework for Understanding Student Difficulties in Quantum Mechanics.....	1225
<i>John D. Thompson, Bahar Modir, Eleanor C. Sayre</i>	
Team Awareness Support for Collaborative Argumentation in Higher Education: A Qualitative Multiple-Case Study	1227
<i>Maria Fysaraki, Frank Fischer, Heinrich Hußmann, Karsten Stegmann</i>	

Gender Differences in STEM Career and Educational Choices of Alumni of an Urban, Museum-Based After School Program.....	1229
<i>C. Aaron Price, Faith Kares, Gloria Segovia</i>	
Utilizing Eye Tracking Technology to Promote Students' Meta-Cognitive Awareness of Visual STEM Literacy.....	1231
<i>Stephen Sommer, Leighanna Hinojosa, Joseph Polman</i>	
Embodied Actions to Support Spatial Thinking in STEM: Structural Diagrams in Organic Chemistry.....	1233
<i>Dane DeSutter, Mike Stieff</i>	
How Teachers Seek and Learn From Negative Feedback.....	1235
<i>Mei Wang, Jiaming Cheng, Hyojin Kim, Xiaozhe Yang, Jian Zhao, Youqun Ren</i>	
Collaborative Knowledge Networks to Facilitate Knowledge Building in Robotics: A Longitudinal Study.....	1237
<i>Ahmad Khanlari, Marlene Scardamalia</i>	
Design Experiments Towards Practice-Based Learning Analytics: A Student Perspective.....	1239
<i>Manolis Mavrikis, Mutlu Cukurova, Nina Valkanova, Annelie Berner, Rose Luckin</i>	
Provoking Mathematical Play Through Hidden Deep Structures.....	1241
<i>Caroline Williams-Pierce</i>	
Acceptance Model of Learning Technologies in Media Commons.....	1243
<i>Gi Woong Choi, Barton K. Pursel</i>	
Obtaining Rich Data in Augmented Reality Settings: A Comparison of Three Data Collection Approaches.....	1245
<i>Eleni A. Kyza, Yiannis Georgiou, Markos Souropetis, Andria Agesilaou</i>	
Learning Progressions as Tools for Classroom Practice.....	1247
<i>Vanessa de León, Erin Marie Furtak, Deb Morrison, Rebecca Swanson, Kathy Kiemer</i>	
A Tri-Level Partnership to Support and Spread Knowledge Building in Ontario.....	1249
<i>Monica Resendes, Marlene Scardamalia, Mary Cordeiro, Linda Massey, Karen Dobbie, Lindsay Sirois</i>	
Building Sustained, Knowledge Creating Networks in Education.....	1251
<i>Monica Resendes, Marlene Scardamalia, Mary Cordeiro, Linda Massey, Karen Dobbie, Lindsay Sirois</i>	
Structuring Authentic Open Inquiry in an Undergraduate Science Lab Course as an Epistemic Onramp to Professional Physics.....	1253
<i>Nicholas C. Wilson, Vera Michalchik</i>	
From Classroom Interaction to Clinical Reasoning: An Interactional Ethnography of PBL in Speech and Hearing Sciences.....	1255
<i>Susan Bridges, Anita MY Wong, Cindy Hmelo-Silver, Carol Chan, Judith L. Green</i>	
The Redesign of an Extensive Learning Environment for Medical Education.....	1257
<i>Timothy Charoenying, Dragan Trninic</i>	
How do Learners With Different Epistemic Beliefs and Needs for Closure Approach Instructor's Feedback to Project?.....	1259
<i>Kun Huang, Victor Law, Xun Ge</i>	
Augmenting Learning From Physical Museum Exhibits With Personal Mobile Technology.....	1261
<i>Kher Hui Ng, Hai Huang, Shanker Selvamurthy, Masharrat Juzar, Nurul Assyifa Ahmad Sabri, Claire O'Malley</i>	

Balancing Expression and Structure in Game Design: Computational Participation Using Studio-Based Pedagogy	1263
<i>Benjamin DeVane</i>	
Exploring the Development of Scientific Argumentation Practices Among First-Year STEM Undergraduates Through a Writing-to-Learn Approach	1265
<i>Margaret M. Lucero, Patricia Serviss</i>	
Exploring Middle School Students' Sense Making of a Computer Simulation About Thermal Conduction	1267
<i>Nitasha Mathayas, David E. Brown, Robb Lindgren</i>	
Colors of Nature: Exploring Middle School Girls' Notions of Creativity in an Art/Science Academy	1269
<i>Blakely K. Tsurusaki, Carrie Tzou, Laura Carsten Conner, Mareca Guthrie</i>	
Developing Gesture Recognition Capabilities for Interactive Learning Systems: Personalizing the Learning Experience With Advanced Algorithms.....	1271
<i>Michael Junokas, Nicholas Linares, Robb Lindgren</i>	
Collective Regulation of Idea Improvement in Knowledge Building Discourse.....	1273
<i>Chunlin Lei, Carol K. K. Chan</i>	
Impact of Theory Improvement and Collective Responsibility for Knowledge Advancement on the Nature of Student Questions	1275
<i>Ahmad Khanlari, Marlene Scardamalia, Derya Kici, Suresh Tharuma</i>	
Where the Rubber Meets the Road: The Impact of the Interface Design on Model Exploration in Science Inquiry	1277
<i>Engin Bumbacher, Zahid Hossain, Ingmar Riedel-Kruse, Paulo Blikstein</i>	
Practitioners' Track	
The ICLS 2016 K-12 Practitioners' Track	1281
<i>Jaime Koh, Choon Lang Quek</i>	
A Comparison of Video Production Styles in Mathematics Flipped Classroom: Examining Students' Preferences	1282
<i>Chung Kwan Lo, Khe Foon Hew</i>	
Refashioning Education as a Knowledge Creating Enterprise: Growing Capacity for Knowledge Building in a Grade 1 Math Class.....	1286
<i>Cindi Chisholm, Heather Fleming, Lizanne Lacelle, Monica Resendes, Marlene Scardamalia</i>	
Creating Knowledge Building Communities: Three Case Studies at the Classroom, School and Board Level	1290
<i>Monica Resendes, Marlene Scardamalia, Karen Dobbie, Jason Frenza, Emma Nichols, Ross Edgar, Francis Noventa, Linda Massey, Mary Cordeiro</i>	
Learning Argumentation Using Web 2.0 Tools.....	1294
<i>Susan Gwee</i>	
Designing and Implementing the Investigative Skills in Secondary Science Curriculum: A Case Study in Singapore	1298
<i>Hoe Teck Tan, Mi Song Kim</i>	
From Passive to Active Learning in A-level Mathematics Classroom	1302
<i>Puay San Chan</i>	
Pop Quizzes – A Journey to Use Formative Assessment to Grow Students and Develop Teachers	1306
<i>Ziyang Chen, Enhui Grace Tan, Chong Chieh Yap, Sian Hong Edith Ang, Wei Seng Andy Loo, Yangming Aw, Muhammad Helmi Bin Ahmad Bamadhaj, Suhailin Mohamed Sahed, Abigail Abraham, Fauziah Kani Abdul Waduth</i>	

Knowledge Building Pedagogy and Technology: Enacting Principle-Based Design in History Classroom.....	1310
<i>Melvin Chan, Chew Lee Teo, Yu Ling Lee</i>	
Nurturing Positive Learning Outcomes: The Role of the Interlocutor	1314
<i>Esther Joosa, Sumitra Pasupathy</i>	
Understanding Teacher’s Principle-Based Practice in Sustaining Knowledge Building Practice in a Science Classroom.....	1318
<i>Mohd Noor Hishamuddin Haslir, Chew Lee Teo, Shahizha Bte Mohd, Yu Ling Lee</i>	
Knowledge Building for Students With Low Academic Achievement	1322
<i>Bing-fai Lee, Carol K.K. Chan, Jan van Aalst</i>	
Developing Students’ Early Science Literacy – A Holistic Approach to Science Learning.....	1326
<i>Liyun Wong, Kim Yeow Chiam, De Qi Chen</i>	
Workshops	
Towards Next Steps for the CSCL Community: Advancing Science and Informing Real World Collaboration in Web 2.0.....	1333
<i>Ulrike Cress, Carolyn P. Rosé</i>	
Revisiting Learning Communities: Innovations in Theory and Practice	1335
<i>Yotam Hod, Dani Ben-Zvi, Katerine Bielaczyc</i>	
‘Jugaad’: Transgressions Within Research Methodologies	1338
<i>Sameer Honwad, Anne Kern, Heila Lotz-Sisitka, Shivaraj Bhattarai, Christopher Hoadley</i>	
Organizing Design-Based Implementation Research in Research-Practice Partnerships: A Workshop	1342
<i>William R. Penuel, Philip A. Bell, Alain Breuleux, Elizabeth S. Charles, Barry J. Fishman, Therese Laferrrière, Susan McKenney</i>	
Situating Multimodal Learning Analytics	1346
<i>Marcelo Worsley, Dor Abrahamson, Paulo Blikstein, Shuchi Grover, Bertrand Schneider, Mike Tissenbaum</i>	
How Students Learn in East Asian Cultures and How That Learning May Evolve in the Future.....	1350
<i>Xiaoqing Gu, Lung-Hsiang Wong, Tak-Wai Chan, Hajime Shirouzu, Heisawn Jeong, Charles Crook, Siu Cheung Kong</i>	
Embodiment and Designing Learning Environments.....	1353
<i>Robb Lindgren, Andrew Manches, Dor Abrahamson, Sara Price, Victor R. Lee, Mike Tissenbaum</i>	
Computer-Based Learning Environments for Deep Learning in Inquiry and Problem-Solving Contexts.....	1356
<i>Minhong (Maggie) Wang, Paul A. Kirschner, Susan M. Bridges</i>	
Early Career Workshop	
The ICLS 2016 Early Career Workshop.....	1361
<i>Julia Eberle, Nikol Rummel, Manu Kapur, Paul A. Kirschner</i>	
Leveraging Classroom Talk to Promote Educational Equity.....	1362
<i>Sherice N. Clarke</i>	
The Social Organization of Play, Embodied Cognition, and Failure in STEM Education.....	1364
<i>David DeLiema</i>	
Supporting Students’ Development of Integrated Knowledge Incorporating the Content and Practices of Science	1366
<i>Sarah J. Fick</i>	

An Emerging Research Agenda on Humanistic Learning Communities.....	1368
<i>Yotam Hod</i>	
Learning Through Teaching: Towards a Socio-Cultural Theory of Teaching Practice Development	1370
<i>Hee-jeong Kim</i>	
How Preparation Activities Affect the Process and Learning Outcomes of Peer Collaboration	1372
<i>Rachel Lam</i>	
Exploring How Teachers' and Students' Personal Alternative Conceptions With Scientific Concepts Influence Teaching Practice	1374
<i>Margaret M. Lucero</i>	
Learning-by-Making and Educational Equity: STEM Learning and Identity Development in a School-Based Makerspace.....	1376
<i>Antti Rajala</i>	
Social, Perceptual, and Conceptual Factors of Learning With Multiple External Representations in Educational Technologies.....	1378
<i>Martina A. Rau</i>	
An Inclusive Framework for Understanding the Role That Communities of Practice Play in New Digital Literacies, Barriers to Participation and Implications for Equity	1380
<i>Gabriela T. Richard</i>	
Designing Science Lessons With a Focus on the Demands of the Language of Science	1382
<i>Lay Hoon Seah</i>	
Advancing Surgical Education: What Does It Mean to Think Like a Surgeon?.....	1383
<i>Sarah Sullivan</i>	
Children's Social and Emotional Development in Collaborative Learning.....	1385
<i>Jingjing Sun</i>	
Flexibility and Adaptability of Collaborative Learning Scaffolds.....	1386
<i>Freydis Vogel</i>	
Learning Mathematics Through Designed Digital Experiences	1388
<i>Caroline Williams-Pierce</i>	
Doctoral Consortium	
ICLS 2016 Doctoral Consortium Workshop	1393
<i>Kristine Lund, Sadhana Puntambekar, Jan van Aalst</i>	
Fostering a Disciplinary Stance in Higher Education History Learning.....	1395
<i>Uzi Zevik Brami, Iris Tabak</i>	
Improving Conceptual Knowledge Acquisition in Online Courses by Adding Collaborative Learning Elements.....	1396
<i>Julia Erdmann</i>	
Exploring the History of Education: Designing for Transition Into Undergraduate Initial Teacher Education and Towards Professionalism.....	1397
<i>Paul Flynn</i>	
Young Children Teleological Explanations for Natural Phenomena: Assessment Methods and Pedagogical Approaches.....	1398
<i>Jonathan Halls</i>	

Computer Support for Group Learning of Physics Models	1399
<i>Lisa A. Hardy</i>	
Leveraging Teacher Noticing for Sustained Idea Improvement in Students' Knowledge Building Inquiry	1400
<i>Darlene Judson</i>	
Co-evolutionary Dynamics Between Teacher Learning and Organizational Learning in the Process of ICT-enabled Pedagogical Innovations.....	1401
<i>Leming Liang</i>	
Supporting Teachers to Develop a Holistic Conceptualization of Science Practices: A Framework to Transform Teachers' Classroom Practice and Improve Students' Understanding of Science Practices	1402
<i>Nicole D. Martin</i>	
Mobile Learning in Secondary Classes: The Use of Tablet Devices as a Learning Tool for Fostering Inquiry Learning and Self-Regulated Learning	1403
<i>Hannelore Montrieux, Tammy Schellens</i>	
Iterative Design, Development, and Evaluation of Scaffolds for Data Interpretation Practices During Inquiry	1404
<i>Raha Moussavi, Janice Gobert</i>	
Redesigning Problem-Based Learning in Medical Education: Contrasting Solutions to Improve Consolidation	1405
<i>Alisha Portolese, Michael J. Jacobson, Robbert Duvivier, Lina Markauskaite</i>	
Becoming a University Student: Tracing Living-Learning Community Students' Engagement in Big "D" Discourses	1406
<i>Andi M. Rehak</i>	
Design of Automated Guidance to Support Student Agency and Knowledge Integration in Science Learning	1407
<i>Charissa Tansomboon</i>	
Indexes	
Author Index.....	A1-A7
Keyword Index	K1-K12

Keynote Presentations

Educational Neuroscience: A Field Between False Hopes and Realistic Expectations

Elsbeth Stern, Professor for Research on Learning and Instruction, ETH Zürich, elsbeth.stern@ifv.gess.ethz.ch

Abstract: Educational Neuroscience has emerged as an interdisciplinary field, mainly stimulated by the improvement of brain imaging techniques and a growing interest in human learning inside and outside of schools. Getting information about human learning and cognition beyond testing or behavior observation by recording brain characteristics or brain functions was highly appreciated by many learning researchers. At the same time, however, educational psychologists were concerned about the uncritical enthusiasm and unrealistic expectations among many parents, teachers, and policy makers. Meaningless slogans like “brain based learning” got attention while at the same time important insights and concepts from classical psychological and educational research were ignored or even brushed aside as being non-scientific. Part of my talk will deal with the question of why this more traditional scientific approach is less popular than brain research, and what psychologists could do to better convey their insights to a broader public, including teachers. This said, however, does not mean to underestimate the added value of considering methods and insights from neuroscience in the research on human learning. How psychological concepts like intelligence, working memory, motivation or dyslexia or dyscalculia have been reshaped in the light of neuroscience will be discussed.

Elsbeth Stern is a cognitive psychologist with special focus on academic learning in science and mathematics. After her PhD in 1987 at the University of Hamburg she held positions at the Max-Planck-Institute for Psychological Research in Munich. 1994 she became Professor of Educational Psychology at the University of Leipzig. In 1997 she started a research group on Scientific Cognitive at the Max-Planck-Institute for Human Development in Berlin. Since October 2006 she is professor of Research on Learning and Instruction at ETH Zurich where she is head of the teacher education program. From 2012-2015 chaired the Department of Humanities, Social and Political Sciences.

Elsbeth Stern has been researching the interaction between intelligence and knowledge in cognitive functioning in the age range of 4 to 25. How knowledge transfer can be fostered by the use of visual-spatial cognitive tools is one of her major topics. She has widely published in refereed journals and books and has been member of the board of many journals.

Moreover, she is involved in international discussions on the potentials and limits of cognitive neuroscience on improving academic learning.

Transformative Learning in Design Research: The Story Behind the Scenes

Yael Kali, University of Haifa, yael.kali@edtech.haifa.ac.il

Abstract: Design research, from its inception and until today, when it constitutes a leading method in the learning sciences, has sought to explore how learning is shaped by design. Of particular interest has been the type of learning known as *transformative*, one that results not so much in the recognition of new facts about matters under study as in a reorganization of the ways of looking at, and thinking about, those matters. Yet, whereas the design research literature is full of reports about transformative learning of students, much less attention is paid to transformative learning frequently experienced in parallel by the researchers themselves. The focus of this talk is on those unplanned, often surprising ways in which researchers arrive at insights that revolutionize their thinking about the phenomena they investigate. As illustrated by the few cases in which researchers did write on the winding roads they travelled to arrive at those insights, reflective analyses of our own learning can become quite useful for other researchers, and in those instances in which we conduct our studies in collaborative partnerships with teachers and school leaders, it can also be of great value to practitioners and, eventually, to their students.

In this presentation I introduce the term *Design Researchers' Transformative Learning (DRTL)* and create a conceptual framework for studying, and reporting about, the phenomenon it signifies. The main tenet of this framework is that design research provides a fertile ground for transformative learning among those who conduct it. This kind of learning happens due to two types of processes that are inherent to design research: (a) boundary crossing in teacher-researcher partnerships, and (b) blending analytical and creative mindsets. I illustrate the use of the DRTL framework by analyzing two case-study design research projects portraying these types of processes. In sum, I suggest to expand the dual-focus view of design research, which includes developing theory and promoting design knowledge and practice, to a triple-focus view that includes DRTL processes as well. Embracing DRTL as a third focus of design research will require the community of the learning sciences to consider alternative means for sharing and publishing design research.

Yael Kali is an associate professor of technology-enhanced learning at the Faculty of Education, University of Haifa, and the director of the Learning In a NetworKed Society (LINKS) Israeli Center of Research Excellence (I-CORE). Using a design-based research approach, Yael Kali explores technology-enhanced learning and teaching at various levels, from junior high school to higher education. Together with her students in the TEL-Design team, she focuses on the role of design, and design principles for supporting Computer Supported Collaborative Learning (CSCL), and for teacher professional development, in a Teachers as Designers (TaD) approach. Yael Kali has been a faculty member at the Department of Education in Technology and Science at the Technion-Israel Institute of Technology for seven years, and a Co-Principal Investigator at the Technology Enhanced Learning in Science (TELS) centre, headquartered at the University of California, Berkeley. She has also served as a visiting scholar at the Centre for Research on Computer Supported Learning & Cognition (CoCo) in the Faculty of Education and Social Work, University of Sydney. She currently serves as an Associate Editor for the journal *Instructional Science*.

The Diffusion of Inquiry Based Practices in the Singapore Education System: Navigating Eddies of the 21st Century

David Hung, Azilawati Bte Jamaludin, and Yancy Toh
dean.oer@nie.edu.sg, azilawati.j@nie.edu.sg, yancy.toh@nie.edu.sg
National Institute of Education, Nanyang Technological University, Singapore

Abstract: We are in the age of uncertainty, flux, and ambivalence. Exponential paces of computing, bandwidth, and technological mobility are creating processes akin to swirling eddies that rapidly reshape boundaries, practices and interactions. Personal identities can be reworked and remade; and context and content are remixed constantly. However, educational systems even among the most successful ones remain rather intransigent and often conventional. Students thus struggle with the disconnect between what they are exposed to in different affinity spaces, as seen in the gulf between their informal out-of-school learning experiences and their relatively structured experiences faced in schools. To bridge the gap, there is now an increasing call for life-long learning epitomized by a distinctive shift from learning as accumulating knowledge as 'assets' for life to participating in agentic endeavours of constantly constructing knowledge amid the changing flows of knowing. While we recognize the pertinent need for schools to embrace change-oriented toward equipping our young ones to successfully 'ride through' eddies-what are the school-orchestrated formal and semi-formal learning practices which can support students' inquiry and agentic learning? We argue that traditional practices are not all to be thrown away. Rather, in adopting a hybridization of old and new pedagogies, it necessitates epistemic shifts towards valuing not only disciplinary content and domain-specific skills but also a qualitative value shift towards 'transferable' understandings and abilities of the 21st century. Mitigating many tensions and managing this change process first occurs for teachers and school leaders, foregrounding the parallel need to attend to teachers' capabilities and professional stance, as they ride through the eddies of educational reform.

Against this backdrop, this presentation first explicates the kinds of inquiry pedagogies and dispositions necessary for engendering epistemological shifts, such as science inquiry, mobile learning, knowledge building, reasoning and argumentation, makerspaces, design and technology, and simulations and the use of multiple representations that schools and teachers can create for and with the students. Drawing upon case studies of such exemplar pedagogical innovations already implemented in Singapore schools, we highlight that the successful enactment of innovative, inquiry based learning is fundamentally premised on (i) expanding the repertoire of pedagogies; (ii) valuing both disciplinary content and process learning outcomes; and (iii) systemic policy mechanisms for reform that balances both centralization and decentralization leverages-tenets of education reform phenomena observably resonating in other high performing East Asia systems, albeit in varying ways and degrees of enactments. To this end, we bring to focus the diffusion of innovative and interactional practices for scalable practices that entails identifying a continuum of innovation diffusion models with a characterization from planned deep to planned wide. Within these models, we describe the socio-and-technical infrastructures needed to enable the change of learning practices and cultures which require support from the decentralized school-cluster structures and leadership, and centralized education policies. The process of what we term, scalable epistemic apprenticeship in the diffusion within and across schools is described, unpacking nuanced ideas of (i) creating ecosystems of inquiry practices; (ii) the importance of ecosystem carryovers; and (iii) the importance of foregrounding leadership apprenticeships to provide a framework for mitigating change and uncertainties, for navigating eddies through the lens of pedagogical innovation diffusion and interactionalities. Hence, a model of scalable epistemic apprenticeship for the diffusion of inquiry practices is presented, oriented towards turning fragments and pockets of change into widespread culture of systemic innovation. Epistemic shifts are observed in our teachers and school leaders as they navigate the interactionalities across the multiple levels of the system, and the process of apprenticeship remains salient in successfully riding through the eddies of the 21st century.

David Hung is a Professor at the National Institute of Education of Nanyang Technological University, Singapore. He received his PhD from the National University of Singapore in 1997. David is also the Associate Dean at the Office of Education Research and the Head of the eduLab initiative. These appointments require him to delve into the change of social-cultural practices in the context of schools in Singapore. He specializes in social cultural orientations to cognition, in particular, communities of practice and apprenticeship forms of learning.

In 2004, David initiated the set-up of the Learning Sciences Laboratory. He has been involved in the ICT MasterPlan initiative by the Ministry of Education since its inception in 1997 and has witnessed its evolution over the years in terms of student-centred inquiry learning mediated through technologies. A significant part of the present effort is on translation and the diffusion of educational innovations, which the eduLab initiative is trying to advance.

David is presently serving as Contributing Editor for Educational Technology, Editor for Learning: Research and Practice, and an International Advisory Board member for the Asia Pacific Education Researcher. With over 100 journal papers to his name, he is presently editing books which document the ecology of diffusion efforts in the system. He discusses the socio-technical considerations, including leadership, and the centralization-decentralization leverages needed at the classroom, school, cluster, and system levels.

Invited Symposia

Beyond Tried and True: The Challenge of Education for Innovation

Carl Bereiter (co-chair), Institute for Knowledge Innovation and Technology, University of Toronto,
carl.bereiter@utoronto.ca

Marlene Scardamalia (co-chair), Institute for Knowledge Innovation and Technology, University of Toronto,
marlene.scardamalia@utoronto.ca

Thérèse Laferrière (co-chair), Centre of Research and Intervention for Student and School Success (CRISAS/CRIRES), Université Laval, Therese.Laferriere@fse.ulaval.ca

Linda Massey, Leading Student Achievement Project, Ontario Principals' Council, lmassey@principals.ca

Bruce W. Shaw, Literacy and Numeracy Secretariat, Ontario Ministry of Education, Bruce.Shaw@ontario.ca

Shirleen Chee, Divisional Director, Educational Technology Division (ETD), Singapore Ministry of Education,
shirleen_chee@moe.gov.sg

Seng Chee Tan, Centre for Research & Development in Learning, CRADLE, Nanyang Technological
University, sengchee.tan@nie.edu.sg

Chew Lee Teo, Learning Partnership with Educational Technology, Education Technology Division, Singapore
Ministry of Education, Teo_chew_lee@moe.gov.sg

David Istance (discussant), Centre for Educational Research and Innovation (CERI), Innovative Learning
Environments project, OECD, david.istance@oecd.org

Abstract: Singapore and Ontario, Canada, have been rated as among the top education systems in the world in terms of both student achievement and progressive management. This symposium brings together leaders and researchers working in these systems to discuss efforts to go beyond present achievements and to address new imperatives to educate for innovation. Both systems are experimenting with Knowledge Building as an approach to meet this need. How they propose to do this while also upholding other educational goals forms one facet of the symposium. The other facet concerns how this work fits into the larger picture of education for innovation as seen from the points of view of research in the learning sciences and OECD's studies of innovation in education.

Keywords: knowledge creation, knowledge building, technology, innovation, design

Introduction

Century-old controversies in education have taken on new shape and urgency in what OECD has termed an “innovation-driven” society (OECD, 2010). Knowledge creation and innovation are becoming constant, pervasive, and essential and are increasingly dependent on collaborative work with knowledge and ideas. OECD's Centre for Educational Research and Innovation is now turning attention to work in the area of innovative pedagogies—an extension of their longstanding Innovative Learning Environments project. The goal is to identify and understand innovative pedagogies that improve outcomes and engagement of young people, and to build an international community of expertise regarding innovative pedagogies. At the same time, ministries of education worldwide are creating plans and new educational means to advance education for innovation.

This symposium brings the leader of OECD's Innovative Learning Environments project into conversation with Ministry of Education and Principals' Council representatives from two major school systems—Ontario, Canada, and Singapore. According to the McKinsey report (Mourshed, Chinezi, & Barber, 2010), Singapore and Ontario are among the top five education systems in the world. The report labels these systems as “sustained improvers,” which means that they have seen “five years or more of consistent rises in student performance spanning multiple data points and subjects” (p. 11). “Sustained improvers,” according to this report, have also successfully moved past a centralized system to a distributed model where schools, teachers and administrators take up more responsibility for developing and implementing effective instructional practice grounded in innovation, collaboration and peer-to-peer learning (p. 20). This trajectory of system improvement, as described in the report, “is all about turning schools into learning organizations” (p. 111). Along with other high-performing school systems, Singapore and Ontario are often visited by educators seeking to improve results in their own schools. The focus of this symposium, however, is not on how Singapore and

Ontario got to where they are but on where they are going from here to raise achievement even further and to address emerging needs for innovative and knowledge-creating capacity.

Both of these groups are experimenting with the same innovative educational approach, known as Knowledge Building (Scardamalia & Bereiter, 2014). Its underlying goal is to recreate educational institutions as knowledge creating enterprises; the pedagogy is defined more straightforwardly as producing knowledge of value to a community and continually improving it (Scardamalia & Bereiter, 2003). In contrast to learning, which is an internal change in competence and dispositions (internal to individuals or to communities of practice), Knowledge Building is an overt social process producing public objects such as inventions, designs, explanations, interpretations, theories, histories, solutions, proofs, and plans. The key to doing this in educational settings is establishing *collective responsibility for idea improvement* as a socio-cognitive norm. An important part of Knowledge Building as a research program is developing technology to assist students in exercising such responsibility. Current developments include software usable by students to map “idea threads” in their group discourse (Zhang, et al, 2015), compare their usage of domain vocabulary to that of more knowledgeable groups (Resendes, et al, 2015), identify and organize promising ideas in their online notes (Chen, Scardamalia & Bereiter, 2015), and examine the communication structure in their class or group (Matsuzawa, Oshima, Oshima, Niihara, & Sakai, 2011).

Knowledge Building is synonymous with knowledge creation as carried out in innovative organizations, but it reflects the differences in context and challenge when knowledge creation is carried out with educational ends in view (Bereiter & Scardamalia, 2014). Whereas most knowledge creating organizations can draw on already-developed talent, educational Knowledge Building must develop capacity for knowledge creation within a novice population. Whereas knowledge creation usually has a clear meaning in relation to the goals of organizations that undertake it, in education it often gets lumped together with generic constructivist learning approaches such as inquiry- and project-based learning. This tends to obscure the essential concept of students taking responsibility for idea improvement.

In many school systems, concerns about low achievement are so intense that there is little elbow room for experimenting with novel approaches such as those championed by learning scientists—let alone mechanisms for collaborative innovation networks (Gloor, 2006) and other initiatives that favor creative work with knowledge and ideas. This limits pursuit of important new objectives and practices; in turn, constrained circumstances for exploration limit demonstration of the potential of new pedagogies and technologies. However, both Ontario and Singapore school systems are committed to improving achievement for all students by internationally recognized standards as well as exploring new possibilities. Thus these systems have the potential to serve the world educational community as trailblazers, venturing beyond well-traveled pedagogical routes and developing evidence and know-how that can serve other school systems in their own efforts to pursue increasingly high-level objectives.

School systems that take on this trailblazing role face a number of problems. They face all the well-known problems of systemic change and “scaling up” (Fullan 2012; Looi & Teh, 2015; Bishop, O’Sullivan, & Berryman, 2010); but in addition they face a two-cultures problem, with a practitioner culture and a design-science culture that may share top-level goals but that may differ significantly in vocabulary, in priorities, in the “grain size” of perceived problems, and perhaps most importantly in the way high-level goals are conceptualized. This symposium provides an opportunity for such differences to come into view and to be addressed. The emphasis, however, will be on possibilities for constructive, positive action and progress toward construction of a shared “problem space” (Newell & Simon, 1972). The work to be reported, in both Ontario and Singapore, is in early stages. Yet within each context there are newsworthy advances on several fronts: refocusing education as a knowledge creating enterprise, professional development networks forming that show complex educational change can sometimes proceed at a rapid pace, school-university-government partnerships demonstrating the power of multilevel engagement, and technology supporting self organization around idea improvement, in and out of school.

Plan of the symposium

The plan of the symposium is to devote the first 10 minutes to introductory comment, followed by 40 minutes of presentations from the architects of Ontario and Singapore initiatives, 10-minute commentary on research directions by collaborating learning scientists and 10-minute analysis in light of OECD studies of educational innovations (cf. OECD, 2015). The remaining 20 minutes will be opened to audience participation.

Education for innovation

Marlene Scardamalia, Presidents’ Chair in Education and Knowledge Technologies and Director, Institute for Knowledge Innovation and Technology, OISE/University of Toronto

Currently schools face the innovation imperative of meeting governmental and private-sector objectives for developing capacity to live and work in an ‘innovation-driven’ knowledge society (OECD, 2010). This new challenge does not lessen the longstanding imperative to increase achievement by international standards. In this session the role of Knowledge Building/knowledge creation in international contexts and research programs will be presented, describing how learning scientists with experience in Knowledge Building are working closely with practitioners and administrators in both Ontario and Singapore to craft a practical program that will yield demonstrable results in terms of both knowledge acquisition and knowledge creation.

Ontario, Canada: Knowledge building and leading student achievement

Linda Massey, Associate Director of Professional Learning, Ontario Principals’ Council, and Chair of the Leading Student Achievement Project (LSA)

Bruce W. Shaw, Director, Leadership and Implementation Branch, Literacy and Numeracy Secretariat, Ministry of Education, Ontario

Linda Massey will discuss how the work on Knowledge Building is being framed within an evolving LSA Theory of Action. This model of action incorporates information inputs from stakeholders as well as from research and emphasizes the role of leadership, especially on the part of school principals and special-purpose district leaders. In moving forward beyond “tried and true,” LSA promotes collaboration across all system levels and working partnerships with universities and with the Ministry of Education’s Literacy and Numeracy Secretariat. The three Ontario principals’ associations develop and deliver webinars and face-to face programs aimed at better equipping principals as leaders in the pursuit of educational goals at the classroom level. Such leadership in turn requires enabling school leaders and teachers to function at full professional levels. Toward that end, LSA assists in the building of effective Professional Learning Communities within schools and Leadership Networks/Principal Learning Teams across districts. The LSA theory of action (Leithwood, 2014) has been so well received that the LSA model and learning strategies are now being offered to an international clientele and provide a means of sharing home-grown knowledge and skills with a worldwide educational community. LSA sees itself not only as a user of research but also as a collaborator in research. It is not only working closely with Knowledge Building researchers at the school level, but it is also supporting the development of group-level assessment technology to empower students and teachers in 21st-century creative knowledge work. LSA participants share a commitment to both the advancement of learning and to the well-being of all students. This dual goal virtually demands the kind of multi-level and cross-sector collaboration that LSA and the three Ontario principals’ associations are seeking to bring about.

Bruce W. Shaw will describe the initiating project. To ensure that all students receive needed support, the Literacy and Numeracy Secretariat of the Ontario Ministry of Education sponsors Student Achievement Officers of the Secretariat to build strong relationships within the education community and to work cooperatively with boards to build capacity within schools and districts to improve student learning and achievement. It was in fact two of these Student Achievement Officers who first brought to the attention of the Secretariat the relevance and promise of Knowledge Building. Their approach to literacy was already based on paying close attention to students’ thinking, and Knowledge Building was seen as a way of giving students’ thinking an even larger role. Still to be determined is whether a similar focus on students functioning as a knowledge-building community can fit equally well into the advancement of numeracy. That is a matter of special interest in Ontario at this time because, as in with many other jurisdictions, progress in mathematics is lagging behind progress in literacy. However, the interest in Knowledge Building goes well beyond that, to the four top-level goals set for education in the Province: achieving excellence, ensuring equity, promoting well-being, and enhancing public confidence. All four of these goals present challenges that require venturing beyond “tried and true,” but at the same time they suggest risks in doing so. To meet the challenges and reduce the risks, the Secretariat is employing the following strategies:

- Full implementation of Knowledge Building in selected sites—with students, teachers, and administrators engaged in decision making—rather than superficial implementation everywhere at once.
- Building effective Professional Learning Communities within and among schools and within boards and across boards to ensure quality implementation.
- Using the group-level analytics in Knowledge Forum technology to provide formative feedback for students and teachers and to foster increasingly high levels of agency.

- Creating practitioner networks that are not only sharing networks but innovation-generating and problem-solving networks.
- Contributing to a progressive program of educational research spanning goals for achievement, innovation and well-being.

Singapore: Learning partnership in educational technology

Shirleen Chee, Divisional Director, Educational Technology Division (ETD), Singapore Ministry of Education
Seng Chee Tan, Deputy Director, Centre for Research and Development in Learning, Nanyang Technological University

Chew Lee Teo, Lead Specialist, Singapore Ministry of Education

Shirleen Chee will discuss ministry initiatives in “Masterplan 4.” She will discuss ETD’s concerted effort to bring about innovative classroom practice with technology and projects with innovative pedagogies that have shown evidence of improving overall learning outcomes and engagement of students in a formal learning setting as well as professional develop programmes that aim to sustain innovative practice by building capacity of teachers to deepen the pedagogical use of ICT to transform student learning.

Seng Chee Tan will discuss how researchers and teacher educators work with the schools and ministry to develop teachers’ capacity in designing and implementing Knowledge Building in schools. The National Institute of Education (NIE) has been offering formal courses, from pre-service to graduate levels, to deepen teachers’ understanding in knowledge creation. In addition, researchers from NIE have been working with teachers to design, implement, and examine the enactment of knowledge building in schools.

Chew Lee Teo will discuss Knowledge Building practice in Singapore. Partnerships with schools support principled implementation of Knowledge Building practice within schools, and connect teachers across Singapore in understanding and adapting Knowledge Building practices to bring about positive impact on teaching and learning in Singapore with technologies.

This work will be discussed within the context of ministry initiatives in “Nurturing an Ecology for a Sustained Shift in Classroom Practice with Technology.” The Learning Partnership in Educational Technology (LPET) is committed to 21st century teaching and learning with information and communication technologies (ICT) in Singapore classrooms. LPET partner schools create a sustaining culture for this work. Efforts include reframing roles of teachers and students, promoting evidence-based discourse, and fostering new relationships and networks within schools, viewing each player’s role in a rich, complex, and synergistic environment.

Partnerships focus on deepening domain knowledge and developing 21st century competencies simultaneously, moving one classroom, one school at a time, supported by school leaders and school teachers. LPET partner schools co-develop ICT practices to bring students’ thinking and learning to the center of classroom practices. This heightened understanding of students’ emerging learning captured by ICT creates the impetus for teachers to design interactions within the learning experience and environment, deepening domain knowledge and 21st century competencies.

Learning scientists: Research directions

Carl Bereiter, Co-Founder Institute for Knowledge Innovation and Technology, OISE/University of Toronto,
Thérèse Laferrière, Chair, Centre of Research and Intervention for Student and School Success (CRISAS/CRIRES), Université Laval,

Marlene Scardamalia, Presidents’ Chair in Education and Knowledge Technologies and Director, Institute for Knowledge Innovation and Technology (IKIT), OISE/University of Toronto.

By undertaking experimental implementations of Knowledge Building, the Ontario and Singapore schools are doing more than engaging in a local school-researcher partnership, important as that may be. They are linking up with and playing a role in an international network of researchers and innovators who can contribute significant ideas and technology and who will in turn learn from the work going on in Ontario and Singapore. Indeed, during 2015 Ontario teachers and administrators have already conducted a webinar with Knowledge Building teachers in Atlantic Canada, and Singapore teachers have taken part in Knowledge Building Summer Institutes in Canada and New Zealand. 2016 will see the launch of the Building Cultural Capacity for Innovation initiative, which will bring a variety of educational institutions from a variety of cultures into a collaborative effort to achieve goals similar to those discussed in this symposium.

Often innovation in schools means *adopting* an innovation already developed and packaged in some form of deliverables. Knowledge Building has been around for several decades, but it is not a package. It

continues to evolve in theory, pedagogy, and supportive technology (Scardamalia & Bereiter, 2014). Reflecting the ferment in the learning sciences as a whole, it is in a period of accelerating change. The three speakers in this segment of the symposium will briefly highlight some of the new directions that should have direct implications for work going on in the connected schools. The role in Knowledge Building of assessment at the group level is an emerging issue in Quebec, and the collaborative model deployed in the Remote Networked Schools (RNS) is now getting attention in both urban classrooms and remote schools in Auvergne, France (Laferrière, Allaire, Breuleux, et al., 2015). Group-level feedback to aid students in knowledge-building discourse is showing that when students are provided with ways to compare the vocabulary of their discourse with that of more knowledgeable groups they use this feedback in conceptual growth; that they similarly take advantage of tools to identify and focus discussion on promising ideas and to extract and display threads of ideas. Semantic network analysis is also being put within students' reach (Matzuzawa, et al., 2011).

Other aspects of knowledge building inviting further development are embodied cognition, hands-on knowledge building, intellectual engagement, and emergent and rotating leadership in student groups. Learning analytics, a development that is raising interest throughout the learning sciences, is being investigated with a view to using automated analyses to support students' epistemic agency, as distinct from the prescriptive uses to which such analyses are often put. A learning challenge that has been identified but not yet pursued is transliteracy—the ability to construct coherent knowledge out of the bits and pieces of information common in internet media. On other fronts, progress is being made in moving practitioner networks from being limited to sharing and social support to being productive innovation networks. An important problem on which we expect progress to be made in the ongoing work with Ontario, Singapore, Quebec, and other schools is defining an optimum role for teacher guidance and instruction within a classroom knowledge building culture. The premise is that knowledge creation—which means students not only generating ideas but taking collective responsibility for idea improvement—can actively promote literacy, numeracy, and subject matter mastery, but that the teacher constitutes a unique resource in this process and needs to play a role that maximizes long-term benefit both to the students and to the society that will depend on them.

OECD pedagogical innovations: Reflections and next steps

David Istance is a senior analyst in OECD's Education and Skills Directorate, Centre for Educational Research and Innovation (CERI). He heads CERI's Innovative Learning Environments project and is developing new work on innovative pedagogy. His most recent OECD publication is *Schooling Redesigned: Towards Innovative Learning Systems* (2015). He earlier headed the Schooling for Tomorrow project, designed and wrote the initial volumes of *Education Today: the OECD Perspective and Trends Shaping Education*, and edited an international reader on lifelong learning with Tom Schuller and Hans Schuetze (Open University Press, 2002). Other significant works include *Trends Shaping Education*, OECD, 2008 and 2010; *The Nature of Learning: Using Research to Inspire Practice*, OECD, 2010; *What Schools for the Future?* OECD, 2001; and *Education and Equity in OECD Countries*, OECD, 1997. He headed OECD's review of *Scotland's Curriculum for Excellence* and is currently working with the policy review division on Aboriginal education. As suggested by these important and highly relevant works, David is uniquely qualified to reflect on the Ontario and Singapore initiatives, to provide perspective from his work with educators worldwide, to discuss challenges and opportunities for the creation of an international design lab to advance education for innovation, and to compare the work reported in this symposium to innovative learning environments and pedagogies he and his team have investigated.

Issues for discussion

The following are issues that may be introduced as questions to guide the audience discussion:

- What constitutes “tried and true”? Have project-based learning and inquiry, for instance, become well enough established that they are part of “tried and true” rather than innovations?
- To what extent should a school system's involvement go beyond “try it and see if it works” to “try it and make it work”?
- How much lead time should be allowed for an innovation to succeed before established criteria of success and failure are applied?
- What are the prospects for adopting a new approach that achieves important new objectives but does not do quite as well or at least does no better than the existing approach on common measures of achievement?

- How much weight should be given to outcomes that can only be assessed qualitatively or by self-report—for instance, intellectual engagement, epistemic agency, well-being, and knowledge creation?
- What is the place for assessment at the group level (of discussion quality, for instance, or collaborative knowledge building or maintenance of intellectual and disciplinary norms) as distinct from assessment at the individual level, which now dominates educational assessment?
- Can practitioner networks be transformed from sharing networks to innovation networks (Gloor, et al, 2012)?
- What are the prospects for collaboration between high-performing school systems and struggling school systems? How could such collaboration be a win-win?

Significance of the symposium for the learning sciences community

Both of the education systems involved in the symposium have had long-standing working relations with research-intensive academic institutions (the National Institute of Education in Singapore and the Ontario Institute for Studies in Education in Ontario), with movement of personnel between organizations not uncommon. Nevertheless, the collaboration that figures in this symposium is distinctive in the extent to which learning science research is involved. This symposium speaks in direct and concrete terms to a concern voiced by many ISLS members: How can the learning sciences have more influence on educational policies and practices?

Significance of the symposium for system-level educational policies and practices

The symposium will address the following issues:

- Overall approach to educational improvement and the perceived potential of Knowledge Building within the larger plan.
- How scaling up from pilot experiments to full-scale implementation is best managed.
- How to provide for continuing design improvement after large-scale implementation.
- Promising ideas on the horizon that have not yet been translated into usable innovations.
- Risks inherent in carrying out research-based design experiments in actual schools, how serious these risks are, and what measures can be taken to minimize them.
- Next-generation software and analytic tools: How intelligent can they be? How intelligent *should* they be?

References

- Bereiter, C. & Scardamalia, M. (2014). Knowledge building and knowledge creation: One concept, two hills to climb. In S. C. Tan, H. J. So, J. Yeo (Eds.) *Knowledge creation in education* (pp. 35-52). Singapore: Springer.
- Bishop, R., O'Sullivan, D., & Berryman, M. (2010). *Scaling up education reform*. Wellington, New Zealand: NZCER Press
- Chen, B., Scardamalia, M., & Bereiter, C. (2015). Advancing knowledge building discourse through judgments of promising ideas. *International Journal of Computer-Supported Collaborative Learning (ijCSCL)*, (DOI) 10.1007/s11412-015-9225-z
- Fullan, M. (2012). *Stratosphere: Integrating technology, pedagogy and change knowledge*. Toronto, ON: Pearson Canada.
- Gloor, P. A. (2006). *Swarm creativity: Competitive advantage through collaborative innovation networks*. Oxford, UK: Oxford University Press.
- Gloor, P. A., Grippa, F., Putzke, J., Lassenius, C., Fuehres, H., Fischbach, K., & Schoder, D. (2012). Measuring social capital in creative teams through sociometric sensors. *International Journal of Organisational Design (and Engineering)*, 2(4), 380-401
- Hong, H. Y., & Sullivan, F. R. (2009). Towards an idea-centered, principle-based design approach to support learning as knowledge creation. *Educational Technology Research and Development*, 57(5), 613-627.
- Istance, D., Schuller, T., & Schuetze, H. (Eds.). (2002). *International Perspectives on Lifelong Learning*, Open University Press.
- Laferrrière, T., Allaire, S., Breuleux, A., Hamel, C., Law, N., Montané, M., Hernandez, O., Turcotte, S., &

- Scardamalia, M. (2015). The Knowledge Building International Project (KBIP): Scaling up professional development using collaborative technologies. In C. K. Looi & L. W. Teh (Eds.), *Scaling educational innovations* (pp. 255-276). Singapore: Springer-Verlag.
- Looi, C.-K., & Teh, L.-W., (Eds.) (2015). *Scaling educational innovations*. Singapore: Springer-Verlag.
- Mourshed, M., Chinezi, C., & Barber, M. (2010, November). *How the world's most improved school systems keep getting better*. London: McKinsey and Company.
- OECD. (1990). *The teacher today*. OECD Publishing, Paris.
- OECD. (1992). *High quality education and training for all*. OECD Publishing, Paris.
- OECD. (1997). *Education and equity in OECD countries*. OECD Publishing, Paris.
- OECD. (2001). *What schools for the future?* OECD Publishing, Paris. DOI: 10.1787/9789264195004-en
- OECD. (2008). *Trends shaping education*. OECD Publishing, Paris. DOI: 10.1787/9789264046627-en
- OECD. (2009). *Education today: The OECD perspective*. OECD Publishing, Paris. DOI: 10.1787/9789264059955-en
- OECD. (2010). *The OECD innovation strategy*. OECD Publishing, Paris. DOI:10.1787/9789264083479-en
- OECD. (2015). *Schooling redesigned: Towards innovative learning systems*. Educational Research and Innovation, OECD Publishing, Paris. DOI: <http://dx.doi.org/10.1787/9789264245914-en>
- Matsuzawa, Y., Oshima, J., Oshima, R., Niihara, Y., & Sakai, S. (2011). KBDeX: A platform for exploring discourse in collaborative learning. *Procedia: Social and Behavioral Sciences*, 26, 198-207.
- Newell, A. & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Resendes, M., Scardamalia, M., Bereiter, C., & Chen, B. (2015). Group-level formative feedback and metadiscourse. *International Journal of Computer-Supported Collaborative Learning*, 10(3), pp 309-336.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In *Encyclopedia of education* (2nd ed., pp. 1370-1373). New York: Macmillan Reference, USA.
- Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences (Second Edition)* (pp. 397-417). New York: Cambridge University Press.
- Tan, S.C., So, H.J., & Yeo, J. (Eds.). (2014). *Knowledge creation in education*. Singapore: Springer.
- Zhang, J., Chen, M.-H., Tao, D., Lee, J. Sun, Y., & Judson, D. (2015). Fostering sustained Knowledge Building through metadiscourse aided by the Idea Thread Mapper. In T. Koschmann, P. Häkkinen, & P. Tchounikine (Eds.), *Proceedings of the International Conference on Computer Supported Collaborative Learning (CSCL 2015)*. International Society of the Learning Sciences.

Future of the CSCL Community

Sten Ludvigsen (chair), University of Oslo, stenl@iped.uio.no
Heisawn Jeong, Hallym University, heis@hallym.ac.kr
Cindy E. Hmelo-Silver, Indiana University, chmelosi@indiana.edu
Nancy Law, The University of Hong Kong, nlaw@hku.hk
Ulrike Cress, University of Tuebingen, u.cress@iwm-tuebingen.de
Peter Reiman, The University of Sydney, peter.reimann@sydney.edu.au
Manu Kapur, National Institute of Education (NIE) at Singapore, manu.kapur@nie.edu.sg
Nikol Rummel, Ruhr-University Bochum, nikol.rummel@rub.de

Abstract: Learning take place under different affordances and constraints make up by people and a range of computational tools. The conference theme for the learning science conference this year - transformation of learning and empowering learners - are foundational for the CSCL community. To learn in collaboration with computer support can mean to develop new cognitive and epistemic orientations, social and emotional dispositions and deep engagement in participation in activities and communities.

As a field within the learning sciences the CSCL community will transform itself as social, cultural, cognitive and technological changes occur. The CSCL community contributes to these developments in various ways, by offering new pedagogical and technological design, computational tools and deep analysis of how and what people learn. The unique combination of the new possibilities that technologies can offer and the empirical analysis of what people choose to do together create a kaleidoscope for new scientific insights into people's learning. In this symposium we will emphasize the new challenges for the CSCL community for the next the years.

Keywords: computer supported collaborative learning, CSCL journal and community

Relevance, background and introduction

In the fields of learning sciences and computer supported collaborative learning (CSCL) the main aim is to understand and explain how and what people learn (Sawyer & Nathan, 2014). To transform learning and empower learners is at the heart of the agenda for the CSCL community. Transformation of societies at micro-, meso- and macro-level changes the dynamics and conditions for learning. More specifically collaboration and learning with computational support create new demands for learning in schools and in working -life. New computational tools and infrastructures create new forms of interdependencies with the social mind (Tomasello, 2014). New tools make new actions and activities possible. This means that people can do different things and new task in companion with technologies. Learning in collaboration often involves social, cognitive and emotional aspects and computational tools that affect processes and outcomes.

Since the early 1990s, CSCL has been established as an interdisciplinary field (e.g. Ludvigsen & Mørch, 2010; Stahl, Koshchman & Suthers, 2014; Jeong, Hmelo-Silver, & Yu, 2014; Tang, Tsai, & Lin, 2014; Stahl, 2015). Computer-supported collaborative learning builds on different scientific fields, such as the learning sciences, communication studies, computer science (human computer interaction, computational linguistics, etc.) and some branches of the social sciences.

Methodologies from a number of fields are a part of the repertoire of CSCL researchers. More concretely, methods from experimental psychology, analysis of social interaction, design studies and field studies are used in the CL part of the CSCL field, while in the CS part methods are connected to development of hardware, software and interfaces and the use of formalism (e.g. Tchounikine, Mørch & Bannon, 2009). We emphasize here that the defining features of CSCL are the interdisciplinary dependencies between theories of learning, collaboration and computer science.

The CSCL community – the next decade

The International Journal of Computer-Supported Collaborative Learning (ijCSCL), the CSCL conferences, and work in the many strong CSCL communities around the world have contributed leading research in how we can understand the human-computer/technology connection (or computational link) and how and why new technologies can support learning in specific settings. The different forms of CSCL research raise basic and applied questions, which communities should strive to address. Additionally, CSCL research contributes to an

important field of knowledge within learning sciences.

The CSCL is the leading community in defining research frontiers about human-computer connections. With a focus on learning in collaboration and computer support, the CSCL community has contributed to a wide range of studies in many different knowledge domains. These studies have a common core (i.e., collaborative learning and computational support), while the domains, contexts, and methodologies vary. The community has a clear strength in that it has maintained a strong focus on how to understand and explain collaboration in relation to computational support. Many communities and their journals contribute with studies on computers and education as well as computer assisted learning. However, the specific aspects addressed by the CSCL community create a cumulative structure that broader, thematic-oriented community cannot achieve. The profile of the community has contributed to its strong reputation as a high quality, interdisciplinary community in the fields of learning sciences and educational research.

To consolidate and further develop its position as a high-quality community the leaders on the field must engage the community in developing new ways of understanding and explaining learning in collaboration with computer support. When new trends arise such as new designs for learning in mathematics and science, games and simulations in various domains, or MOOCs, and where new quantitative methods and techniques are used, we must be open to new ways of understanding and investigating collaboration. Some of the new methods and techniques are often connected to concerns about big data and learning analytics (here used as an umbrella term), but other datatypes also emerge as sources of understanding the mechanisms of collaboration.

The mentioned trends are just examples of “hot topics” today; they were put on the agenda at the 2015 CSCL conference in Gothenburg for both invited and regular symposia and papers (Lindwall et al 2014). The community should be the frontrunner in creating and defining new research trends. We now see that new methodologies (e.g., learning analytics) and other means of collecting subjective and objective data are emerging. These new approaches, methods, and techniques will challenge how we conceptualize collaborative learning with computer support, and the community must be open toward such endeavors. As a community, we have also been open to investigating new ways of supporting collaborative learning as new computational tools are developed. Furthermore, the community has taken a critical stance on different theoretical stances, and we have established and used new methodologies in CSCL studies. Being open to new trends and taking a critical stance have been the strengths of the CSCL community.

The vision for the future must include making efforts to create and follow-up on current trends in the field. We must also ensure that the journal receives a high number of research papers that can challenge the frontiers within the fields of CSCL and learning sciences.

Global contributions

The CSCL community is a global community. One important aim is to recruit additional members from all continents around the world, especially those who are not currently strongly represented in the community. Although the Asian community has representatives from Japan, Hong Kong, and Singapore, other countries and regions are not as well represented. For example, the continent of South America is underrepresented, and countries in southern Europe could have more of a prominent presence as well. Furthermore, the African continent is not well represented in our community. Efforts to promote the CSCL field should be seen as a long-term strategic effort, but we should also create concrete targets for the upcoming period.

Societal impacts

Development within the CSCL field takes advantage of the cumulative growth of knowledge in its own context as well as in the relevant contexts of learning sciences and computer sciences. In recent years, policy-related questions have been raised concerning whether advanced CSCL tools and environments can be used in all types of educational institutions and in more informal settings. We can identify growth and increased diversity in empirical fields where CSCL studies are increasingly important (e.g. workplaces, higher education settings, and increased uses of mobile technologies). One important question becomes how can CSCL research influence social practices?

CSCL research can model practices that represent new ways of designing environments for teaching and learning, and as researchers and developer engaged in projects we can demonstrate how CSCL tools can be used. The CSCL field can impact social practices by modelling how deep learning can take place. To create change on a large scale involves much more than scientific studies and developmental work. Local and national politics are often part of large-scale changes, as indicated by research on school system reforms. Researchers in Singapore have interpreted broad policy imperatives in educational policy with CSCL, and worked together with teachers to implement and do research on advanced CSCL practices in classrooms (Looi et al, 2010). Also in Hong Kong research teams have managed to influence educational policies and have used CSCL as part of their

overall framework for improving their educational system and the performance of their teachers and students (e.g. Chan, 2011; Law 2010).

CSCL design-based research that has been tested in naturalistic environments has provided the research community with important information on how to make use of new technologies in schools and higher education settings. Such studies make the culture of schooling more visible. By culture of schooling we here refer to norms, values and expectations. Norms, values and expectation can be general and work across subjects or they can be domain specific (usually both) is part of conventions and practices that can enhance or hamper students learning.

Most CSCL studies are conducted with rather low numbers of participants. This means that CSCL studies are often based on small-scale experiments, quasi-experimental designs, or design-based studies in natural contexts. In experimental studies, specific hypotheses are tested. Conversely, in design-based and explorative studies, one investigates what and how students can learn under specific conditions. All these types of studies should be perceived as part of the research design that provides novel contributions to CSCL. When we ask questions that cut across different studies, specific patterns emerge that reveal how we can better create designs for deep learning. These different research designs, assumptions, and perspectives do not need to be fused together. Rather, they can be seen as incremental steps toward an advanced understanding of how and what students learn in collaboration with new technological tools in environments that have been designed and built using generalized knowledge from the CSCL field. Variations in research designs reveal how different perspectives contribute to CSCL.

While the CSCL design often emphasizes that students' work should be modelled on (parts of) scientific practices, many schools and classrooms are based on other social or cultural conventions. This means that CSCL designs can create tension and discontinuity in students' and teachers' (educational) practices. Such tensions and discontinuities can create 'seeds for change' to encourage new practices that are based on state-of-the-art learning principles.

Global impact

A new large scale assessment study that was performed by the OECD in 2015 (November) provided a strong message about the potential impact of collaborative learning in the 21st century (OECD 2015). The PISA Collaborative Problem Solving (CPS) study involves peer collaboration using specific tools in an assessment of twenty-first century skills. Large scale studies involve particular methodological challenges in CSCL. Leaders in the CSCL field have contributed to this work (e.g. Friedrich Hesse, Pierre Dillenbourg, Looi Chee Kit and others). The fact that OECD included collaborative problem solving as one of the most important twenty-first century skills reveals the impact made by the CSCL community. It also indicates that we can expect growing interest among stakeholders in the field of education as well as growing interest from other research communities that have populated the CSCL community up to today.

Perspectives, orientations and multiple layers

In CSCL research, one can identify influential studies that are based on the cognitive, socio-cognitive or the socio-cultural perspective. The different orientations imply that their analytic attentions are directed towards different aspects of learning and human cognition. The most important difference is how collaboration is accounted for. Within the cognitive and socio-cognitive perspective, individual processes and outcome measures are normally assessed. The socio-cultural studies have mostly been concerned with the investigation of emerging interactions and practices. Conceptually, one could frame the different units of analysis in CSCL as three interdependent layers, all of which are needed to understand and explain learning, human cognition and development with computational tools. As a community we need variation of unit of analysis and levels of descriptions (Ludvigsen & Arnseth, in press).

CSCL: An interdisciplinary field of research and practice

As we will show in this symposium CSCL research is diverse when it comes to perspectives and methods. The community conducting both quantitative and qualitative studies, and make use of mixed methods approaches. We explore ways to discriminate between what can be explained either by social interactions or by individual students' actions. In order to understand collaboration with computer support we need experimental studies, quasi-experimental studies, naturalistic studies, randomized controlled trials, and a wide range of methods. Furthermore, we must continue to develop multiple perspectives based on basic models that come from cognitive, socio-cognitive, or socio-cultural research. This means that both experimental and naturalistic settings are required to further explore key issues in the CSCL field.

Below follows the themes and problems that leading CSCL researchers envision and predict will create some of the grand challenges the coming years. These challenges are related to theories, methodological issues and how new representational technologies and practices will influence collaborative learning.

Methodological and theoretical traditions in CSCL research

Heisawn Jeong, Hallym University and Cindy E. Hmelo-Silver, Indiana University

As an interdisciplinary research fields, CSCL is built on diverse research traditions that contributed greatly to the diversity and vibrancy of the field, but at the same time created confusion about what is an acceptable research. We have been engaged in a long-term project that examines CSCL research practice from a number of different dimensions including research method, theoretical frameworks, technology, and outcomes (Faulkner, Hartley, Hmelo-Silver, & Jeong, 2016; Jeong, Hmelo-Silver, & Yu, 2014). In particular, we have examined methodological practices of CSCL research in terms of research design, settings, data sources, and analysis methods along with their relationship to theoretical frameworks. It revealed that CSCL research consists of four distinct method-theory clusters.

In this symposium, we will present an updated analysis of the method-theory clusters over ten year period (2005-2014) with an additional focus on the role of the *ijCSCL* journal. Reflection on the methodological and theoretical practices of CSCL can help us to brainstorm how different research traditions can benefit from and complement each other and what role *ijCSCL* can play in helping the field to advance in a more integrated manner

Methodological challenges in CSCL community

Ulrike Cress, University of Tuebingen

Currently there is a big discussion going on about methodological issues in the social sciences (including psychology, learning science, empirical research) and the validity of results from their empirical studies. More and more scientists criticize that journals have a tendency to publish only significant results and neglect non-significant ones (Wagenmakers, Wetzels, Borsboom, van der Maas & Kevit, 2012)

As a consequence, whole disciplines may deliver highly biased results, where randomly found pattern represent more noise than real regularities. Many journals (e.g. *Journal of Media Psychology*) started to react on this issue by explicitly calling for “pre-registered studies” (Nosek & Lakens, 2014). They allow submissions where the review procedures start already before the studies are conducted and where the study design, its theoretical foundation and methodological quality is reviewed. If the study is accepted, the results are published independently of their statistical significance. Journals furthermore started to give replication studies much more value. They expect that results pattern are replicated in other studies and by other authors of labs.

The presentation will describe these current tendencies and show why this is relevant for *ijCSCL* to take into account these issues and find a strategy to deal with them.

A semiotic turn in CSCL?

Peter Reimann, University of Sydney

I predict that we will see a kind of semiotic turn in CSCL, with a focus on materiality; a rising interest in the kind of notional and representational systems that are used when people collaborate in particular practice fields. Semiotics is the study of sign systems, their symbolic as well as physical qualities (Eco, 1979). While there was a certain interest in the first phase of CSCL--the discussions forums, online forums--in semiotic aspects of collaboration, those first generation semiotic devices were designed for the purpose of asynchronous communication and exchange (‘discussion forum’, ‘thread’). They were not so much informed by people's practices and activities. In more recent years, we've seen a continued interest in these systems, and a surging interest in talk, in synchronous communication. A particularly active area that yielded numerous ideas for representational notations as been research on computer-supported argumentation (Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2012).

The new semiotic turn should focus on artifacts that are representative of people's practices, rather than artifacts designed specifically for the purposes of communication and learning. For instance, the blueprints that building engineers and architects use, the symbol system that musicians use, the specialized document types and codes medical practitioners use. There has been more interest on practice-related notations and artifacts in

CSCW than in CSCL (e.g., Turner, Bowker, Gasser, & Zacklad, 2006), and still comparatively little work in CSCL that engages with authentic artifacts and their role in collaboration and learning.

As an example for what CSCL research with a semiotic perspective could look like, think of Dan Suthers's early work on the guidance function of specific notional systems (e.g., Suthers & Hundhausen, 2003), but now with a focus on notations and artifacts that have a more discipline/profession-specific grounding and are more practice-based.

I can see a number of benefits of the 'new semiotic turn': For instance, content would become more important again; we are currently perhaps too much focused on the analysis of the collaboration process (Reimann & Yacef, 2013). But without a concern for content, process remains hard to understand. Another benefit would be the development of stronger ties between CSCL and CSCW. Thirdly, CSCL would become more relevant for vocational and professional learning because we would now be studying and supporting collaborative learning around a range of artefacts much wider than dedicated 'knowledge' artifacts such as concept maps and math equations. Furthermore, a semiotic perspective on collaboration could contribute to HCI research (de Souza, 2005) and to the development of task-related applications that support learning in (collaborative) practice, in addition to getting a task done (solving a problem).

The question I want to raise is what the reasons might be that practice-related artefacts play still such a little role in CSCL. Why are they left behind? Maybe it is because they require specialized knowledge, and most of CSCL researchers are not at the same time engineers, doctors, musicians, accountants? Maybe it is because these kind of artefacts are difficult to analyze computationally? Maybe it is because we still make a strong distinction between learning and work, at least in K-12, arguably even in studies that take place in the tertiary sector?

Timescales in CSCL research

Manu Kapur, National Institute of Education (NIE) of Singapore and Nikol Rummel, Ruhr University of Bochum, Germany.

Learning unfolds over multiple timescales. CSCL can range from shorter time scale ad-hoc groups (in the order of minutes to hours and days) to longer timescale groups and communities (in the order of weeks to months, and even years). Although there is an increasing trend towards longer timescale studies (e.g., Forte, 2015; Siqin et al., 2015), it seems that majority of CSCL research still tends to focus more on shorter timescales, leaving important developmental and learning trajectories under-explored. Much as we must continue to explore shorter timescale collaborative dynamics, we need to redouble our efforts towards theorizing, designing, and developing methods of analyzing longer timescale dynamics in CSCL, especially if we want CSCL research to make a significant ecological contribution to theory and practice. Such a push will minimally involve a commitment to 4Cs: a) stronger Coupling between collaborative learning theory and principled ways of designing for CSCL, leading to a specification of design principles and the learning mechanisms they embody (Sandoval, 2014), b) exploiting Complementarity and convergence among multiple analytical methods, including machine learning (e.g., big data) and computational modeling, but in ways that inform theory and design, c) Coordination of theory, design and analysis not only across timescales but also between CSCL and non-CSCL contexts, because learning across longer timescales cannot be restricted only to CSCL contexts, and d) interdisciplinary Collaboration between learning scientists, computer scientists, statisticians and learning analytics experts, technologists, and domain experts.

CSCL beyond group cognition?

Nancy Law, University of Hong Kong

Stahl (2015) provides a succinct summary of the first 10 years of ijCSCL, describing the strong focus even within CSCL research to focus on the performance and behavior of individuals, and argues for the need to focus on group cognition. In this symposium, I argue that CSCL researchers need to study collaborative problem solving and knowledge construction in real world, authentic settings beyond classroom and formal educational settings. There are emerging forms of CSCL that are important in organizations and open collaboration communities that cannot be well understood if we confine our theoretical lens to groups of individuals. Expanding our focus beyond the formal educational contexts will bring exciting new theoretical dimensions into our research, and possibly connect us to communities that we have not considered as cognate before.

The classic form of problem solving focuses on well-defined problems and is well described by Polya (1973), consisting of four key steps: understanding the problem, devising a plan, carrying out the plan, and reviewing. Studies of collaborative inquiry have largely been carried out in similar problem contexts. Hence,

how to achieve inter-subjectivity and a common problem representation has been a key concern in the CSCL literature (Stahl, Koschmann and Suthers, 2006; Hesse et al., 2015). Studies of group cognition have contributed much to our understanding of the social dimensions of learning and provided convincing evidence that what a group can achieve goes beyond the sum of the knowledge and expertise of the individuals involved, and that the quality of the interactions matter. On the other hand, in many situations, the basis for a successful collaboration may not be shared representations and shared goals.

Actually, knowledge building as defined by Scardamalia and Bereiter (2003), which is very familiar to the ijCSCL community, does not focus on a group. The focus of the collaboration is on idea improvement, which is to be achieved through intentional knowledge building discourse characterized by the 12 knowledge building principles. In their design-based research settings in classrooms, they have specifically avoided the formation of groups even in classroom learning contexts in order to allow for free and open exploration of ideas. This form of collaboration is analogous to collaboration within the scientific community, which is borderless.

Another form of collaborative problem solving is that commonly found in project teams in business and industry that require interdisciplinary collaborations of members with different expertise and roles. In this type of collaboration, there is a well-defined team structure, but members have varying degrees of overlap in their expertise, have different roles, and contribute differently to the project. Success depends on team members sharing situational awareness (i.e. the knowledge of what each other is doing), understanding of the organizational goals, the nature of the project at hand and the approach/methodology for its accomplishment. When team members need to communicate for collaboration purposes, they need to be able to share at higher levels of abstraction beyond the specific concepts and representations used in tackling their own specific task within the team (Tran, 2008).

The question I would like to pose is what is the scope collaborative learning would our community wish to define for ourselves? Do we consider the examples of collaboration outside of the formal classroom as CSCL? When computational devices are so powerful and prevalent in taking over the solution of well-structured problems, learning is a necessary part of the everyday workplace. Are these other forms of CSCL also relevant to the educational community? I would like to argue that the examples of non-group based CSCL described above are different forms of network-based learning. Many of us in the CSCL community see teacher collaborative learning as crucial to bringing about change in pedagogical practice in the classroom. However, effective teacher learning cannot be brought about through group-based collaboration alone, and are much better effected through innovation networks (Law, Liang and Lee, under review). To me, this is one major motivation for the CSCL research community to go beyond the study of group collaboration.

Challenges for the CSCL community in the next decade

The different contributions in the symposium highlight some of the most likely changes that will occur in the CSCL community. Below we emphasize some of the main challenges.

New theoretical challenges. The main issue concerns how we should conceptualize and analyze collaborative learning. This is related to the micro-, meso-, and macro-dimensions of collaboration in specific situations and over longer time periods. The problem of timescales and how they can be analytically connected is emerging as a challenge that needs more attention.

New methodological challenges. This challenge is closely related to the theoretical challenge. Most CSCL studies deal with rather short timespans, while new methods and techniques make it possible to follow students/people for longer periods of time. We can say that collaboration is distributed in time, space, and across organizations. In psychology and the social sciences, longitudinal studies with a mix of data sources are now used. The research design that we now use will need to be critically examined in order for us to develop new research designs that can create more valid results. The question then becomes: what does this mean for the CSCL community, and how can we still work to understand collaborative mechanisms — not just the results of collaboration?

The landscape of new technologies and infrastructures. MOOCs and other environments make it possible to collect huge amounts of data. To connect such data in order to understand collaborative learning is still a challenge. Collaborative learning is the point of fixation for CSCL studies. The question becomes how different data sources and analyses can give us new insights about this phenomenon.

Semiotics or meaning-making will be central in the future. The artefacts that are used in education and in professional work are getting more and more advanced. Artefacts used by engineers and medical staff in their work contain inscribed knowledge that is rather advanced, creating knowledge gaps for newcomers to a profession. The inscribed knowledge would need to be understood and articulated. In schools, increasingly advanced statistics, simulations of abstract phenomena, and visualizations have become part of everyday

education in many countries. The question become: what does it take to create meaning from very advanced artefacts?

A question related to the semiotic turn is how do we educate students and other actors to work with multiple resources. The semiotic work involves students working with multiple resources, creating radical change in the collaboration and cognitive efforts involved. While many traditions try to capture this phenomenon by following only the individual student/actor, the CSCL community already has a repertoire for analyzing how students and others talk about certain aspects of an artefact. This is not only about interpreting single resources, but also institutional work with multiple pathways of justification.

The last challenge that we put forward here concerns 21st-century skills and in-depth learning. The CSCL community can and will contribute with models for learning that can show how such learning can take place in different domains and settings. The question here becomes: how can such efforts be part of the evidence chain that leads to new educational policies?

Organization of the symposium

The speakers will present their main ideas in 7-8 minutes. The audience will work together in small groups and report back to the symposium organizers. The last 5-10 minutes will be set aside for a summary of future challenges for the CSCL community.

References

- Chan, C. K. K. (2011). Bridging research and practice: Implementing and sustaining knowledge building in Hong Kong classrooms. *International Journal of Computer-Supported Collaborative Learning*, 6, 147-186.
- de Souza, C.S. (2005). *The semiotic engineering of human-computer interaction*. Cambridge, MA: MIT Press.
- Eco, U. (1979). *A theory of semiotics*. Bloomington, IN: Indiana University Press.
- Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In *Assessment and teaching of 21st century skills* (pp. 37-56). Springer Netherlands.
- Faulker, R., Hartley, K., Hmelo-Silver, C. E. & Jeong, H. (2016). Technology-based learning environment for STEM education. To be presented at AERA 2016.
- Forte, A. (2015). The new information literate: Open collaboration and information production in schools. *International Journal of Computer-Supported Collaborative Learning* 10 (1), 35-51.
- Jeong, H., Hmelo-Silver, C. E., & Yu, Y. (2014). An examination of CSCL methodological practices and the influence of theoretical frameworks 2005–2009. *International Journal of Computer-Supported Collaborative Learning*, 9, 305–334.
- Law, N. (2010). Teacher Skills and Knowledge for Technology Integration. In B. McGaw, E. Baker & P. Peterson (Eds.), *International Encyclopedia of Education* (3rd ed., Vol. 8, pp. 211-216). Oxford, UK: Elsevier.
- Law, N., Liang, L.M. & Lee, Y. (under review). What Happens to the Innovation when Project Funding Ends? *Learning Architecture Matters!* Paper submitted to the International Conference on Learning Sciences 2016.
- Lindwall, O., Häkkinen, P., Koschman, T. Tchounikine, P. & Ludvigsen, S. (Eds.) (2015). *Exploring the Material Conditions of Learning: Opportunities and Challenges for CSCL*. The Computer Supported Collaborative Learning Conference (CSCL) 2014, Gothenburg, Sweden. The International Society of the Learning Sciences.
- Looi, C.K., So, H-J., Toh, Y. & Chen W. (2011), The Singapore experience: Synergy of national policy, classroom practice and design research, *International Journal of Computer-Supported Collaborative Learning*, 6, 1, pp. 9-37.
- Ludvigsen, S. & Mørch, A.I. (2010). Computer-Supported Collaborative Learning: Basic Concepts, Multiple Perspectives, and Emerging Trends, In P. Peterson; E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education*, Third edition, Volume 5 (pp.290-296). Oxford, UK:Elsevier.
- Ludvigsen, S. & Arnseth, H.C. (in press). Computer-supported collaborative learning (CSCL). Perspectives and challenges. In Sutherland, R. et al (Eds.) *The STELLAR TEL Reader*.
- Nathan, M.J. & Sawyer, R.K. (2014). Foundation of the Learning Sciences. In Sawyer, R. Keith (Ed.).*The Cambridge Handbook of the Learning Sciences* (pp. 21-43). Cambridge University Press. New York.

- Noroozi, O., Weinberger, A., Biemans, H., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSCCL): A synthesis of 15 years of research. *Educational Research Review*, 7, 79-106.
- Nosek B. A., Lakens D. (2014). Registered reports: a method to increase the credibility of published results. *Soc. Psychol.* 45, 137–141.
- OECD (2015). PISA 2015. Collaborative problem solving. www.oecd.org/pisa/pisaproducts/
- Polya, G. (1973) *How To Solve It: A New Aspect of Mathematical Method*. Princeton, New Jersey: Princeton University Press.
- Reimann, P., & Yacef, K. (2013). Using process mining for understanding learning. In R. Luckin, S. Puntambekar, P., Goodyear, B. Grabowski, J. D. M. Underwood & N. Winters (Eds.), *Handbook of design in educational technology* (pp. 472-481). New York: Taylor & Francis.
- Sandoval, W. (2014) Conjecture Mapping: An Approach to Systematic Educational Design Research, *Journal of the Learning Sciences*, 23(1), 18-36.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge Building. In J. W. Guthrie (Ed.), *Encyclopedia of Education* (2nd ed.). New York, USA: Macmillan Reference.
- Siqin, T., van Aalst, J., & Chu, S. K. W. (2015). Fixed group and opportunistic collaboration in a CSCL environment. *International Journal of Computer-Supported Collaborative Learning* 10 (2), 161-181.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, 2006, 409-426.
- Stahl, G., Koschman, T. & Suthers, D. (2014). Computer-Supported Collaborative Learning. In Sawyer, R. Keith (Ed.). *The Cambridge Handbook of the Learning Sciences* (pp. 479-500). Cambridge. University Press. New York.
- Stahl, G. (2015). A decade of CSCL. *International Journal of Computer-Supported Collaborative Learning*. 10, 337-344.
- Suthers, D.D., & Hundhausen, C.D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *The Journal of the Learning Sciences*, 12(2), 183-218.
- Tang, K. Y., Tsai, C. C., & Lin, T. C. (2014). Contemporary intellectual structure of CSCL research (2006-2013): A co-citation network analysis with an education focus. *International Journal of Computer-Supported Collaborative Learning*. 9(3), 335-363.
- Tchounikine, P., Mørch, A.I. & Bannon, L. (2009). A Computer Science Perspective on Technology- Enhanced Learning. In Balacheff, N., Ludvigsen, S., de Jong, T., Lazonder, A., Barnes, S. (2009). *Technology Enhanced Learning: Principles and Products* (pp. 275-288). Springer Verlag.
- Tomasello, M. (2014). *A Natural History of Human Thinking*. Cambridge, MA, London. Harvard University Press.
- Tran, M. Q., & Biddle, R. (2008, November). Collaboration in serious game development: a case study. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (pp. 49-56). ACM.
- Turner, W. Bowker, G., Gasser, L., & Zacklad, M. (2006). Information Infrastructures for Distributed Collective Practices. *Computer Supported Cooperative Work (CSCW)*, 15, 93-110.
- Wagenmakers E., Wetzels R., Borsboom D., van der Maas H. J., Kevit R. (2012). An agenda for purely confirmatory research. *Perspect. Psychol. Sci.* 7, 632–638.

Analytics of Social Processes in Learning Contexts: A Multi-Level Perspective

Carolyn P. Rosé (co-chair), Carnegie Mellon University, cprose@cs.cmu.edu
Dragan Gaesevic (co-chair), University of Edinburgh, dgasevic@acm.org
Pierre Dillenbourg, Swiss Federal Institute of Technology Lausanne, pierre.dillenbourg@epfl.ch

Yohan Jo, Gaurav Tomar, Oliver Ferschke
yohanj@cs.cmu.edu, gtomar@cs.cmu.edu, oliverf@cs.cmu.edu
Carnegie Mellon University

Gijsbert Erkens, Anouschka van Leeuwen, Jeroen Janssen, Mieke Brekelmans
g.erkens@fss.uu.nl, A.vanLeeuwen@uu.nl, J.J.H.M.Janssen@uu.nl, m.brekelmans@uu.nl
Utrecht University

Jennifer Tan, Elizabeth Koh, Imelda Caleon, Christin Jonathan, Simon Yang
jen.tan@nie.edu.sg, Elizabeth.koh@nie.edu.sg, imelda.caleon@nie.edu.sg, christin.jonathan@nie.edu.sg,
simon.yang@nie.edu.sg
National Institute of Education, Nanyang Technological University, Singapore

Abstract: In the past two decades, the field of Machine Learning has not only greatly expanded in terms of the plethora of increasingly powerful modeling frameworks it has provided, but has also birthed the applied fields of Educational Data Mining and Learning Analytics. Learning Analytics has blossomed as an area in the Learning Sciences, promising impact for various stakeholders working at different educational levels, such as Instructional Designers, Students, Instructors, Policymakers and Administrators. This symposium offers a taste of cutting edge work across each of these levels, with a common emphasis on analytics applied to social processes.

Keywords: learning analytics, social analytics, discourse analytics

Introduction

Over the past two decades, the fields of Educational Data Mining and Learning Analytics have received growing prominence in research, policy and public literature. The fields evolved from the applied disciplines of Machine Learning, Intelligent Tutoring Systems and Data Mining, which in turn have their own roots in Applied Statistics, Cognitive Science, and Computational Linguistics. This invited symposium explores the niche domain Learning Analytics is establishing within the Learning Sciences and to offer a taste of the impact it is having at multiple levels within that sphere. The symposium will open new opportunities for bridge building and examine the commonalities for collaboration across related fields.

The earliest manifestations of Learning Analytics within the Learning Sciences community connected particularly with research into collaborative learning. A decade ago at CSCL 2005, attention was given to a vision for the field over next 10 years, and an important active ingredient in that formulation was the presence of machine learning. In particular, automating analysis of collaborative processes for making support adaptive and dynamic was one of the topics discussed. Nevertheless, in the same conference, none of the sessions were named in a way that acknowledged machine learning or data mining as constituting an area. Instead papers offering modeling technologies were “hidden” within other sessions on topics like Argumentation or Interactivity. From 2005 forward, a trend of increasing attention into machine learning was evident. There was a plenary keynote and two workshop keynotes on dynamic support for collaboration at the first Kaleidoscope CSCL Rendez-Vous held in Villars Switzerland. The trend of increasing attention continued at the 2007 conference, where papers in the area became more frequent. They appeared in sessions entitled “Tools & Interfaces” or “Methods of Scaffolding”. In particular, “Methods of Scaffolding” gives evidence of beginnings of a vision that a new form of scaffolding for collaboration was becoming possible. In this session, Frank Fischer, leader in the area of scripted collaboration, presented a paper on the vision for fading of scripted support for collaboration. A major shift was apparent by the 2009 conference where there were two workshops with related work, one on Intelligent Support for CSCL and another on Interaction Analysis and Visualization,

where automated and semi-automated analytic technologies were featured topics. In the main conference, a session was included on “Scripts & Adaptation” and another on “Data Mining and Process Analysis”.

In light of the center of gravity within the CSCL community, this symposium offers a taste of cutting edge work across each of these levels, with a common emphasis on analytics applied to social processes.

Could a model of educational design enhance learning analytics?

Pierre Dillenbourg

Many educational scenarios rely intensively on social processes. Therefore, if we want to use analytics to investigate learning processes, one first needs to model the social structure of the learning activities. This social structure is often implicit, but, if the designer would make it explicit, the collected data could be processed in a more accurate way. For instance, if the same lesson includes individual and team activities, the analytics need to map individuals to groups and vice-versa. We therefore developed a language for the purpose of modeling pedagogical scenarios, represented as directed geometrical graphs. Another rationale for this modeling language is that, by making the social structure explicit, rich learning activities that have been designed by the learning sciences community, generally for small classes, could be scaled up for use with thousands of participants.

The vertices of these graphs represent learning activities (in black on Fig. 1) and the edges capture the pedagogical relationship between activities (in red on Figure 1). The objective of modeling orchestration processes has emerged as a challenge in CSCL since typical learning scenarios have evolved towards multi-plane learning scenarios. An example of such a scenario is the group formation algorithm displayed between the second and third activity on Figure 1. The graph edges are associated with data operators that implement these algorithms. A graph of data operators constitutes a workflow, which enables the construction of scenarios that are more sophisticated than those currently implemented in MOOCs. The proposed modeling language is not only relevant in learning technologies, it also allows researchers in learning sciences to formally describe the structure of any lesson, from an elementary school lesson with 20 students to an online course with 20,000 participants. This modeling language formalizes the design of learning scenarios. With that in mind, we can then consider the analytics side.

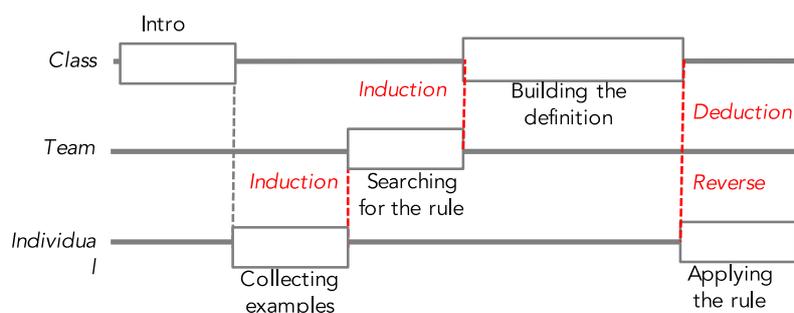


Figure 1. Example of an orchestration graph describing the following scenario: The teacher explains the goal—to find the rule that calculates the number of diagonals in a polygon from the number of edges. He assigns a number between 3 and 8 to each student. Each student draws a polygon with the number of edges assigned to him. Then, students form teams of 4 made up of students who drew a polygon with a different number of edges. They try to find the rule. After a while, each team presents its solution(s) to the class. The teacher compares the invented rules, proposes counter-examples that disprove some proposed rules, and ends up writing the formal rule. Finally, he asks students to apply the rule (G– Deduction edge label) in a reverse way (T– Reverse edge label) to calculate the number of edges of a polygon with 35 diagonals.

Let us consider that the learner state is detected at the end of each activity displayed in the graph (if finer grain analytics are required, any learning activity can be decomposed into sub-activities). In that way, the graph not only describes the pedagogical design of the scenario, but it also captures the sequence of learner states. More precisely, the sequence captures one dimension of the learner modeling process, which in turn combines 3 sources of information (and is hence represented as a cube on Figure 2):

- Horizontally: since a learning path is a time series, a learner's state can be inferred from his previous state. Let's define $x_i(s)$ as the state of the student s at the end of the i^{th} activity (a_i). The state of the class at time i is a vector with the distribution of students in each state. When two successive activities

have been completed, the two state vectors form a transition matrix $M(i, i+1)$ in which the data in cell (r,c) is the percentage of students who have moved from $x_i(s)=r$ to $X_{i+1}(s)=c$. The average entropy across the rows of M determines the weight of the edge between a_i and a_{i+1} : the lower this entropy, the better the current state of a learner can be predicted from her previous state.

- Vertically, the state of student a - $x_i(s_a)$ - is inferred from the state of another learner b at the same activity $x_i(s_b)$, for instance because s_a and s_b had very similar learning paths so far. The state of student may also be inferred from the state of the whole class, for instance if the student has always been one standard deviation above the class average. Simply stated, if the system has no information about the state of Mike in the third activity but knows that 90% of other learners failed it, it may infer that Mike has a good chance to fail as well.
- The third axis represents the cognitive diagnosis process, i.e., inferring the state of a learner from her behaviour denoted $b_i(s_a)$. This behavior representation refers to her answers, video player actions, assignments, gaze path, blood pressure, etc. In HMMs, behavioural variables are the observable variables and the learners' states are hidden states. This inference process can be a simple mapping between the answers in a quiz and a set of states or it can include a more complex interpretation processes, for instance when analysing gaze traces. Here we have also applied the notion of entropy to describe the probability that the inferred state matches the actual cognitive state of the learners.

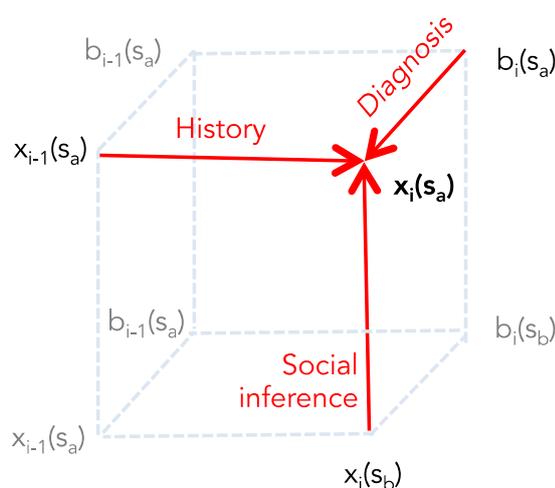


Figure 2. The Learning Analytics cube.

The connection between the graph structure and the cube corresponds to the hypothesis expressed by the title of this talk: I expect that a formal description of the designed educational structure should empower learning analytics applied to this structure. This approach has the ability to use data more insightfully than a naïve approach that ignores the underlying instructional design. Imagine a set of sensors placed in a car: they would generate better deductions if they were associated with a functional model of the car structure that explains the relationship between data points. The same should be true for education. For instance, the work on open learning analytics might lead us to aggregate data across MOOCs, which is a fantastic opportunity for learning sciences, but would produce meaningless results without caution. There is a risk to place in the same data set quizzes in which the learner reasons for 5 seconds (a memory question) alongside quizzes where they work 50 minutes (e.g. computing the noise ratio in a complex system and choosing among one of the proposed values). There is a risk to crunch as if they were similar scores from individual exercises, and even more an issue if team task scores are included as well. The need to describe structures can be addressed with existing taxonomies from researchers such as Bloom, Gagné, and D’Hainaut, using metadata standards (e.g. IMS LD) or with the proposed modeling language.

Learning analytics to support students: Enabling automated interventions

Carolyn P. Rosé, Yohan Jo, Gaurav Tomar, and Oliver Ferschke

An important research problem in learning analytics is to expedite the cycle of data leading to the analysis of student needs and the improvement of student support. On the basis of the importance of social interaction in learning, this work proposes a pipeline that includes data infrastructure for a common representation of social interaction data from multiple platforms; a probabilistic sequence model to analyze the effects of social connections on students' learning paths; and a social recommender system to support students for acquiring positive social capital.

The foundation of computational analytic work is representation of data. Much of our published work in analytics of collaboration in discussion has been focused on either chat data (Howley et al., 2013) or transcribed face-to-face discussion (Ai et al., 2010; Clarke et al., 2013). These can both be represented in a simple, uniform, flat sequence of text segments, each contributed by one speaker. However, when expanding to learning in MOOCs or learning in other online contexts such as opensource communities, the form that the discussions may take becomes more diverse as they are embedded in a variety of platforms. They may even occur simultaneously through multiple separate streams. To that end we offer a publically available data infrastructure we call DiscourseDB (<https://discoursedb.github.io/>), which enables translation of data from multiple streams into a common, integrated representation. The interface level representation is translated down into Discourses, with embedded Discourse Parts consisting of Contributions, which may be related to one another through Relations, and which are associated with content that can be associated with Annotations. This common representation enables combining data across communication streams and applying common modeling technologies.

As a concrete example, consider connectivist Massive Open Online Courses (cMOOCs) that include environments like the competency-based learning platform ProSolo, featured in a recent edX MOOC called Data, Analytics, and Learning (DALMOOC) (Rosé et al., 2015). In these environments data is rich and heterogeneous. In ProSolo, for example, student behaviors formally within the environment include follower-followee relations, posting wall notes including updates and goal notes, and commenting on notes. Students also engage in threaded discussions, blog and comment on blog posts, and tweet. These behaviors occur within accounts in other linked online community spaces. In a proof-of-concept using data from DALMOOC, we have transformed data from wall post comments, blogs and blog comments, and Twitter into DiscourseDB, and applied probabilistic graphical modeling techniques to identify typical student learning trajectories that could be supported through social recommendation.

Once the data of interest has been represented in a way that is generalizable across sources, the next step is to model student trajectories, especially as they relate to their observed social connections. This analytic approach enables us to identify opportunities where interventions can positively impact student trajectories. We propose a model that automatically extracts student learning paths composed of discussions across multiple platforms and active social engagement. This model aims to detect the pattern of students' learning paths conditioned characteristics of their social connections in a follower-followee network and thereby inform us of the influence of different configurations within the social space on student behavior. We define a student state in terms of the discussed topics and the document types used for discussions (e.g. forum, Twitter, blog), and identify these states in a bottom-up fashion through an integration of graphical probabilistic modeling techniques. Given data, the model infers a set of meaningful states along with the topics and document types for each state. The learned topics may be interpreted as informing us about students' interests and inclinations and how they evolve over time. The learned states provide insight into the ways students adopt different social interaction practices at different times.

Interpretation of the learned model from DALMOOC revealed an important problem. On the positive side, we saw that students who used the ProSolo affordances for setting learning goals persisted longer in the course, did more hands on practice, and spent more of their time in the environment doing course relevant work. Furthermore, students who chose to follow other students who had set concrete goals for the course, became increasingly engaged in course relevant activity. However, students in both of these categories were few and far between. Most students were found not to take advantage either of goal setting affordances or follower-followee affordances.

Nevertheless, the situation is not hopeless. Data mining technologies again applied over the common representation obtained through DiscourseDB enables a potential solution. A popular method for recommender systems is matrix factorization, which identifies a latent representation that connects recommendations to those the recommendations are made to (Yang et al., 2014). A feature-aware matrix factorization approach is able to combine data about preferences with an arbitrary feature representation that augments the latent state representation with information deemed to be potentially valuable in making the recommendations. In particular, we designed a context-aware matrix factorization model that uses features extracted from students' goal setting behavior as additional features. A corpus based experiment shows that our system can find

appropriate followees who are not only qualified as positive role models but also relevant based on a model of affinity learned by means of a feature aware matrix factorization approach. Thus, our solution involves an analytics enabled intervention made possible through the DiscourseDB data infrastructure.

Learning analytics to support teachers: Regulating teaching practices through analytics in CSCL

Gijsbert Erkens, Anouschka van Leeuwen, Jeroen Janssen, & Mieke Brekelmans

As explained in the introduction to this symposium, Learning Analytics (LA) can be beneficial for multiple stakeholders. In contrast to approaches that directly target students (such as the work by Rosé et al.), in this contribution we consider LA aimed at supporting *teachers* during the phases of diagnosing, intervening, and evaluating students' activities during computer-supported collaborative learning (CSCL).

In collaborative learning, knowledge is constructed through discourse with other students by sharing and discussing resources and jointly building on task products. Teachers play an important role during the problem solving activities of groups of students. They do so by stimulating meaningful interaction between group members and offering support when needed. The support can be needed when groups experience problems with the task content or with the regulation of the task (cognitive or meta-cognitive level). Support may also be needed with regard to the process of collaboration or to the regulation of collaboration (social or meta-social level). Based on a diagnosis of the progress and quality of a group's activities, the teacher has to decide whom the intervention will be aimed at and what type of intervention is most suited (direct instruction, hints, supporting questions, etc.). Lastly, the teacher has to evaluate, again by diagnostic observation, whether the intervention achieved its intended effect. This is a very demanding task because of the number of collaborative groups in a classroom, the time pressure, and the multitude and multidimensionality of information that is needed in this cycle of diagnosis, intervention and evaluation. Our hypothesis is therefore that LA could be supportive to teachers.

Within a CSCL environment, all students' actions are typically logged automatically and in real time. In most CSCL environments these actions represent clicks on task resources, interaction with task products and to communication within groups or more broadly within the class. These logs may serve as input for learning analytics tools and provide teachers with information about groups of students regarding both task and collaborative activities. LA can be used for real-time assessment to support teachers' moment-to-moment decision making in diagnosis, intervention, and evaluation for multiple groups simultaneously within a class (or even multiple classes), thereby possibly supporting adaptive teaching. This means that teachers, instead of having to monitor all activities of all groups separately, can receive a summary or overview of the situation regarding multiple groups more easily, which would aid the teacher in providing timely assessment and support.

Another function LA has the potential to fulfill is to analyze and report characteristics of collaboration that otherwise require diagnosis at multiple time points. Aggregating aspects of collaboration that cannot be reduced to a single event into a visible summary means information about such processes is more easily accessible to the teacher. For example, during computer-supported collaborative assignments, because group communication is logged, LA could provide the teacher with up-to-date reports about collaborative processes that would otherwise be hard to keep track of. When LAs are used for the purpose of supporting teaching practices, choices must be made. In this contribution we consider two of those choices, namely 1) focus and granularity of presented information, and 2) distribution of decision making between the teacher and the learning analytics system. Both choices will ultimately affect the teaching practices of the teacher.

Focus and granularity of information

The first choice we consider is which type of information to use as input to LA, and thus the focus of diagnosis presented to teachers. As mentioned, a common distinction in the field of CSCL is the difference between cognitive, task related, and social aspects of collaboration. A further question is at what level of granularity the information is presented. Information can either be given in *real time* as the collaboration unfolds, or as later as an aggregated measure. Table 1 represents the LA tools we have investigated in our research, categorized by focus and granularity of information.

Concerning real time measures, the Shared Space tool (top left) shows the level of agreement and disagreement within a group discussion in a graph alongside the chat, with right representing agreement and left representing disagreement. Furthermore, horizontal placement of chat utterances similarly indicates a status related to agreement and disagreement as computed based on the presence of discourse markers in the utterances themselves. The Concept Trail (top right) shows the occurrences of task relevant concepts (and synonyms) in utterances of the students in a graph on top of the chat window.

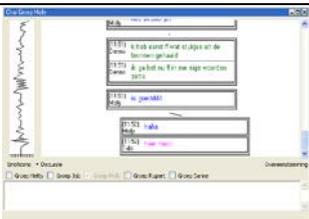
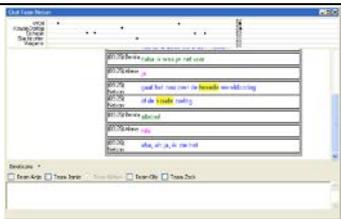
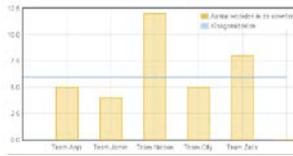
Concerning aggregated characteristics, Participation statistics (bottom left) offers a diagram of proportional participation of group members in chat and productivity tools. Task progress statistics (bottom right) show the relative task progress between the groups in the class as well as their average progress.

In two prior studies, we investigated the effects of LA concerning social activities and cognitive activities on teacher diagnosis and interventions (van Leeuwen, Janssen, Erkens, & Brekelmans, 2014; 2015). When provided with the LA tools the focus of diagnoses or interventions of the teachers were directed by the type of information shown: cognitive or social. Effects of granularity were less clear.

Distribution of decision making between teacher and learning analytics system

LA can support multiple stakeholders. In our research, we have provided LA to *teachers* that represent information on degree of agreement, equal participation, use of task concepts and task progress, but do not give advice or recommendations. In principle, advice and recommendations could automatically be provided to *students* by the learning analytics system as well, as was suggested in the contribution by Rosé et al. A combination of targeting teachers and students is also possible, as demonstrated by the Tan et al. contribution.

Table 1: Focus and granularity of learning analytic tools

Granularity	Focus of LA tools	
	Social activities	Cognitive activities
Continuous characteristics of discussion	Shared Space: Occurrence of agreement and disagreement 	Concept Trail: Use of task relevant keywords 
Cumulative quantity of activities	Participation Statistics 	Task progress Statistics 

Thus, we see that LA tools can be placed on a continuum of how much control and decision making is left to the teacher. On the one end of the continuum are tools that are solely aimed at supporting the teacher in the phase of diagnosis, and the interpretation of the information shown by the tools remains the teacher's purview, as are the intervention and evaluation thereof. Further on the other end of the continuum are LA that are used in scripted scenarios in which tools automatically give alerts or even deliver intervening actions to signaled problems. In the end the question is whether LA tools are used by teachers to regulate their teaching practices or whether LA are used to regulate the teaching practices of teachers.

Learning analytics to support policy: Identifying and fostering 21st century collaborative, critical and connective literacies among diverse learners

Jennifer Tan, Elizabeth Koh, Imelda Caleon, Christin Jonathan, Simon Yang

A major global educational challenge today lies in the question of how schooling systems, policies and practices can more effectively foster 21st century literacies and enhance educational equity among diverse learners—not only at the individual level but more importantly at the collective level. This is with particular sensitivity to the highly-networked, technology-mediated social and learning contexts of modern life. Alongside improved understandings of the dynamic and non-linear nature of 21st century skills and their constitutive socio-interactive processes, policymakers and administrators worldwide are increasingly cognizant of the limitations of conventional modalities of assessment and pedagogic designs. Consequently, many are actively partnering with design-based learning scientists and researchers to explore the affordances that contemporary social learning analytics can bring to bear on this educational imperative of our time.

This contribution provides an exemplification of one such exploratory multi-stakeholder effort in the K-12 schooling sector in Singapore. To this end, we showcase *WiREAD*, a web-based collaborative reading and social learning analytics environment aimed at fostering senior school students' critical reading skills. An important goal of this work is deepening reading engagement levels, and promoting self-regulated and collaborative knowledge construction in the literacy domain, during and beyond formal English class time. Its primary objective is that of motivating and scaffolding students to develop richer dialogue and quality interactions with peers around multimodal texts, thereby enhancing their personal connection to and appreciation for reading as a highly relevant, generative and meaningful social practice.

To achieve this, the techno-pedagogical design of *WiREAD* focuses on 2 key learning affordances: online peer interactions around reading, and the social learning analytics dashboards for students and teachers.

First, the web-based social reading and discussion tool was underpinned by Vygotskian socio-constructivist theories (Stahl, Koschmann & Suthers, 2006) and a Multiliteracies pedagogical framework (Tan & McWilliam, 2009). The Multiliteracies framework comprises 4 essential dimensions of effective 'new literacies' enculturation in learners: situated practice, overt instruction, critical framing, and transformed practice. The micro-level of pedagogical scaffolding scripts comprise 7 critical lenses (message, purpose, audience, assumption, point of view, inference, impact of language/visuals) and 5 critical talk types (*I think that*, *I think so because*, *I agree*, *I disagree*, *I need to ask*), in turn informed by Paul-Elder's (2001) 'wheel of critical reasoning' and our own work on dialogic indicators of collective creativity and criticality (Tan, Caleon, Jonathan and Koh, 2014). These frameworks have served as a meta-schema for guiding students' collaborative critique of texts on *WiREAD* in that students were required to tag each of their comments/replies with 1 critical lens and 1 critical talk type (Figure 3). Each critical lens and critical talk type tag contained a 'popover' that provided students with question prompts and sentence starters, thereby providing students with a constant reference illustrating how each tag could be used to critique texts more deeply.

Second, the social learning analytics affordance of *WiREAD* was designed with the aspiration of providing rich, meaningful and timely formative feedback to students and teachers, so as to help monitor varying levels of socio-interactive reading engagement and progress. In this way, adaptive modifications can be made to learning strategies by students and pedagogical practices by teachers to improve process related learning outcomes. To achieve this, the individual student and consolidated class-level teacher dashboards comprise a range of social learning analytics visualizations that include social networks analytics, discourse and content analytics, dispositional analytics, and achievement analytics (Figure 4).

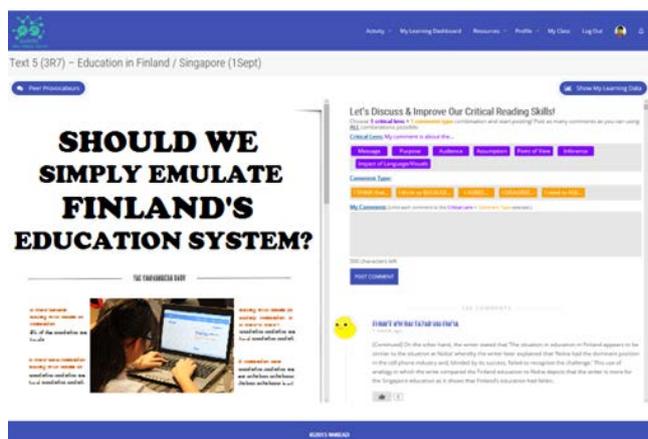


Figure 3. Texts, critical lenses, talk types, and popovers.

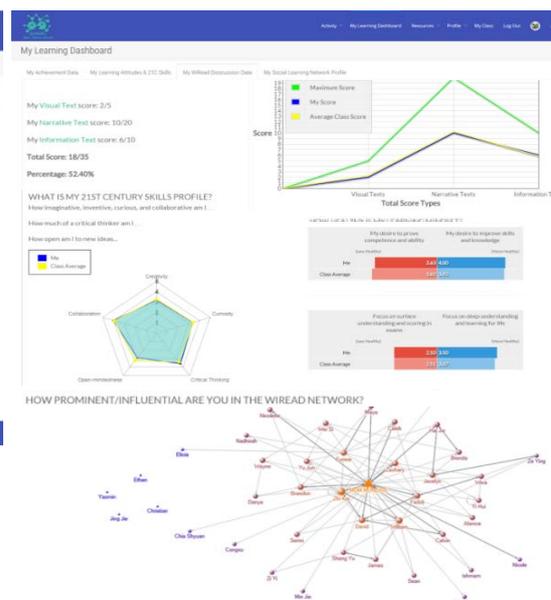


Figure 4. Dashboard visualizations (selected only).

Drawing on a combination of 4 critical stakeholder perspectives—a facilitator-researcher, 3 English teachers, a policymaker and participant students (N=114)—the possibilities of *WiREAD*'s social learning analytics will be highlighted in this contribution. This includes visualizations for 1) making visible and motivating students' agentic development of 21st century collaborative, critical and connective literacies, and 2)

early identification of and support for disengaged and at-risk learners. This will be complemented by a critical discussion of the pedagogical paradoxes and complexities that accompany these possibilities. We situate this discussion against the backdrop of a higher call by policymakers and funding stakeholders to engender broader adoption, translation and diffusion of 21st century technology-enhanced social learning and analytics innovations such as *WiREAD* beyond one ‘seed innovation’ school within a relatively centralized and high-performing Singapore education system. In doing so, we foreground the educational promises and problems that can arise as the ‘rubber’ of well-intentioned learning innovations ‘hits the road’ of entrenched socio-institutional beliefs and practices in mainstream schooling.

Conclusions

Considering work in this young, emerging area at all levels, we see many common concerns, with multiple stakeholders influencing and benefitting from analytics in each. We see common interests in data representation, modeling technologies, and feedback to end users. We see promise of impact, but far to go in terms of serious deployment and adoption. This innovative work promises to challenge and extend what is possible both in the field of Learning Analytics and in the field of Learning Sciences through bridge building. In the discussion we will reflect on the current state of this work and discuss next steps for productive synergy between fields.

Acknowledgements

This work was funded in part by NSF grant ACI-1443068.

References

- Ai, H., Sionti, M., Wang, Y. C., Rosé, C. P. (2010). Finding Transactive Contributions in Whole Group Classroom Discussions, in *Proceedings of the 9th International Conference of the Learning Sciences, Volume 1: Full Papers*, pp 976-983
- Clarke, S., Chen, G., Stainton, K., Katz, S., Greeno, J., Resnick, L., Dyke, G., Howley, H., Adamson, D., Rosé, C. P. (2013). The Impact of CSCL Beyond the Online Environment, *Proceedings of the Computer Supported Collaborative Learning (CSCL) Conference 2013, Volume 1*
- Howley, I., Mayfield, E. & Rosé, C. P. (2013). Linguistic Analysis Methods for Studying Small Groups, in Cindy Hmelo-Silver, Angela O'Donnell, Carol Chan, & Clark Chin (Eds.) *International Handbook of Collaborative Learning*, Taylor and Francis, Inc.
- Paul, R., & Elder, L. (2001). *Critical Thinking: Tools for taking charge of your learning and your life*. Prentice Hall, Upper Saddle River, NJ.
- Rosé, C. P., Ferschke, O., Tomar, G., Yang, D., Howley, I., Alevin, V., Siemens, G., Crosslin, M., Gasevic, D. (2015). Challenges and Opportunities of Dual-Layer MOOCs: Reflections from an edX Deployment Study, *Proceedings of the Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 2*
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, 2006, 409-426.
- Tan, J. P-L., Caleon, I. S., Jonathan, C. R., & Koh, E. (2014). A dialogic framework for assessing collective creativity in computer-supported collaborative problem-solving tasks. *Research and Practice in Technology Enhanced Learning*, 9, 3, 411-437.
- Tan, J. P-L., & McWilliam, E. (2009). From Literacy to Multiliteracies: Diverse Learners and Pedagogical Practice. *Pedagogies: An International Journal*, 4, 3, 213-225
- Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education*, 79, 28–39.
- Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers & Education*, 90, 80–94.
- Yang, D. & Rosé, C. P. (2014). Constrained Question Recommendation in MOOCs via Submodality, *Proceedings of the 2014 ACM International Conference on Information and Knowledge Management*.

Full Papers

Scaling Studio-Based Learning Through Social Innovation Networks

Natalia Smirnov, Matthew W. Easterday, and Elizabeth M. Gerber
nsmirnov@u.northwestern.edu, easterday@northwestern.edu, egerber@northwestern.edu
Northwestern University, Delta Lab, Ford Design Center

Abstract: In this paper, we articulate a model for scaling Studio-Based Learning (SBL) via a *Social Innovation Network (SIN)* – a distributed community of design studios collaborating to solve social problems. We report findings from a case study of one SIN called Design for America (DFA), using methods of “ethnography of infrastructure” (Star, 1999) that combine interviews, surveys, and analysis of members’ communication on various channels to understand the ways SBL can be orchestrated as a distributed learning community. We argue that principled design and use of cyberlearning tools and organizational routines can foster sociability and trust among members and promote routing of resources across the network, thereby alleviating orchestration challenges and infrastructuring a more effective environment for innovation and design learning.

Keywords: design, innovation, networks, orchestration, studio-based learning, cyberlearning

Introduction

To solve society’s most pressing crises such as climate change, refugee displacement, and access to healthcare, 21st century education should prepare social innovators, people who can recombine resources in creative ways in order to tackle complex, ill-structured problems (Trilling & Fadel, 2009). However, innovation is difficult to learn, as it depends both on acquiring skills of creative problem-solving, and access to a favorable *innovation ecosystem*—the resources, skills, collaborators and experts who can guide and support social innovation work (Gundry et al, 2011). Learning sciences is well positioned to improve innovation learning through research on project-based learning, distributed cognition, and adaptive expertise (Blumenfeld et al, 1991; Hutchins, 1995; Hatano & Oura, 2006). Innovation learning on the other hand is a fruitful discipline for exploring the most interesting of learning sciences inquiries: the integration of STEM, arts, and social policy domains within pressing, real-world problems; and the design of learning environments and technologies to support authentic, expert-guided communities of novice innovators.

There is a timely opportunity to design and mobilize cyberlearning tools to support innovation learning through: (a) scaffolding studio-based learning via affordances of computer-mediated technologies; and (b) connecting design studios across institutions to facilitate sharing of resources, interdisciplinary collaboration and knowledge building. In this paper, we describe how the design studio model offers a promising prototype for an online learning community to support innovation, what we call a *social innovation network* or *SIN* (Gerber and Easterday, 2015). We develop our proposal through a case study of Design for America (DFA), a network of 29 university-based studios working to create local and social impact through interdisciplinary design. Using an “ethnography of infrastructure” approach (Star, 1999), we analyze the key features of this social innovation network and the ways different practices and technological systems support and constrain accomplishment of its unique learning goals.

Studio-based learning

The studio model or Studio-Based Learning (SBL) is a promising approach to designing learning environments that can promote both deep disciplinary learning and creativity (Sawyer, 2012). Characterized by open-ended iterative problem solving, expert coaching, and a culture of sharing and critique, the studio model is especially well suited for supporting complex, project-based forms of learning like innovation. While prevalent in creative domains such as art and architecture, and increasingly utilized in novel informal STEM learning environments, the studio model has not been extensively researched in the learning sciences (Sawyer, 2012). Even less attention has been given to the opportunity to use digital and networked tools to scaffold and scale SBL.

While SBL might be organized similarly to and incorporate features of other learning arrangements, such as classroom, apprenticeship, or informal learning (Lave & Wenger, 1991; Ito et al, 2009), recent ethnographies of studio-based learning (SBL) have found that the studio has a particular set of norms as a community of practice, where students are expected to: (a) iteratively generate and refine design solutions by incorporating peer and instructor feedback; (b) frequently communicate design ideas visually and verbally, and

(c) collaborate with peers to give and receive help in achieving learning goals (Cennamo et al, 2011). Studio instructors scaffold these activities through assignments that constrain the complexity of problems, coaching novices through expert feedback, and explicitly reminding students of these cultural norms during critiques. Additionally, Sawyer (2012) found that the studio curriculum thrives on project complexity, engaging learners in problems that require multiple and diverse solutions. Working through this complexity requires continuous externalization of students' insights, initial directions, and partial solutions, exposing them to expert and peer feedback. Thus, the SBL model requires a culture of open sharing, capacity for self-regulation, collaboration, and ongoing iteration. These unique practices make the design studio an especially fruitful context for learning social innovation. However, these distributed and collaborative practices are difficult to scale, because they depend on a complex orchestration of diverse resources and stakeholders, in the context of uncertain and ambiguous problems.

Social innovation networks

Online learning communities are increasingly being used to increase access to resources, solve orchestration challenges and help make classroom learning more authentically aligned to real-world contexts (Bruckman, 2006). But different online platforms afford different social and learning interactions. Online learning communities should reflect the goals and organizational structures of prototypical learning communities, whether that prototype is a classroom, a tailor apprenticeship, or a neighborhood samba school (Bruckman, 2006). Our goal is to create a widely adopted cyberlearning environment that will support innovation. To that effect, we have been developing the Loft (<http://loft.io>)—an authentic online learning community based on the studio model. This platform enables learners and instructors to share progress, exchange feedback, collaborate with each other on complex real-world problems, and develop transferrable skills by participating in scaffolded design challenges. While cyberlearning tools can enhance innovation learning within a single studio context by providing a system for project documentation and resource exchange, internet technology can also enable us to take the affordances of studio-based learning further, by connecting people, resources, ideas, and feedback exchange *across studios*. Thus, we are also developing a new genre of a design learning community that is dependent on radical and distributed connectivity. We call this new genre a *social innovation network (SIN)*.

We define a *social innovation network (SIN)* as a *technologically connected, distributed community of design studios working collaboratively to solve social problems and share design resources*. A social innovation network echoes the vision of a *collaboratory*, developed in the 1950s by William Wulf of National Science Foundation. Enabled by networked technology, a collaboratory is “a center without walls, in which the nation’s researchers can perform their research without regard to geographical location—interacting with colleagues, accessing instrumentation, sharing data and computational resource, and accessing information in digital libraries” (Wulf, 1993). In the case of a SIN, the network-enabled collaboration supports the design of social innovations instead of scientific research.

In this paper, we study a specific SIN called Design for America (DFA). DFA is a network of extracurricular studios at 29 universities that brings together interdisciplinary teams of students, faculty, professionals, and alumni to solve real world problems through human centered design. Universities host on-campus DFA studios where student teams work on innovation projects throughout the academic year, by following a systematic design process provided in a downloadable DFA Process Guide and detailed on the Loft platform through interactive challenges. University students who choose to participate in a DFA studio form project teams with other studio members, identify challenges affecting their local community, such as reducing hospital-acquired infections or reducing water waste in cafeterias, and work over the course of a year to understand user needs, ideate, prototype, test, and implement solutions. In the process, they partner with community organizations, and receive coaching from local design professionals, support from more experienced DFA mentors at the national office (5 full-time DFA staff), and regular feedback from their peers on campus. Because the studios are self-organized and participation of all the stakeholders is voluntary, the organization of the studios is very loosely structured and adapted to meet the culture and rhythm of the respective college campuses. Rather than mandating a particular set of administrative roles and activities, the five full time DFA national staff provide ongoing support to the different studios’ emerging challenges and needs through bi-annual campus visits, regular video conference check-ins, email newsletters with tips and suggestions, and a yearly Leadership Studio that brings together studio leaders from all the DFA campuses for one week to learn the design process, share best practices, and troubleshoot common studio problems. Since this community is distributed across the entire country, DFA leaders and members heavily rely on digital communication tools, including Google docs, Slack and the Loft.io platform, to

connect and collaborate with each other and support studio-based learning. Our study of DFA was designed to answer the following questions:

1. How does DFA currently orchestrate distributed Studio-Based Learning (SBL)?
2. What are the affordances and constraints of existing communication and collaboration tools and practices DFA leaders and members use for realizing the goals of the distributed SBL community?
3. Which principles might we use to better design socio-technical systems (jointly optimized organizational routines and cyberlearning tools) to facilitate learning and social innovation in this type of learning organization?

Methods

This inquiry into the communication and collaboration practices of DFA members and leaders followed an “ethnography of infrastructure” (Star, 1999) approach—examination of the “boring things” that reveal the background workings and assumptions of a complex socio-technical system. We followed this approach because we noticed that even though we had designed and implemented a platform specifically for Studio-Based Learning (e.g. the Loft), DFA members were using other tools in addition (e.g. Facebook, GroupMe, Google Docs, etc) to connect and collaborate with each other, both within and across studios. Thus, the existing cyberlearning platform was not adequately meeting the needs of members at the *network level*. Members and leaders of DFA repeatedly expressed desire for more resource sharing and coaching specifically at this scale. We wanted to understand why DFA members chose to use certain communication and collaboration tools, how they used them (e.g. via what routines) and what those tools afford and constrain in terms of learning and collaboration. Studying both tools and routines helps us deepen our understanding of this learning community and generate design principles for developing new tools and routines to better support the goals of this system. We believe that while specific technological designs can afford particular practices and interactions, technologies themselves do not determine social behaviors. Rather, social behaviors are mediated by users’ mental models, circumstances, goals and tactics (Geels, 2004, Fischer, 2007). As such, we emphasize the need for joint optimization of both organizational routines and technologies, and for attending to continuous emergence and adaptation mechanisms within a socio-technical system (Trist, 1981; Bruckman, 2004).

Data collection and analysis

We studied the network-level practices, strategies, and goals of DFA in three ways. First, we collected and analyzed communication and collaboration practices of network members at the network level, or *across* studios. For example, we collected posts on DFA’s Student Network Facebook Page, cross-studio discussions on Loft, and other platforms and channels (GroupMe, Text, GoogleDocs). Second, we conducted interviews with network leaders about their mental model of the DFA network and its goals, and the tools and routines they use to support DFA studios to support their project teams and connect to each other. Third, we surveyed 72 studio leaders about what supports are needed for managing their studios and the current communication and collaboration tools in use by studio members.

This data collection was iterative, with each phase of collection and analysis informing the next round of questions. For example, analysis of DFA social media practices prompted us to consider the network leaders’ decision to setup and use particular tools and not others. We were interested in this group in particular because they are primarily responsible for orchestrating learning at the network level, whereas studio leads and faculty coaches support individual studio learning. In interviews, we first asked the five DFA leaders (4 full-time fellows and 1 administrator) to draw a “map of the DFA community” while talking aloud about the different actors they were illustrating. This drawing served as a useful representation for the rest of the interview, providing a reference for the multiple stakeholders of the network and illustrating the ways different digital tools were used to communicate and collaborate by certain members of the DFA community and not others. We analyzed the interview transcripts using a grounded theory approach, first open coding transcripts to generate emerging themes, then specifying a set of descriptive codes, and applying these codes systematically to the entire data set (Charmaz, 2014). The coding was performed by the first author and discussed and refined in weekly research meetings with the rest of the research team. This analysis revealed that individual DFA leaders’ communication practices were idiosyncratic and relied primarily on anecdotal evidence rather than analytics or extensive understanding of member needs. This finding prompted us to conduct need-finding interviews and develop a survey to understand the needs of DFA student members and the digital tools they already use to communicate and collaborate with each other. The survey was completed by 72 studio leaders at the annual all-network training conference. We

analyzed the digital tool use and needs of DFA members, comparing them with our findings from DFA staff interviews. Finally, we continued to collect examples of cross-studio communication on various platforms, triangulating our emerging thematic categories with new evidence (Yin 2013; Stake, 1995). As we advanced in our analysis, we compared empirical findings to existing theoretical constructs (Snow, Morrill & Anderson, 2003), and created explanatory matrices and networks (Miles, Huberman, Saldaña, 2013) to model relationships between various parts of the system (See tables 1 and 2).

Findings

How does the Studio model scale?

Studio-based learning (SBL) is characterized by a set of norms and practices including project complexity, public critiques, and expert coaching (Sawyer, 2012; Cennamo, 2011). These norms are realized in the *DFA studio-level organizational practices* and *afforded by the cyberlearning tools*, including the Loft (Easterday et al, 2015; Rees Lewis et al, 2015). In this study, we have zoomed out to examine the learning infrastructures at the network level that allow DFA to realize SBL model across institutions to support studios to carry out design and innovation work (see Table 1). In the table below, we present a concise summary of this 3-level dynamic, from the SBL theoretical model to studio practices to network organization. The two columns on the right present the findings specific to this study. The *DFA goals and practices* column presents examples of network-building activities from our ethnography, while the rightmost column summarizes *organizational needs to inform future cyberlearning tools and routines*. These needs combine suggestions and wishes of DFA network leaders and members, articulated in interviews and surveys and ideas generated by the research team through the analysis of the data. Some potential ideas for adapting existing practices and creating new tools and routines are listed in parenthesis.

Table 1: Studio-Based Learning (SBL) Model at the Studio and Network Levels

Features of the Learning Environment	Studio-Based Learning (SBL)	How DFA orchestrates SBL <i>within studios</i>	Tools and routines to support SBL <i>within studios</i>	DFA goals and practices to orchestrate SBL <i>across studios</i>	Organizational needs for tools & routines to support distributed SBL
type of problem solved	Students work on complex real-world problems	DFA Teams partner with local community organizations to work on real-world dilemmas in healthcare, education, etc.	Tools to manage project complexity (Scoping Wheel, Design Canvas)	Studio leads confront many real-world challenges, such as member recruitment, leadership organization, managing relationships with partners, working with coaches, etc.	Help studio leads anticipate, externalize and respond to common studio challenges (Studio guides, FAQs, training, mentoring)
support provided by instructor(s) or system	Scaffolding through constrained assignments, expert modeling and coaching, and prompts	Studio leads guide teams through phases of the design process, using DFA process guide and LOFT.io tutorials	Design Process Scaffolds (Guides, Design Challenges, Scoping Wheel, Coaching Stands, regular mentoring calls)	National staff support studio leads to facilitate learning within their studios (e.g. recruitment, critiques) but must manage: (a) varied studio calendars; (b) busy student schedules; (c) unreliable communication practices; and (d) limited time of 5 staff	Just-in-time reminders and prompts for how to facilitate a studio-wide activity (email newsletters; social media reminders; Group Calls)
demands on learners	Students self-regulate goal-setting and progress management	Project Teams consistently meet independently to work on projects	Tools and routines to support goal setting and planning (to-do lists, Workbench, Calendar)	Studios must adapt process and governance to their own campus structure and culture; studio leads might have little to no experience leading an organization	Support studio goal planning and flexible adaptation to campus culture. (e.g. Group Calls, Campus Visits)
organizational	Regular	Studio crits allow	Tools and routines	Annual DFA Leadership	Make resources and

routines	exhibition/ visualization of progress & problems (Critiques or Crits)	teams to share progress and receive feedback	to support progress sharing and feedback with coaches and mentors (Virtual Stands, feedback features on Loft)	Studio provides opportunity for studio leads to share best studio management practices and tools	progress visible across studios more regularly (Cross-studio meetups; email updates; Tool Expos; Group Calls; Facebook/Twitter Updates; Studio Dashboard)
role of peers and mentors	Reliance on peer collaboration and expert mentoring to achieve learning goals	DFA projects are collaborative; teams are supported by coaches and faculty advisors via critiques and stands	Tools and routines to foster peer support and help-seeking (Cheers, Feedback Routines, Stands)	National staff want DFA members to connect to each other directly to share resources and provide help and activate the power of the DFA alumni network for mentorship	Facilitate help seeking, giving and information routing across studios based on similarities (Help Forum, Tagging, Cross-studio meetups, Networking activities, Recommender system); Attract and retain professional mentors
timescale of work	Cycles of Ongoing Iteration	DFA are teams encouraged to continue working on their projects after first year	Non-linear tools for authoring and archiving multiple solution prototypes (Challenges, Feedback)	Network practices change as network grows; Members join studios on a rolling basis, require onboarding. Studio leadership turns over as students graduate.	Crowdsource curriculum authoring and revision, preserve traditions and train new leadership (Wiki, Leadership Ladder)

DFA orchestrates Studio-Based Learning – guided, collaborative, iterative, real-world problem-solving – both within and across studios (Table 1). Running a DFA studio is its own form of a complex, real-world challenge that requires ongoing problem refining, help-seeking, iteration, and just-in-time learning. Just like in an SBL context, there is no one right answer to the problem of “how to run a studio” and expertise is distributed across the community of other novice learners and volunteer coaches. To support this SBL network, the organizational leaders need to “scale”–replicate and adapt–the practices of SBL instruction to 29 universities, 120+ projects and 900+ active student members. The limited capacity of 5 full-time staff to effectively support all these stakeholders and their needs presents significant orchestration challenges. From our interviews with network and studio leaders, we found that the primary goals and needs of this SIN are to connect network stakeholders with each other to promote resource sharing, help-seeking and giving, and orchestrate social and pedagogical support for the studios and teams to follow the design process. Leaders and members of the network adapted various information and communication technologies (ICTs) to orchestrate resource sharing and collaboration across studios. In our interviews, DFA national staff mentioned 37 ICTs that they either currently used or considered using for DFA activities, including social media, video conferencing, project management and collaboration tools. In our survey of DFA studio leaders, students mentioned 18 separate digital platforms that they use on a daily basis. Below we analyze the affordances and constraints of the most frequently used ICTs for supporting this Social Innovation Network.

How do ICTs support and constrain SINs?

To illustrate the use of ICTs to orchestrate distributed SBL, consider the following vignette of an hour-long video call on Adobe Connect with leaders from the 29 DFA studios across the country. Three group calls happened at the same time, with several studios per group, facilitated by 1 national staff member each. Two weeks prior, studio leads received an email asking them to fill out a Google Form and list their studio’s upcoming events and current challenges. The results were organized by the DFA national staff in a Google Spreadsheet, based on similarity of the goals, the kinds of help each studio needs, variability of experience, and “diversity” of studios–one of the goals of this call is make studio leads talk with leaders from different universities than they did last month. 10 participants from 3 different time zones joined one of the video calls. 3 of these attendees had trouble with their microphones, so they had to type in the chat window instead of talking. One student came in late; two participants had to leave early to go to other meetings. The students introduced themselves and discussed how each of their studios was dealing with various types of design feedback – from studio peers, professional mentors,

and target users. Studio leaders shared exemplary practices (e.g. “a workshop on presenting your work” or “we do a weekly pin-up session with professionals”), others surfaced concerns (“are we losing sight of users?”). The DFA staff member facilitated the conversation, took notes in a Google Doc, and tried to acknowledge and respond to the contributions from students typing in the chat while others spoke out loud. With less than 10 minutes left, the facilitator asked each person to name their takeaways and reflect on how the call went. Participants were surprised by the diversity of practices across campuses and wished to know more about each studio’s different projects. Then the facilitator said “let’s take a screenshot!”, counted “1-2-3” and pressed a several keys to capture a still of her computer screen. As the call ended, the DFA leader sighed, “I’m always so tired after these!” and then immediately, “I have to tweet this now.”

The example above captures only some of the tools DFA members currently use to orchestrate distributed, student-led studio-based learning. One of the most prominently used tools in this community is Facebook Groups: there is a Facebook Group for each DFA studio, one just for studio leads, one for the entire network, one for alumni, and a group for DFA-ers interested in traveling together. Facebook was also mentioned as the most widely used tool in the DFA member survey. The group messaging app GroupMe is used by local team members to manage their collaboration and by studio leads to keep in touch after annual and regional meetups. A few studios have recently adopted the platform Slack as a way to manage communication between teams and mentors within a studio. This combination of ICTs is tactical—an ad-hoc assemblage of available and familiar resources to achieve emerging goals (De Certeau, 1984). While DFA network uses ICTs in ways that are creative, adaptive and practical, our study also helps to highlight the limitations of these tools, as they have not been designed explicitly or strategically to support distributed studio-based learning. For example, DFA members complained that the Facebook feed doesn’t allow easy searching, archiving or re-organizing of information that could be potentially useful to future studio leaders or network members. Similarly, Google Docs lack standards for organization and archiving. As one of our informants said, “My personal hell is being lost in someone else’s GoogleDoc.” Despite its capacity to facilitate virtual face-to-face meetings, much of the time during teleconferences on Adobe Connect may be spent managing internet connections or software issues. In the table 2 below, we summarize the features and qualities of ICTs most commonly used by DFA members that were addressed as salient in our interviews and observations.

Table 2: Features and Affordances of ICTs for supporting distributed SBL

	Features and affordances of ICTs	Facebook Groups	Group Me	Google Docs	Google Hangout	Adobe Connect	Email	Loft	Slack
<i>Ease of onboarding</i>	Familiar to participants	✓	✓	✓	✓		✓		
	Mobile App	✓	✓		✓		✓		✓
<i>Intra-team collaboration</i>	Collaborative Authoring / Production			✓				✓	
	Project-Management Tools							✓	
	Shared/cloud storage of documents	✓		✓				✓	✓
<i>Network wide resource sharing</i>	Many-to-many conversation	✓	✓		✓	✓			✓
	User- and category- tagging, sharing	✓							✓
	Feedback, voting, polling, annotation	✓		✓				✓	
	Alerts, notifications	✓	✓				✓	✓	✓
<i>Affective Infrastructure</i>	Emojis, reactions, photo posting	✓			✓				✓
	Virtual Face2Face (video/audio)				✓	✓			
	“Less Formal”	✓	✓		✓				
<i>Organizational Strategy</i>	“More Professional”					✓	✓	✓	✓
	Analytics	✓					✓		

	Standardized archiving								✓	
--	------------------------	--	--	--	--	--	--	--	---	--

We organized ICT features into 5 major categories: (1) ease of onboarding, (2) intra-team collaboration features, (3) inter-team/network-wide information routing and resource sharing, (4) affective infrastructure, and (5) functions that support organizational strategy rather than solve an immediate problem. As this analysis demonstrates, there is no one perfect platform that meets all of the needs of this community. While Facebook emerges as extremely versatile with regard to several variables (familiarity, collaborative features, affective expression and so on), it lacks project management tools to facilitate teamwork or archiving tools to organize useful information developed in discussions for future use. In addition, because DFA is both extra-curricular and proto-professional—students volunteer to join DFA and lead studios, but often view it as a stepping stone to potential design jobs—it must balance being fun and social, while also supporting professional skill development. Part of our research project is to articulate design principles for developing and assembling ICTs to support innovation learning. Intentionally designed cyberlearning tools can help to capitalize on affordances of existing technologies while enabling more productive collaboration and resource sharing among SIN members. Next, we highlight two key principles for designing cyberlearning tools to support SBL at the distributed network level: *affective infrastructure* and *information routing*.

Design principles for socio-technical systems to support SInS

Affective infrastructure

Across the different preferred modes of collaboration and communication, network members placed high importance on affective modes of expression (e.g. emojis, pictures, reactions, and video calls). Posting silly photos and videos and using emojis in their conversations helped them build more sociability and trust with their peers, which facilitated other kinds of knowledge building and design resource sharing. The *affective infrastructure* is thus a critical feature of the SBL environment. As designers, we must attend to the ways the learning environment (or technology) is affectively experienced—as fun or formal, as exciting or boring, institutional or personal—and specific infrastructural features that promote trust and sociability, which in turn foster resource sharing and help-seeking.

Information routing

In a distributed SBL community, or SIN, it is impossible for any member of the network to attend to all the information exchanged between various parts of the network and consider its usefulness for his/her own purposes. *Information routing* (Resnick, 2001) is a form of socially-distributed filtering to manage this problem. For example, when resources or opportunities were posted on the DFA Facebook page, members used the Facebook user tagging feature to alert other network members and bring them into the conversation. DFA staff intentionally organized group calls to facilitate information routing between more and less experienced studio leaders. As designers, we must consider features and routines for information routing, such as the ability to tag other users to alert them of a relevant discussion, or automatic targeted notifications based on user-specified interests or roles.

Discussion

Learning Sciences has a long tradition of designing cyberlearning tools to support project- and inquiry-based learning (Edelson, Gordin, Pea, 1999). Our case study of a distributed studio-based learning community extends this tradition by focusing on a new disciplinary domain (design learning), articulating a new genre of a learning infrastructure (*social innovation network*), and proposing two design principles for building socio-technical systems that enhance SInS. Specifically, we have argued that attending to affective infrastructure and information routing will facilitate the creation of communication channels and resource sharing across a distributed network of studios, thus supporting innovation learning and design. These considerations are likely missing from previous ethnographies of SBL because the small scale and face-to-face interaction of typical studio classes allow affective infrastructure and information routing to be taken for granted (Colyvas & Powell, 2006).

References

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26(3-4), 369-398.

- Bruckman, A. (2004). Co-evolution of technological design and pedagogy in an online learning community. *Designing for Virtual Communities in the Service of Learning*, 239-255
- Bruckman, A. (2006). Learning in online communities. In *Cambridge handbook of learning sciences*.
- Cennamo, K., Brandt, C., Scott, B., Douglas, S., McGrath, M., Reimer, Y., & Vernon, M. (2011). Managing the complexity of design problems through studio-based learning. *Interdisciplinary Journal of Problem-based Learning*, 5(2), 5.
- Charmaz, K. (2014). *Constructing Grounded Theory*. Sage.
- Colyvas, J. A., & Powell, W. W. (2006). Roads to institutionalization: The remaking of boundaries between public and private science. *Research in organizational behavior*, 27, 305-353.
- De Certeau, M. (1984). *The practice of everyday life*. Berkeley: University of California Press.
- Easterday, M., Lewis, D. R., & Gerber, E. (2013). Formative feedback in Digital Lofts: Learning environments for real world innovation. In *AIED Workshops*.
- Easterday, M. W., Rees Lewis, D., & Gerber, E. M. (2015). Crowdsourcing critique. In *CSCL crowd 2015, a workshop at the 2015 International Conference on Computer Supported Collaborative Learning*, June 7, Gothenburg, Sweden.
- Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the learning sciences*, 8(3-4), 391-450.
- Fischer, G. (2007). Designing socio-technical environments in support of meta-design and social creativity. In *Proceedings of the 8th international conference on computer supported collaborative learning*.
- Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research Policy*, 33(6), 897-920.
- Gerber, E., & Easterday, M. (2015). Social innovation networks: Models, outcomes, and challenges. In *Proceedings of the MDW IX*.
- Gundry, L. K., Kickul, J. R., Griffiths, M. D., & Bacq, S. C. (2011). Entrepreneurial Bricolage and Innovation Ecology: Precursors to Social Innovation? *Frontiers of Entrepreneurship Research*, 31(19), 3.
- Hatano, G., & Oura, Y. (2003). Commentary: Reconceptualizing school learning using insight from expertise research. *Educational Researcher*, 32(8), 26-29.
- Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive science*, 19(3), 265-288.
- Ito, M., Antin, J., Finn, M., Law, A., Manion, A., Mitnick, S., ... & Horst, H. A. (2009). *Hanging out, messing around, and geeking out: Kids living and learning with new media*. MIT press.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook*. SAGE Publications, Incorporated.
- Rees Lewis, D., Easterday, M., Gerber, E. M., (2015). Supporting Project Scoping: The Scoping Wheel. In *Proceedings of the Harvey Mudd Design Workshop IX*. May 28-30, Claremont, CA, US.
- Resnick, P. (2001). Beyond bowling together: Sociotechnical capital. *HCI in the New Millennium*, 77, 247-272.
- Snow, D. A., Morrill, C., & Anderson, L. (2003). Elaborating analytic ethnography: Linking fieldwork and theory. *Ethnography*, 4(2), 181-200.
- Stake, R. E. (1995). *The art of case study research*. Sage.
- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 377-391.
- Trilling, B. & Fadel, C. (2009). 21st century skills: learning for life in our times. Jossey-Bass: San Francisco.
- Trist, E. (1981). The evolution of socio-technical systems. *Occasional paper*, 2, 1981.
- Wulf, W. A. (1993). The collaboratory opportunity. *Science*, 261(5123), 854-855.
- Yin, R. K. (2013). *Case study research: Design and methods*. Sage publications

Acknowledgments

The authors thank DFA leaders and members who participated in interviews and surveys analyzed for this study. Thanks to Daniel Rees Lewis, Rob Calvey and Stacy Klingbeil for help with data collection, Alex Sher and Andre Mohring for help on data transcription and analysis, and members of Northwestern Delta Lab for ongoing feedback, moral support and inspiration. This work was supported by the National Science Foundation grants IIS 1320693/1217225 and Venture Well.

Multiple Legitimate Language Games in Family Serendipitous Science Engagement

Dana Vedder-Weiss, Ben-Gurion University of the Negev, Beer-Sheva, Israel, vedderwe@post.bgu.ac.il

Abstract: In this paper, I explore the nature of family-based Serendipitous Science Engagement, analyzing a case of three children and a mother engaging with bugs they discovered in their yard. Employing linguistic ethnography methods, the analysis reveals how the absence of designated goals afforded legitimate multiple "language games", which supported each other. Guided participation entailed authentic sense-making modeling and gradual participation in the sense-making game, without abandoning other games. Theoretical and practical implications are discussed.

Keywords: family learning, informal science learning, language game, linguistic ethnography, parent-child interaction, science learning, self ethnography, socio-cultural theory

Introduction

Learning takes place everywhere: in designed environments such as museums and schools, as well as in unstructured ones such as family meals (Tal & Dierking, 2009). Whereas the vast majority of research on out-of-school learning focuses on designed environments, research on serendipitous, spontaneous undesigned learning is scarce. Few studies have documented serendipitous learning, demonstrating how it may afford deep engagement in exploration (Crowley & Jacobs, 2002), skill improvement (Ochs, et al., 1992), and knowledge co-construction (Goodwin, 2007), resulting in the development of science learning pathways (e.g., Bricker & Bell, 2014). However, these studies did not focus on the unique nature of serendipitous science engagement (SSE), which I define as free-choice engagement in an undesigned environment, i.e., voluntary engagement in an environment that does not entail prior educational scientific goal.

To better support SSE, or at least not interrupt it, we need to understand it better. In this study, I provide an account of family-based SSE, obtained through self-ethnography, and discuss how we might explore and conceptualize it, employing language games and guided participation frameworks.

Theoretical framework

Learning as participating

Framed by a socio-cultural perspective, Rogoff and colleagues (1993) suggested viewing learning as a process of "guided participation" in which an experienced member of the community supports novices' learning by bridging between the familiar and the novel, connecting observations, information, behavior, and meanings in a shared endeavor. "Guidance" refers to the "direction offered by cultural and social values, as well as social partners"; "participation" refers to "observation, as well as hands-on involvement in an activity" (Rogoff, 2008, p. 60). Guided participation extends the Vygotskyian notion of *zone of proximal development* by emphasizing tacit knowledge and communication rather than only explicit, didactic, verbal academic instruction; it includes "not only the face-to-face interaction, which has been the subject of much research, but also the side-by-side, joint participation that is frequent in everyday life" (Rogoff, 2008, p. 60).

Guided participation in families (and elsewhere) entails modeling. In modeling, participants observe more experienced members (e.g., parents, siblings, peers) exhibiting a model of "expert performance", which includes internal cognitive processes and tacit knowledge, as well as motivational and emotional processes and behaviors (Schoenfeld, 1985). The situatedness of learning is another important aspect of guided participation (Brown, Collins, & Duguid, 1989), in which people are engaged in an authentic activity facing ill-structured tasks and real-life problems. SSE is inherently situative, as it entails real-life, authentic contexts and unstructured tasks.

In this study, I look into processes of guided participation to explore the nature of SSE. In doing so, I also aim to add to the theory of guided participation in two ways. First, Rogoff et al. (1993) studied child-parent guided participation focusing on shared endeavors or "mutual engagement in the same agenda" (p. 68). They analyzed tasks that entailed a designated set of practices and goals (such as correctly operating novel objects or dressing). However, in everyday situations (and in particular in serendipitous ones), there is not always a designated goal, and therefore the agenda or the relevant practices are not always shared. In this study, then, I ask: what is the nature of guided participation when participants do not share the same agenda and may pursue disparate

goals? Second, Rogoff (1993, 2008) focused on non-academic activities, while I explore how guided participation in informal settings advances children's academic participation (i.e. their scientific practice).

Language games

Building on and expanding Rogoff's (2008) notion of "purposes involved in shared endeavors" (p. 64), I analyze guided participation through the lens of "language games". "Language games" are frames we ascribe to an activity to make sense of it (Wittgenstein, 1953). Language games are mostly implicit; we are not aware of it and of the "rules of the game" that guide us. Nevertheless, the game defines our goals, roles, legitimate courses of action, symbols' meaning, tools, and expectations (Sfard, 2001). The same situation may constitute several games, negotiated from moment to moment as the interaction unfolds. For example, in the classroom, a teacher may frame an activity as "science learning", while students may frame it as "maintaining social status"; The game "science learning" has the objective of developing understanding and skills, and entails teachers' and students' roles; rules such as "teacher asks, students answer"; scientific terms as symbols, and tools such as textbooks. The "maintaining social status" game may have the objective of appearing "cool"; the roles of the "popular kid" or the "geek"; rules such as "exhibit disinterest"; slang as symbols, and tools such as forbidden Smartphones.

Language games are a fruitful prism for analyzing learning processes. Wertsch (1979) analyzed parents assisting their children in solving a puzzle and argued that a child develops as she comes to share the adult's language game. It was also used to explore classroom interaction and student engagement and sense-making processes (Fleener, Carter, & Reeder, 2004; Sfard, 2001).

Methods

Research approach

Studying family-based SSE is inherently challenging: How can a researcher capture serendipitous activity as it unfolds? If s/he cannot anticipate when and where engagement will occur, how can s/he be in the right place at the right time to observe and record it in real time, as it evolves? To confront this challenge, I have conducted a "self-ethnography", documenting SSE in my own family. In self-ethnography "the researcher-author describes a cultural setting to which s/he has a "natural access", is an active participant...The researcher then works and/or lives in the setting and then uses the experiences, knowledge and access to empirical material for research purposes." (Alvesson, 2003, p. 174). Previous studies exemplify the value of self-ethnography by a parent-researcher, affording access to empirical accounts and aspects of the phenomenon (such as continuity) that might not otherwise have been accessible to research, as well as familiarity and pre understanding of the context (Bissex, 1985; Long, 2004; Yoon, 2012). While the challenge of ethnography is usually to "break in" to the lived experiences of others and develop "closeness" to the "native" perspective, the challenge in self-ethnography is to "break out" of the familiar, implicit, and taken for granted. Thus, while in ethnography, the researcher needs to "make the strange familiar", in self-ethnography, s/he needs to make the familiar strange (Alvesson, 2003). To accomplish this, I used linguistic ethnographic methods, which allow the researcher "to get analytic distance on what's close-at-hand" (Rampton, 2007, p. 590).

Context and data collection

I collected the data during 2012-2013, when my family and I lived in Australia (due to my husband's sabbatical leave), and I homeschooled my sons (aged 8.5 and 11). Instead of attending school, they spent most mornings with me, engaging in planned educational and recreational activities (such as math learning, cooking, visiting museums, and cycling), as well as serendipitous activities (such as the one analyzed here). I collected the data by intensive observations, participant-observations, audio-records, and a journal, all documenting my children's engagement with science content and practice. In total, during that year, I audio-recorded 305 engagement episodes (ranging from 30 seconds to 66 minutes). The episodes spanned a variety of everyday settings, such as family meals, car rides, strolling the streets, camping, browsing the web, cooking, and visiting a skate park. All family members were (and are) aware of the study and repeatedly expressed their consent to participate, i.e., to be recorded, analyzed and reported.

The dual role of a parent-researcher raises methodological challenges, such as: (1) the question of whether my awareness of the research induced the children's engagement and thus eliminated the possibility of serendipity. I argue that by adhering to the definition of SSE and closely examining the data, it is reasonably possible to discern SSE from other episodes, even in self-ethnography; (2) the tradeoff between rich, full documentation and authenticity. I argue that it is often possible to identify in the audio recordings when the children were aware of the recording and "acted out" on its behalf, and to account for such acts in the analysis;

and (3) the specificity of the context. clearly the goal is not to generalize but rather to elucidate what SSE may entail in an “extreme case” such as this (i.e., where parents are both scientifically educated and has plenty of time to spend with their children).

In this paper, I analyze one SSE episodes, which in and of itself constitutes an “extreme case” among the other SSE episodes in the dataset: This was the longest recorded episode (66 min), it met the definition of SSE for all participants, it entailed high level of behavioral, cognitive and emotional engagement in both science content and practice, and overall it was particularly fun and exciting.

Analysis

Using linguistic ethnography methods, I repeatedly listened to the data, transcribed it in detail, and brainstormed (on my own and with colleagues) about what was happening and what we found interesting (Rampton, 2007). Then, I used micro-analytic methods to analyze the sequential unfolding of the episode, which included proceeding slowly through the text, asking at each line “What is the speaker doing?”; “Why now?”; “How does this turn of talk respond to what preceded it?”; “What else might have been done here but wasn’t?” and so forth. To account for the temporal dimension of learning, I looked at the textual trajectories (Blommaert, 2005) of discourse across the episode, which involved moving backwards and forwards through time, tracking threads of text, ideas, and concepts, how they were used, when, and by whom.

Initial analysis pointed to the prominence of multiple language-games, directing me to focus the subsequent analysis on two research questions: (1) What games do participants “play”? and (2) How do the games evolve? To answer these questions, I re-analyzed the data using the same method as before, with the addition of systematically and iteratively coding each utterance for its corresponding game(s) and how they evolved, searching for emerging patterns. I conducted this analysis myself, then circled around a draft of the analysis and emerging understandings, and had colleagues acting as “critical friends” review it and comment on it.

The original language of the episode was Hebrew, and I have worked from the Hebrew recording and transcript throughout the analysis. In presenting the case and its analysis, I use the third person to refer to myself (calling myself “Dana” or “Mother”) to represent the analytic distance I have developed over the years that have passed since the data was collected and the numerous cycles of its analysis and presentation.

Data and findings

This SSE episode took place in December 2012. Dana, the mother, was home with her sons, Shahar and Yoav, and with Amit (Shahar's friend). The three children and Dana planned to spend the morning making Pizza. Dana was in the kitchen, while the children hung out in the backyard waiting for her to call them. Suddenly Dana heard them calling. She grabbed the recording device, which was on the table, pressed “Record” and ran outside, finding the children inspecting black and orange insects on a citrus tree, jumping and calling excitedly, “*I’ve never seen this many in my life! ...there’re millions here... Mom, you won’t believe it. There’re millions of ‘em. Here! Here! There! There!*”

For the next 66 minutes, the children were engaged with the insects, with Dana joining them and departing occasionally, leaving the recording device close by. None of them saw those insects before. They observed the insects, collected them into a box, and searched for others in various places in the yard. They discussed topics related to the insects' appearance, behavior, needs, and relations. Dana searched online and found that the insects were Citrus Stink Bugs, which change colors from orange to black as they mature, live on citrus trees, suck their sap, and secrete a repellent substance when distressed. After an hour, the children and Dana turned to the kitchen to make Pizza, as planned, but they continued discussing the experience and the insects.

The analysis revealed that participants were engaged in five prominent language games: Sense-making, treasure hunting, daring, caring and admiring.

The sense-making game

Playing this game, participants construe the activity as one that is about trying to make sense of the phenomenon. This includes trying to specify the bugs (e.g., “*what is it?*”, line18), raising questions (“*what's the difference between the red ones and the black ones?*”, 47), hypothesizing (“*maybe it's a stage in maturation*”, 192), investigating (“*I want to find out what they are and what's their deal*”, 146), exploring (“*we can really explore them like this*”, 481), observing (“*It's possible to look at them through magnifying glass*”, 151), and more.

The treasure hunt game

Playing this game, participants construe the activity as one that is about: (1) searching (as in “*it's fun to look for them*”, 718); (2) finding (“*here's another one. Wow!*”, 30); and (3) pointing and sharing (“*look, a caterpillar*”,

24). Participants maintain this game by expanding the “hunt” in 3 different ways: (1) by expanding the game to additional locales, searching for bugs on other trees and in other parts of the yard; (2) by expanding the game to virtual spaces, searching for pictures of the bug online; and (3) by changing the scale of the game through closer observation of the bugs in the box, searching, finding and pointing out finer features. While such expansion may also serve the sense-making game, at moments its purpose is the mere searching and finding (without attempting any “understanding”).

The daring game

Playing this game, participants construe the activity as one that is about exploring their courage, experiencing the thrill of the extreme, dangerous, and forbidden. Immediately from the beginning and continuing throughout the whole episode, Amit is preoccupied with daring to touch the bugs (e.g., “*I want to take one off [the tree], but I don’t want to touch it*”, 100). He challenges the others to touch it (“*who dares to touch?*”, 102), repeatedly discussing whether they are dangerous (“*it’s probably really dangerous*”, 54) and capable of flying (“*can they fly?*” 15). I relate his repeated reference to flying (15 utterances) to the daring game, as he is “*scared that one will jump on my head or something*” (261) and “*if I touch it, it’ll fly*” (727), and when one indeed flies, he shrieks and yells (and laughs): “*So scary! It flew over me...So crazy...you’d die if it happened to you!*” (1,003-1,008). All three children get very excited when the bugs crawl on each other’s backs. The bugs’ “copulating” behavior repeatedly draws a lot of attention, including the use of foul language (e.g., “*I want them to have sex*”, 588), which I argue is part of the daring game. “Poo” and “Pee” are heard often too, for example regarding how the caterpillar appear (“*It looks like poo*”, 210) or the bugs’ secretions (“*it peed on you*”, 933).

The caring game

While Yoav and Amit are dominant in the daring game, Shahar appears more engaged in the caring game, construing the activity as one that is about taking care of the bugs. Yoav and Amit also participate in this game, and together they prepare a place for the bugs (bringing a box, filling it up with leaves and fruit, and punching holes in the lid), negotiating the conditions that best suit them (e.g., “*we need them to have, to feel comfortable*”, 512), collecting them (“*I want to bring another one*”, 370), keeping them in the box (“*they escaped*”, 897), and empathizing with them (Yo, *poor guy! It’s tough for him, his antenna’s broken.*”, 610).

The admiring game

Playing this game, participants construe the activity as one that is about admiring and getting excited and enthusiastic about various objects. It is manifested through expressions of excitement (such as “wow”, “yo” and “cool”), intonation, laughing, and shrieking. It addresses the quantity of bugs (e.g., “*Mom, you won’t believe it. There’re millions of ‘em*”, 4), size (“*it’s so huge!*”, 992), behavior (“*yo, look! It pooped on me*”, 238), cuteness (“*what a sweetheart...what a cutie-pie*”, 488-489), aesthetic (“*they’re insanely beautiful*”, 729), and the activity (“*it’s so fun*”, 653). Participants are not merely responding emotionally but are busy triggering and maintaining the excitement. For example, after about 33 minutes, when the children are sitting observing the bugs in the box, there is silence and the excitement seems to have faded, Yoav tries to re-ignite it, suggesting ways to make the activity fun again: “*Yoooo I have an idea. It’s the most fun with insects. You take all sorts of sticks like this so they walk above and then they like climb upwards*” (662).

Modeling and guided participation

Dana’s modeling is particularly prominent in the sense-making game. When she first sees the insects, she immediately asks, “*What is it?*” (18), modeling interest in classification. She then tentatively suggests a classification: “*looks like a kind of bug*” (20), modeling caution of jumping into conclusions (as she could have said, “it’s a bug” even if she was not certain). The fact that Dana does not know what the insect is, allows her to authentically model this scientific practice. The children ignore Dana’s classification attempt, pointing to the bugs’ “copulating” and trying to draw Dana’s attention to it “*Yo here they’re copulating... Mom, they’re copulating.... Look, they’re copulating*” (21-26). Dana ignores their calls, refusing to play the daring game. Yet, weaving in the hunting game, she addresses Yoav noticing a caterpillar by modeling sense-making behavior, hypothesizing, “*maybe it’s her caterpillar? And they’re hatching here from the caterpillars now?*” (25). However unreasonable Dana’s hypothesis (caterpillars turn into pupae and not immediately into the mature form), raising it constitutes sense-making behavior and an authentic modeling of raising hypothesis.

Dana’s unreasonable hypothesis is challenged by Yoav over an hour later, when they are making Pizza and Yoav is reasoning that it does not make sense that the caterpillar is a phase in the bug’s life cycle: “*cause, mom, we saw a tiny green one... it doesn’t make sense that that size comes out of the caterpillar*” (1080). Dana accepts his argument and changes her mind: “*what you’re saying makes sense. So maybe they don’t have a*

caterpillar. The caterpillar isn't related" (1081). This both constitutes another example of the children's delayed uptake, and illustrates how the authenticity of Dana's sense-making allowed the children to move, with time, beyond their initial games into the sense-making game.

When Dana finally addresses "copulating", she weaves it into the sense-making game, asking Yoav to provide evidence that the bugs are indeed copulating: "*how do you know?*" (27). In doing so, she models another type of sense-making behavior, driving Yoav to ground his argument: "*Look, their ass is hooked.*" (28).

A little later, Dana suggests using a magnifying glass and searching online, but the children are engaged in a different game, so she follows her interest herself: she goes to the far end of the yard to look for bugs on other trees. While the children are not joining her, they can hear her talking to herself: "*I want to find out who they are and what's their deal and why they're only on the citrus tree...I'm going to see if they're on other trees as well*" (146, 162), and when she returns, she says: "*it's a citrus tree insect, 'cause on that side [of the yard] there's another citrus tree and it also has [bugs] on it. And only on it. Like, there're none on the other trees*" (171). Here, Dana models: (1) Interest in understanding "*what's their deal*"; (2) collecting evidence – although Dana already noticed that the bugs are only on the citrus tree, she goes to collect more evidence; and (c) induction – Dana verbalizes the logic that guides her conclusion: this is a citrus tree insect because I found another citrus tree, and the insects are on this tree and only this tree. These practices are later echoed in the children's behavior, (e.g., going to the far end of the yard to look for more bugs) and in their language (e.g., using the phrase "I'm interested if/why" the same way as Dana).

Then, Dana brings the laptop and searches online, again modeling the inclination to explore ("*I'm going to bring the computer. I'm dying to see what it is*", 350). When surprisingly quickly, she finds the bugs online, she expresses sincere excitement, transferring the treasure hunt game to the virtual space, repeatedly calling "*here it is! Yoo! I found it! Cool! First hit. Look. It's the first picture that came up, So cool.*" (386). Only then does she switch back to the sense-making game, reading the online information aloud. While the children appear disengaged in the sense-making game, they later use this information to participate therein, for example by planning to track changes in the bugs' color over time (lines 686-690), as well as by predicting, explaining, inferring, arguing, providing evidence, and more.

The examples I presented illustrate how Dana, as the more experienced member, modeled ways of participating in the games, mostly (yet not exclusively) the sense-making game. She modeled sense-making behavior and inclination while the children were still immersed in the treasure hunt, daring, and caring games; she did so by exhibiting intellectual curiosity, wondering and querying, collecting data, hypothesizing and making evidence-based inductions and expressing positive affect toward these processes. In addition, she supported the children's participation in the sense-making game by bridging between it and the other games, as well as by recruiting resources to solve the mystery of the insects, and making these available to the children. Although initially the children did not share the sense-making game with her, gradually they expanded the games they were playing (i.e., their definition of the situation) and their ways of participating therein. In doing so, they did not abandon their initial games (e.g., Amit still reflects on the thrilling experience of the bug flying over him even as they are already making pizza).

Discussion

The analysis elucidates legitimate multiple games as a prominent feature of this SSE episode. Participants simultaneously played five language games, with no one participant trying to regiment discourse and impose a certain game over others. This was inherently related to the absence of pre-set goals. Thus, the children could engage in the treasure hunt, admiring, caring, and daring games for a relatively long time before they shifted to the sense-making game. On one hand, when time is limited this might be viewed as a constraint. On the other hand, as this case shows, it may afford the weaving of games into the sense-making game, in an enjoyable, autonomous manner: one game supports the other as the children use new knowledge to make sense of their caring and daring games, in ways that might not be possible when only one dominant game is recognized and/or when time is restricted. The latter is often the case in formal learning environments (such as classrooms) as well as non-formal designed environments (such as museums) (Tal & Morag, 2007). Thus, while SSE inherently affords multiple legitimate games, for these to be productive, abundant time is required. In so saying, I do not mean to imply that regimented discourse is not possible in SSE, but rather suggest that one potential unique characteristic of SSE is legitimate multiple language games over relatively long periods of time.

The analysis also suggests that, at times, when children appear to be ignoring the sense-making game, they are actually attentive to it, picking up ideas, which they later use in delayed uptake (Mercer, 2008). Likewise, while the language game perspective has previously been employed primarily to exemplify how students may appear engaged in learning while they are actually playing other games, in this case the children appeared "playing" while they were actually learning. But what do they learn?

The children's participation evolved in a certain direction as the activity unfolded. This development is associated with observing Dana's practice and engaging therein with her, as well as in Dana's bridging between games. The shifts in the children's participation in the games constitute a process of guided participation. Thus, learning in SSE, may be viewed as the evolution of language games, in particular the sense-making game. Such evolution may entail employing additional definitions of the situation (i.e., joining the sense-making game) and/or expanding the ways of participating in familiar games (i.e., participating in the sense-making game in new ways). These may occur through guided participation, which takes a unique form in SSE, when there is no designated goal and participants do not necessarily share the same definition of the situation. Just as in structured activities (e.g., dressing and solving a puzzle), in SSE, guided participation entails observing others engage in other games, gradually engaging therein. However, while in structured activities, guided participation entails abandoning former games in favor of the "correct" game, in SSE, participants may continue engaging in other games.

Another important aspect of SSE's nature concerns the situatedness of learning. SSE's accidental nature affords authentic sense-making modeling, wherein the "expert" can genuinely exhibit wonder and interest, inquiry behavior, and other scientific practices and inclinations. Such authenticity is harder to obtain in designed engagement (e.g., in classrooms or science centers), when the adult usually knows the answers in advance and is only pretending to play the scientific game (Tal & Morag, 2007). Additionally, in SSE, children may pursue their own queries, rather than pretend to engage in others' queries or try to guess the desired (prescribed) answer to a pretend question, as is often the case in the classroom setting (Dillon, 1982).

Implications

The parental role in family SSE entails employing whatever expertise the parent has to model authentic scientific practice and inclination without imposing it on the children, thus allowing them to participate in multiple games. Achieving such a balance is a delicate endeavor, sometimes counter-intuitive requiring self-restraint; however, it may be essential to the maintenance of SSE. This raises questions regarding the cultural and social resources parents need in order to be able to productively support children's SSE.

Time is another important resource in SSE. Children of middle-class families are increasingly engaged in tightly scheduled, organized, extracurricular activities (such as science clubs) rather than in unstructured activities (such as outdoor free play) (Lareau, 2011). This study provides an example of the affordances of children engaging in unstructured activities, and suggests that they may be productive when an adult is around to model (not teach), but also when there is plenty of time.

Finally, this study expands Rogoff's (1993, 2008) notion of guided participation by: (a) describing what it may look like in activities that do not entail a designated goal; (b) illustrating how it advances learning not only in non-academic activities (such as dressing), but also in scientific ones.

Limitations and future research

In this study, through self-ethnography, I present and analyze an account of naturally-occurring family-based SSE. Self ethnography's affordances and constraints were shortly discussed in the methods section. Yet additional research exploring SSE in other families, from differing cultures and socio-economic backgrounds, with children of various ages, will shed light on differing patterns of family-based SSE and on issues such as parents' role. More work is also required to investigate SSE without adults' involvement, as well as other aspects of SSE, such as power relations, roles, and identity. In addition, longitudinal studies, which track changes in SSE and in associated structured engagement (e.g., in school), will teach us more about the long-term consequences of SSE. Also, exploring children's SSE as they mature, and portraying SSE trajectories across the life span is recommended.

If we agree that science learning is not restricted to designed environments, and that we are surrounded by opportunities for serendipitous science engagement, it then incumbent upon us to attempt to expand our understanding of such learning. Better understanding of SSE will enable us to better support it, or at least not to interrupt it. This study offers a first step toward such understanding.

References

- Alvesson, M. (2003). Methodology for close up studies—struggling with closeness and closure. *Higher education*, 46(2), 167-193.
- Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A. (Eds.). (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. National Academies Press.
- Bissex, G. L. (1985). *GNYS AT WRK: A child learns to write and read*. Harvard University.

- Bricker, L. A., & Bell, P. (2014). "What comes to mind when you think of science? The perfumery!": Documenting science-related cultural learning pathways across contexts and timescales. *Journal of Research in Science Teaching*, 51(3), 260-285.
- Blommaert, J. (2005). *Discourse: A critical introduction*. Cambridge University Press.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, 18(1), 32-42.
- Crowley, K., & Jacobs, M. (2002). Building islands of expertise in everyday family activity. *Learning conversations in museums*, 333-356.
- Dillon, J. T. (1982). The effect of questions in education and other enterprises. *Journal of Curriculum Studies*, 14(2), 127-152.
- Goodwin, M. H. (2007). Occasioned knowledge exploration in family interaction. *Discourse & Society*, 18(1), 93-110.
- Fleener, M. J., Carter, A., & Reeder, S. (2004). Language games in the mathematics classroom: teaching a way of life. *Journal of curriculum studies*, 36(4), 445-468.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74(1), 59-109.
- Lareau, A. (2011). *Unequal childhoods: Class, race, and family life*. Univ of California Press.
- Long, S. (2004). Passionless Text and Phonics First: Through a Child's Eyes. *Language Arts*, 81(5), 417.
- Ochs, E., Taylor, C., Rudolph, D., & Smith, R. (1992). Storytelling as a theory-building activity. *Discourse processes*, 15(1), 37-72.
- Rampton, B. (2007). Neo-Hymesian linguistic ethnography in the United Kingdom. *Journal of Sociolinguistics*, 11(5), 584-607.
- Rogoff, B. (2008). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. *Pedagogy and practice: Culture and identities*, 58-74.
- Rogoff, B., Mistry, J., Göncü, A., Mosier, C., Chavajay, P., & Heath, S. B. (1993). Guided participation in cultural activity by toddlers and caregivers. *Monographs of the Society for Research in Child development*, i-179.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic press.
- Sfard, A. (2001). There is more to discourse than meets the ears: Learning from mathematical communication things that we have not known before. *Educational Studies in Mathematics*, 46(1/3), 13-57.
- Tal, T., & Dierking, L. D. (2014). Learning science in everyday life. *Journal of Research in Science Teaching*, 51(3), 251-259.
- Tal, T., & Morag, O. (2007). School visits to natural history museums: Teaching or enriching?. *Journal of Research in Science Teaching*, 44(5), 747-769.
- Wertsch, J. V. (1979). From social interaction to higher psychological processes. A clarification and application of Vygotsky's theory. *Human development*, 22(1), 1-22.
- Wittgenstein, L. (1953). *Philosophical Investigations* Oxford. England: Blackwell.
- Yoon, B. (2012). Junsuk and Junhyuck Adolescent Immigrants' Educational Journey to Success and Identity Negotiation. *American Educational Research Journal*, 49(5), 971-1002.

Scientific Reasoning and Problem Solving in a Practical Domain: Are Two Heads Better Than One?

Andras Csanadi, Ludwig Maximilian University of Munich, andras.csanadi@psy.lmu.de
Ingo Kollar, University of Augsburg, ingo.kollar@phil.uni-augsburg.de
Frank Fischer, Ludwig Maximilian University of Munich, frank.fischer@psy.lmu.de

Abstract: To meet high-level problem solving standards, practitioners should optimally solve problems in an evidence-based manner: by using scientific knowledge while engaging in scientific reasoning processes. It can be argued that professionals do this better in groups than individually. Also, heterogeneous groups might have more potential to engage in higher levels of scientific reasoning than homogeneous groups or individuals. The present study investigated these questions in the context of teacher education. 76 teacher students solved a problem case from their future practice either individually or in dyads that were homo- or heterogeneous with respect to their members' problem solving scripts. The results show that although dyads have an advantage on generating hypotheses to explain the problem; individuals engage more in generating solutions. Moreover, especially heterogeneous dyads seem to differ from individuals' approaches. Future studies could investigate further how to use groups' potential in generating hypotheses to facilitate them in generating solutions.

Keywords: scientific reasoning, collaborative problem solving, group composition.

Introduction

Practitioners need to solve and reason on problems in their work on an everyday basis. It can be argued that in many domains, the kinds of reasoning processes that underlie competent solutions can be conceptualized as scientific reasoning. To make appropriate practical decisions, practitioners need to refer to relevant scientific knowledge from their domain and engage in reasoning and argumentation processes that are similar to the processes scientists in the particular domain would undertake (Fischer et al., 2014). However, utilizing scientific knowledge in a science-like way for solving complex problems is frequently a problem for practitioners (Gruber, Mandl & Renkl, 2000).

One possible way to support practitioners during problem solving is by allocating them into groups, i.e. to realize a collaborative problem solving setting (Rummel & Spada, 2005). A general notion is that groups tend to outperform single individuals on inquiry tasks (Lazonder, 2005). Collaborative partners, for example, can engage more in explaining and interpreting a phenomenon (Teasley, 1995), can facilitate better the regulation of the process of inquiry (Lazonder, 2005) and coordinate their ideas in a way that leads to a more abstract or deeper understanding of the problem (Schwartz, 1995). However, group settings do not guarantee better performance (Weinberger, Stegmann & Fischer, 2010). One way to explain these unequivocal findings is by considering group composition as a moderator of the effectiveness of collaborative learning. Group members' understanding on how they should reason during problem solving is guided by their knowledge or problem solving scripts (Fischer, Kollar, Stegmann & Wecker, 2013). The more diverging group members' views are on what processes and knowledge to build the own problem solving process on, the more creative they might be while solving the problem. However, introducing new perspectives might also increase the need for coordination in order to establish a joint problem space (Roschelle & Teasley, 1995).

The main aim of this study is to investigate whether scientific reasoning of future practitioners who solve an authentic problem from their practice differs if they work (1) individually or as a group; (2) as a homogeneous or as a heterogeneous group. These questions are investigated in the profession of teaching.

Solving practical problems as scientific reasoning

Based on the partially diverse conceptualizations of scientific reasoning (Fischer et al., 2014), we differentiate two aspects of scientific reasoning: the process and the content level (Zimmerman, 2000). At the *process* level, scientific reasoning can be viewed as an inquiry process (Klahr & Dunbar, 1988) that is characterized by an engagement in certain epistemic activities while solving a problem (Fischer et al., 2014). As the framework of Fischer et al. (2014) suggests, practitioners might engage in the same epistemic activities as scientists do for the sake of solving a practical problem. First, they need to identify the problem itself (Problem Identification); develop directions that target their further exploration on the problem (Questioning); consider potential explanations of

the problem (Hypothesis generation); take into account or generate further information necessary to understand or solve the problem (Evidence Generation); evaluate the information in the context of their hypotheses (Evidence Evaluation); generate solutions (Constructing artefacts); engage in discussions with others to re-evaluate their thoughts (Communicating and scrutinizing) and sum up their process to come to well-warranted conclusions on how to solve the problem (Drawing conclusions).

At the *content* level, we can look to what extent reasoners build on relevant theoretical concepts and empirical scientific findings of a domain during their problem-solving process. For example, teachers and teacher students should be able to apply teaching methods, theories of learning and relevant research findings from the Learning Sciences in order to solve a given problem (Voss, Kunter & Baumert, 2011), such as dealing with an underperforming student.

Collaborative problem solving and scientific reasoning

People reasoning together bring different perspectives on how to solve a given problem and have the opportunity to share their ideas (Hesse, Care, Buder, Sassenberg & Griffin, 2015). Groups have the potential of better and more innovative reasoning processes as reasoning partners may stimulate each others' cognitive processes (Paulus, 2000). Collaborative reasoning gives the opportunity to critically evaluate each other's as well as one's own ideas (Asterhan & Schwarz, 2009). Teasley (1995) demonstrated that dyads' typically produced more interpretive talk and hypothesis generation than individuals; in contrast, individuals engaged in more descriptive talk. Similarly, Okada and Simon (1997) found that reasoning with a partner resulted in higher engagement in explanatory processes (eg. hypothesis generation) compared to individual reasoning. One reason for such heightened engagement in explanatory behavior may be the need to communicate in a more explicit manner (Okada & Simon, 1997). Another is to think about the limitations of each learners' ideas, via challenging and defending one's own ideas (Asterhan & Schwarz, 2009). More recent studies (Métraiiller, Reijnen, Kneser & Opwis, 2008) also confirmed these findings showing that dyads hypothesize, challenge and justify ideas more often than individuals do. Such dialectic conversational turns occurring in a collaborative scenario might be beneficial for understanding (Schwartz, 1995), performing (Teasley, 1995) and learning (Asterhan & Schwarz, 2009).

On the other hand, groups typically underperform to their potential, a phenomenon identified as productivity/process loss (Weinberger, Stegmann & Fischer, 2010). Several factors have been identified accounting for process losses in collaborative vs. pooled performance settings (Paulus, 2000). Some authors (Weinberger et al., 2010) see process loss a group coordination problem. Also, several cognitive and social factors such as production blocking, cognitive load, downward social comparison, social loafing or evaluation apprehension might account for moderating collaborative performance (Paulus, 2000).

Collaborative problem solving in homogeneous vs heterogeneous groups

One aspect that might explain the ambiguous findings on whether groups outperform individuals or not, is group composition. For example, heterogeneous groups are expected to bring wider perspectives to the interaction than homogeneous groups (Paulus, 2000). A meta-analysis by Bowers, Pharmed and Salas (2000) concludes that although there are mixed results, heterogeneous teams perform better on complex problem solving tasks. More recent empirical research is in accordance with their findings. Studies (Canham, Wiley & Mayer, 2012; Wiley & Jolly, 2003) indicate that while homogeneous dyads perform better on routine-like tasks, heterogeneous dyads have an advantage when applying knowledge to solve novel problems. While heterogeneous group setting might pose higher coordination demands on its members, these demands can also lead to more elaborated reasoning (Teasley, 1995), for the benefit of problem solving. However, beyond the scattered findings (Canham et al., 2012), there is yet much light to shed on how reasoning processes are affected by group composition.

If group composition can indeed account for the ambiguous performance of groups, the question is what attributes of group members should we consider? This study is specifically concerned with individuals' knowledge about problem-solving in the domain in which they are situated. Such knowledge that directs collaborating partners' understanding and behavior in a situation can be described as "scripts" (Fischer, et al., 2013; Schank, 1999). While one learner may understand problem solving as the engagement in multiple epistemic activities in a given sequence, another learner's problem-solving script may include only some activities and even in a very different order. Reasoning partners with such diverse knowledge structures or problem solving scripts might stimulate each other, but may also face coordinational constraints and invest more effort to develop a common ground (Roschelle & Teasley, 1995). So far, however, there is a lack of empirical research investigating the effect of group composition with respect to the members' problem solving scripts.

Research questions

Utilizing scientific knowledge in a science-like way for solving complex problems is frequently a problem for future practitioners (Gruber, Mandl & Renkl, 2000). Collaborative problem solving is a highlighted cross-domain competence (e.g., Hesse et al., 2015; Rummel & Spada, 2005). While studies (e.g., Okada and Simon, 1997) indicate the advantage of collaboration on the engagement in scientific reasoning processes, groups frequently face process losses (Weinberger, Stegmann & Fischer, 2010). Moreover, research illuminating the role of homo- vs. heterogeneous collaboration on scientific reasoning processes is scarce.

These considerations led us to ask two research questions to be addressed in this paper:

- RQ1: To what extent do individuals and dyads of future practitioners differ in engaging in scientific reasoning processes and utilizing scientific knowledge during problem solving?
- RQ2: Does dyadic composition (homo- vs. heterogeneity) with respect to problem solving scripts affect future practitioners' scientific reasoning processes and scientific knowledge utilization during problem solving?

Methods

Participants and design

76 teacher education students (59 female, 17 male, $M_{age} = 21.22$, $SD = 3.98$) from a German university participated in the study. Each participant was randomly assigned to either an individual (16 students, $M_{age} = 22.31$; $SD = 6.73$) or a dyadic (60 students, i.e. 30 dyads, $M_{age} = 20.93$; $SD = 2.85$) condition. Within the dyadic condition, 22 students (11 dyads) reasoned in homogeneous groups ($M_{age} = 20.32$; $SD = 1.84$), i.e. collaborated with a learning partner with a similar problem solving script, while 38 students (19 dyads) reasoned in heterogeneous groups ($M_{age} = 21.29$; $SD = 3.28$), i.e. collaborated with a partner with a dissimilar problem solving script (for the operationalization of homo-/heterogeneity, see below).

Procedure

The study consisted of four steps. In the first three steps students participated individually; in the fourth step, they participated according to their condition: either individually or in dyads. Firstly, after the students arrived individually or in pairs to the study, they filled out an informed consent and a survey on demographic variables. Secondly, they were given a card sorting task to measure their problem solving scripts (see below). Thirdly, they read five printed out presentation slides (took about 5 minutes) that carried scientific content information originating from their introductory psychology class and encompassed short descriptions of theories and concepts (e.g., on strategic use of short-term memory and a classification of learning strategies). Fourthly, participants were presented an authentic problem from their (future) professional practice, which was: "You are a teacher in a school. One of your students receives low grades in comparison to others. The student looks motivated and it seems she understands the content. You know from the parents that she spends enough time on her homework and studies. You as a teacher, please find possible reasons and maybe solutions to the problem". Participants were asked to find reasons and possible solutions (individually or as pairs depending on the condition) for this problem in about 10 minutes. Data from this problem-solving process were used to measure reasoners' engagement in the epistemic activities of scientific reasoning and their use of scientific content. The whole data collection took about one hour.

Dependent variables

Dependent variables were collected during the problem-solving phase that had students (dyads or individuals) solve an authentic problem case from their future professional practice (see above). To make scientific reasoning processes visible, we asked students in the individual condition to think aloud and students in the dyadic conditions to verbally discuss how they would solve the problem. All verbal data were transcribed and segmented into propositional units (Chi, 1997). The proportion of agreement between two segmenters showed reliability with a lower bound of 79.73% (Coder 1) and an upper bound of 85.09% (Coder 2; Strijbos, Martens, Prins & Jochems, 2005). We developed coding schemes to assess students' engagement in (a) epistemic activities of scientific reasoning and (b) scientific content use.

Epistemic activities

All segments were coded with the aid of a coding scheme developed to capture the eight epistemic activities of scientific reasoning proposed by Fischer et al. (2014). Since it proved impossible to reliably separate "evidence generation" from "evidence evaluation", we merged these two categories into one: evidence evaluation. After

coding each segment, the numbers of segments that fell in the same category were summed for each epistemic activity (problem identification, questioning, hypothesis generation, constructing artefacts, evidence evaluation, drawing conclusions, communicating and scrutinizing) and for every transcript (individual or dyadic). The resulting sum scores were included in the statistical analyses. Two independent coders coded 10% of the data for the identification of epistemic activities. Inter-rater reliability was sufficient (Cohen's $\kappa = .68$).

Scientific content use

A second coding scheme was developed to capture for each segment whether or not participants used scientific content. Segments were coded as "scientific content" if speakers referred to scientific theories, concepts or methods. Specifically, we used the following five categories: "Learning strategy" was applied when participants mentioned learning strategy or memory coding-retrieval related topics. "Anxiety" was applied when participants referred to test anxiety or emotional pressure. "Motivation" was used when participants talked about motivation. As the last scientific content code we used the category "Other" such as "Self-fulfilling prophecy"; "Mobbing"; "Mind-map". If the above codes did not apply, the segment was coded as non-scientific content related. For each transcript (individual or dyadic) we merged all content categories to calculate engagement in scientific content use. 5% of the segments were coded by two independent coders with an agreement of Cohen's $\kappa = .82$.

Independent variables

Learning setting: Collaborative vs. individual reasoning

Learning setting was varied by assigning students to the problem-solving phase either as individuals or dyads.

Dyadic composition: Homo- vs. heterogeneity of problem solving scripts within dyads

We defined homo- and heterogeneous dyadic composition by comparing dyadic members' problem solving scripts that were measured during the card sorting task that preceded the problem-solving phase. During the card sorting phase, students were first presented the practice-related problem case described above. Then they were asked to use a set of prefabricated activity cards available on a MS PowerPoint slide to indicate what (epistemic) activities they would perform while solving the presented problem. The eight epistemic activities from Fischer et al. (2014) and five additionally selected activities (e.g., "Giving feedback", "Improvising") were written on the activity cards that were presented to the participants. Besides, five blank cards were provided to give participants the opportunity to note down further activities if they wanted to. Resulting activity sequences were then coded in the following way: We summed those epistemic activity cards that represented activities from the Fischer et al. (2014) model that were selected by both dyadic members. This number represented their shared knowledge component index (SKCI) on scientific inquiry. Then, we calculated a disagreement on sequentiality index (DSI) between dyadic members by calculating how many activities would need to be switched in position so as both members' selection shows the same sequence. Then, we calculated a pooled knowledge component index (PKCI) on scientific inquiry by summing the number of epistemic activities of the Fischer et al. (2014) model that at least one dyadic member had selected. Finally, a homogeneity index was calculated as $(SKCI - DSI)/PKCI$ to account for agreements and, at the same time, controlling for disagreements between dyadic members. A median split on the resulting values divided the sample in homo- and heterogeneous groups: dyads above a homogeneity index above .50 ($N = 11$) were considered as homogeneous groups, dyads with a value of .50 ($N = 9$) or below ($N = 10$) were considered heterogeneous dyads.

Results

Quantitative analysis

Overall, dyads ($M = 138.50$, $SD = 40.16$) talked more than individuals ($M = 88.19$, $SD = 27.74$) did, $F(1, 44) = 19.93$, $p < .001$, partial $\eta^2 = .31$. Thus, total talk was used as a covariate in all subsequent analyses. To investigate RQ1, we calculated a MANCOVA that included learning setting (individual vs. dyadic) as independent variable and all the eight epistemic activities as dependent variables. This MANCOVA revealed a main effect of learning setting, $V = .33$, $F(7, 37) = 2.59$, $p < .05$, partial $\eta^2 = .33$. Follow-up ANCOVAs revealed significant effects of learning setting on the engagement in hypothesis generation, $F(1, 43) = 5.68$, $p < .05$, partial $\eta^2 = .12$; as well as a close-to-significant effect on the engagement in solution generation $F(1, 43) = 3.35$, $p = .07$, partial $\eta^2 = .07$: Dyads engaged more in hypotheses generation (adj. $M = 29.18$, $SE = 2.24$) than individuals (adj. $M = 19.03$, $SE = 3.25$). Yet, individuals engaged longer in solution generation (adj. $M = 48.41$, $SE = 5.5$) than dyads (adj. $M = 35.21$, $SE = 3.8$). Regarding scientific content use, an ANCOVA with learning setting as independent variable and

the engagement in scientific content use as dependent variable revealed no significant effect of learning setting on the engagement in scientific content use, $F(1, 43) = .04, p = .84$.

Regarding RQ 2, we preliminarily found a significant main effect of dyadic composition on total talk, $F(2, 43) = 10.35, p < .001$, partial $\eta^2 = .33$. Subsequent post hoc tests revealed that both homogeneous ($M = 146.55, SD = 38.47$) and heterogeneous ($M = 133.84, SD = 41.39$) dyads talked more than individuals ($M = 88.19, SD = 27.74$) did, $p < .001$. Yet, there was no difference between homogeneous and heterogeneous dyads in total talk, $p = .36$. In light of these results, amount of total talk was used as a covariate for all analyses regarding RQ 2. To investigate RQ2, a MANCOVA with dyadic composition as independent variable, all epistemic activities as dependent variables and total talk as covariate revealed a significant multivariate main effect of dyadic composition on the engagement in epistemic activities, $V = .56, F(14, 74) = 2.04, p < .05$, partial $\eta^2 = .28$. Follow-up ANCOVAs showed significant effects of dyadic composition on the engagement in hypothesis generation, $F(2, 42) = 3.85, p < .05$, partial $\eta^2 = .16$; and in solution generation, $F(2, 42) = 4.87, p < .05$, partial $\eta^2 = .19$. LSD pairwise comparisons showed that heterogeneous dyads (adj. $M = 31.21, SE = 2.67$) engaged more in hypothesis generation than individuals (adj. $M = 19.33, SE = 3.22$), $p < .01$. Moreover, heterogeneous dyads (adj. $M = 29.39, SE = 4.31$) engaged significantly less in solution generation than homogeneous dyads (adj. $M = 46.52, SE = 5.85$) and individuals (adj. $M = 47.56, SE = 5.22$), $p < .05$. No other differences were significant, $p \geq .11$. Finally, an ANCOVA including total talk as covariate showed no effect of dyadic composition on scientific content use, $F(2, 42) = .06, p = .94$.

Qualitative analysis

The quantitative analysis revealed a different focus between dyads and individuals: dyads engaged longer in hypothesis generation and less in solution generation than individuals did. The comparison of heterogeneous dyads, homogeneous dyads and individuals led to the observation that the main differences were present between heterogeneous groups and individuals. Therefore, in the following, we are presenting examples from a (heterogeneous) dyadic discourse and an individual think aloud in order to briefly introduce the special characteristic of dyadic verbalizations that might have led to a different focus from the individual think alouds (names occurring in the excerpts are fictitious). The dyadic discussion stems from the conversation of a female (age = 20, elementary school teacher-to-be) and a male (age = 21, high school teacher-to-be) participant. The individual think aloud comes from a female (age = 18, high school teacher-to-be).

Excerpt 1: Heterogeneous dyadic reasoning

Sarah has initiated the conversation after about 10 seconds of the recording. After 30 seconds she is starting expressing her explanations for the problem case:

- Sarah: So I could imagine, that maybe it's exam fever that she has... when she is writing it on the paper, and maybe she cannot ask when she does not understand something or...
- Tim: As for me, I think one difficulty, [as I am] just reading, could even be a wrong learning strategy. She seems to study at home. However, for a teacher, it is not always totally visible how the student learns when it is a big class. Yes, I would also start with first, to think it over, where exactly the difficulties can occur. So what could be responsible for it after all.

Sarah has come up with the idea of exam fever and exemplifies it with a hypothetical episode how she thinks exam fever might influence the students' performance. Tim refers to the content of the slides that describe learning strategies and introduces a second plausible hypothesis. He evaluates the case information they have about the student by arguing with possible limitations of the observations. Then, Tim further opens the discussion on hypotheses by moving one step back in exploration: "what could be responsible for it after all". By these moves, Tim has changed the reference of exploration from Sarah's initial idea on exam fever. Indeed, Sarah continues the conversation from this altered point:

- Sarah: Yes, maybe the case is that she learns only to short term memory at home, and she tries to learn a large amount of the materials again short before the exam, and with that a lot of time is passing by... well, is lost. That's also apparently why the parents say that she learns fluently, and that is however,

not so effective. When she is sitting at the exam, she is confused or is forgetting it again.

Tim: Ok, how will we try to find it out?

Sarah elaborates on the idea that Tim has brought up beforehand and she also takes into account case evidence to justify her point. This accommodation regarding the topic and reasoning may be a coordination attempt to reduce ambiguity and to develop shared understanding. Tim is turning back to the original question, how they could make sure about the reason for the problem, causing a temporary confusion for Sarah who cannot answer his question. Shortly after the previous episode she comes up with an idea on a solution for the problem:

Sarah: Maybe you can also see that you describe another learning strategy to her, and look if she learns differently for the following exam, and maybe it turns out indeed better.

Tim: Okay, so to reflect then, if it has brought something. Yeah, ok.

[~40 seconds talk on the ideas how to solve the problem]

Tim: I think, what would be very important too, to go in the direction, if she might have totally different reasons, personal or so, yeah.

The initial problem solving phase here breaks by jumping back to the unclosed hypothesis generation phase. Tim's continuous turning back to the uncertainty of the explanations of the problem is a critical characteristic of this conversation as it does not let ideas for solutions evolve. After this episode, they continue discussing possible explanations for the problem and planning ways to generate more evidence to be able to find out the "real" explanation. Only in the very end, they reach the point of solution generation and do not elaborate on that longer than 45 seconds until the end of the discussion:

Sarah: Or as a teacher you can also try to give feedback to the student, that she does not need to be anxious from the writing, because her oral performance is very, very good, and [this way] you may take a bit the anxiety from the writing away, and say: try to imagine yourself, that I set the questions to you orally, and you write then down what you would say otherwise.

Tim: Yeah, you can also, if it is... it can depend also on the time pressure. [It] would also be a possibility. That you then simply take it away by letting all the students write longer. I cannot let [only] one student write longer.

The overall conversation took about seven and a half minutes and it seems that the dyad could not establish a common ground for further collaborative engagement in solution generation. Challenging questions seemed to lead to enhanced productivity from one side (considering engagement in hypothesis generation), but it seemed to hinder constructive collaboration (building up on each other's initial ideas) for solution generation.

Excerpt 2: Individual reasoning

Lisa starts 20 seconds after the recording and immediately considers explanations for the problem case:

Lisa: Well, I think that the problem for sure does not lie in learning, rather maybe in the pressure that sometimes comes from the parents, or from friends, or from external influences. So it can be that the student is simply anxious, she has exam fever, and at all of the exam situations she feels herself overloaded, and maybe from this pressure from the parents or from friends who have better grades than she, [she] is pressed or feels pressed. Of course, it can also be that her learning strategy is not really good, rather she learns like, she memorizes, and with that she has not understand it well, rather – like it is nowadays said – bulimic learnt. She learns everything and forgets everything.

During the first minute of her talk Lisa generates two potential hypotheses about the student's problem and gives brief explanations about these ideas. She seems to consider both ideas as equally possible ("Of course, it can also be...") without weighting their plausibility or, in contrast with Sarah and Tim, at least considering further information collection so as to be able to reason in a more evidence-based manner. In fact, her ideas are merely descriptive, consisting of claims and clarifications without engaging in deeper reasoning. After one minute she continues with solution generation:

Lisa: Therefore, it could make sense [...] to talk with the student about certain learning strategies, what options are there, how can someone really keep the stuff in the head, and not to forget it again, so how it is anchored in the mind. Those were such strategies...uhm...yeah, where you like, where she herself, what is really important, has a motivation to learn. That means that she is interested in the topic.

During the second minute, Lisa tries to solve the problem by building on her second hypothesis (learning strategy problem). When thinking about strategies for deeper knowledge processing, she mentions that the student should be motivated, should realize the value of the topic. This does not sound like a consequent solution to the problem of "bulimic learning" though, especially because the engagement or motivation of the student was not mentioned by Lisa earlier, when she was thinking of the potential explanations for the problem. It is a question though, whether engagement in deeper reasoning during the hypothesis generation phase would have triggered the appearance of additional ideas such as "motivation" and, consequently, lower the risk of incoherency between hypotheses and solutions.

In the remaining about five and a half minutes, Lisa's focus is on how to solve the problem. Although through some segments she is examining the potential of involving other parties (parents, school psychologist) to find out more information about the student and what her problem might be, this happens in the context of solutions. As a matter of fact, 75% of the remaining segments were coded as solution generation.

Conclusions

This study aimed to address two questions. Our first research question targeted if collaborative or individual setting is more beneficial for engagement in scientific reasoning processes and knowledge utilization. Our second research question was concerned with whether heterogeneous or homogeneous dyads show different engagement in scientific reasoning processes and knowledge utilization.

Quantitative analysis of the scientific reasoning processes showed a different pattern for collaborative vs individual reasoning. Dyads engaged in hypothesis generation longer (Okada & Simon, 1997), but at the same time in solution generation shorter than individuals. Separating dyads based on homogeneity vs heterogeneity of dyadic members' scripts proved to be meaningful. Heterogeneous dyads showed longer engagement in hypothesis generation but to a less extent in solution generation compared to individuals and homogeneous groups. However, to investigate where this difference originates from, we compared excerpts of a heterogeneous dyadic discourse with an individuals' think aloud.

The qualitative analysis demonstrated that the pattern of the amount of talk found in the quantitative analysis reflected also on the time of talk participants discussed hypotheses and solutions. Moreover, the way they used case evidence or planned the collection of further evidence was also in accordance with such an epistemic focus. We believe that dyads' focus on hypothesis generation can at least partially be explained by a coordination process that is necessary, especially for heterogeneous groups, to establish a common ground (Roschelle & Teasley, 1995; Schwartz, 1995). Coordination demands emerged in the initial stage of the conversation when participants also dealt with the explanations of the problem. Developing hypotheses may particularly afford the opportunity for developing shared meaning in dyads. Earlier studies also concluded that constructive engagement in hypothesis generation or experimentation might originate in the aim of knowledge co-construction (Okada & Simon, 1997).

Besides the advantage of dyads in more elaborated engagement in hypothesis generation, their solutions seemed to be shorter. Yet, they showed more coherence with their initial ideas than the individual excerpts. It is a matter of further analysis to investigate whether dyadic solutions - although they were shorter - showed more epistemic stability in terms that they emerged from well-evaluated and established hypotheses.

In the context of professional development, evidence-based collaborative problem solving is an important skill to acquire (Rummel & Spada, 2005). Engaging teacher students in complex problem solving processes in a collaborative setting might be beneficial to develop such skills. Furthermore, problem solving can benefit from

the diversity of collaborating partners if collaboration is appropriately structured (Rummel & Spada, 2005). For such efforts, it seems to be important to establish a grounding phase where participants can explore the problem, develop explanations and evaluate their own ideas (Asterhan & Schwarz, 2009). However, collaborators might need further scaffolding on their engagement in a solution generation phase, to build on the productive advantage of their initial exploration phase. Also, for teachers, who solve problems individually, it is important to reason critically enough while thinking of potential explanations of a problem and do not necessarily jump into conclusions by jeopardizing coherency and validity of their ideas.

References

- Asterhan, C. S., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science*, 33(3), 374-400.
- Bowers, C. A., Pharmed, J. A., & Salas, E. (2000). When member homogeneity is needed in work teams. A meta-analysis. *Small Group Research*, 31(3), 305-327.
- Canham, M. S., Wiley, J., & Mayer, R. E. (2012). When diversity in training improves dyadic problem solving. *Applied Cognitive Psychology*, 26(3), 421-430.
- Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. *Educational Psychologist*, 48(1), 56-66.
- Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., Neuhaus, B., Dorner, B., Pankofer, S., Fischer, M., Strijbos, J. W., Heene, M. & Eberle, J. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, 2(3), 28-45.
- Gruber, H., Mandl, H., & Renkl, A. (2000). Was lernen wir in Schule und Hochschule: Träges Wissen? In H. Mandl & J. Gerstenmaier (Eds.), *Die Kluft zwischen Wissen und Handeln* (pp. 139-156). Göttingen, Germany: Hogrefe.
- Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 37-56). Amsterdam, Netherlands: Springer
- Klahr, D., & Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. *Cognitive Science*, 12, 1-48.
- Lazonder, A. W. (2005). Do two heads search better than one? Effects of student collaboration on web search behaviour and search outcomes. *British Journal of Educational Technology*, 36(3), 465-475.
- Métraiiller, Y. A., Reijnen, E., Kneser, C., & Opwis, K. (2008). Scientific problem solving in a virtual laboratory: A comparison between individuals and pairs. *Swiss Journal of Psychology*, 67(2), 71-83.
- Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Science*, 21(2), 109-146.
- Paulus, P. (2000). Groups, teams, and creativity: The creative potential of idea-generating groups. *Applied Psychology*, 49(2), 237-262.
- Roschelle, J., & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer-Supported Collaborative Learning* (pp. 69-197). Berlin, Germany: Springer.
- Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *The Journal of the Learning Sciences*, 14(2), 201-241.
- Schank, R. C. (1999). *Dynamic memory revisited*. Cambridge, United Kingdom: Cambridge University Press.
- Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *The Journal of the Learning Sciences*, 4(3), 321-354.
- Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. (2006). Content analysis: What are they talking about? *Computers & Education*, 46(1), 29-48.
- Teasley, S. (1995). The role of talk in children's peer collaborations. *Developmental Psychology*, 31(2), 207-220.
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952-969.
- Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers In Human Behavior*, 26(4), 506-515.
- Wiley, J. & Jolly, C. (2003) When two heads are better than one expert. In R. Alterman D. Hirsh (Eds.), *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society* (pp 1242-1246). Boston, Massachusetts: Lawrence Erlbaum Associates Inc.
- Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, 20(1), 99-149.

Acknowledgments

This research was supported by the Elite Network of Bavaria under Grant K-GS-2012-209.

Combining Exploratory Learning With Structured Practice to Foster Conceptual and Procedural Fractions Knowledge

Nikol Rummel, Ruhr-Universität Bochum, nikol.rummel@rub.de

Manolis Mavrikis, UCL Knowledge Lab, University College London, M.Mavrikis@ioe.ac.uk

Michael Wiedmann, Ruhr-Universität Bochum, michael.wiedmann@rub.de

Katharina Loibl, University of Education, Freiburg, Ruhr-Universität Bochum, katharina.loibl@rub.de

Claudia Mazziotti, Ruhr-Universität Bochum, claudia.mazziotti@rub.de

Wayne Holmes, UCL Knowledge Lab, University College London, W.Holmes@ioe.ac.uk

Alice Hansen, UCL Knowledge Lab, University College London, A.Hansen@ioe.ac.uk

Abstract: Robust domain knowledge consists of conceptual and procedural knowledge. The two types of knowledge develop together, but are fostered by different learning tasks. Exploratory tasks enable students to manipulate representations and discover the underlying concepts. Structured tasks let students practice problem-solving procedures step-by-step. Educational technology has mostly relied on providing only either task type, with a majority of learning environments focusing on structured tasks. We investigated in two quasi-experimental studies with 8-10 years old students from UK ($N = 121$) and 10-12 years old students from Germany ($N = 151$) whether a combination of both task types fosters robust knowledge more than structured tasks alone. Results confirmed this hypothesis and indicate that students learning with a combination of tasks gained more conceptual knowledge and equal procedural knowledge compared to students learning with structured tasks only. The results illustrate the efficacy of combining both task types for fostering robust fractions knowledge.

Keywords: conceptual knowledge, procedural knowledge, exploratory learning environments, intelligent tutoring systems, mathematics education

Introduction

Robust domain knowledge consists of two types of knowledge, namely conceptual and procedural knowledge (Anderson, 1987; Rittle-Johnson, Siegler, & Alibali, 2001). Previous research has mostly focused on fostering either type of knowledge. This work explores how a combination of different types of instructional support provided by educational technology can foster both conceptual and procedural fractions knowledge.

Conceptual knowledge can be defined as implicit or explicit understanding about underlying principles and structures of a domain (Rittle-Johnson & Alibali, 1999). The focus of this type of knowledge lies on understanding why, for example, different mathematical principles refer to each other and on making sense of these connections. Procedural knowledge can be defined as knowledge about and application of procedures (Rittle-Johnson & Alibali, 1999). Procedures are an action sequence of, for instance, mathematical problem-solving steps (Rittle-Johnson & Alibali, 1999). The main aspect of procedural knowledge is in knowing how to apply a rule in order to solve a problem. According to Anderson's ACT-R model (e.g., Anderson, 1982), procedural knowledge becomes implicit with increasing practice.

Both types of knowledge develop over the same period of time (Canobi, Reeve, & Pattison, 2003; LeFevre et al., 2006) and evolve in a relationship of mutual dependence (Rittle-Johnson & Koedinger, 2009; Rittle-Johnson et al., 2001). Conceptual and procedural knowledge develop iteratively "with increases in one type of knowledge leading to gains in the other type of knowledge, which trigger new increases in the first" (Rittle-Johnson et al., 2001, p. 347).

While the development of these types of knowledge thus coincides (Rittle-Johnson et al., 2001), learning activities are thought to differ in which type of knowledge they primarily foster (Koedinger, Corbett, & Perfetti, 2012). Exploratory learning activities provide space for students to discover the underlying (mathematical) principles by abstracting concrete information and constructing schemata, thus primarily fostering conceptual knowledge. Structured practice activities introduce problem-solving procedures step-by-step and offer repeated (structured) practice opportunities for acquiring and deepening of these procedures (Anderson, Boyle, Corbett, & Lewis, 1990), thus primarily fostering procedural knowledge.

Educational technology has mostly focused on supporting either one or the other type of learning activity. Exploratory learning environments (ELEs), often referred to as micro-worlds, allow students, for example, to manipulate representations (e.g., Mavrikis, Gutiérrez-Santos, Geraniou, & Noss, 2013) and crucially, to explore the mathematical relationships between and within the representations and their underlying concepts (Hoyles,

1993; Thompson, 1987). ELEs can support students in these activities by encouraging reflection and self-explanation (e.g., Mavrikis & Gutiérrez-Santos, 2009). Intelligent Tutoring Systems (ITS) guide students through solving problems step-by-step, and offer immediate feedback so that students can automatize the problem-solving procedure bit by bit (Koedinger & Corbett, 2006; VanLehn, 2006). This feedback is typically directed more at problem-solving rather than at understanding the underlying concepts.

Given that conceptual and procedural knowledge develop in mutual dependence (e.g., Rittle-Johnson et al., 2001), it is somewhat surprising that prior work in the learning sciences and educational technology has focused on fostering *either* procedural knowledge with structured tasks (within ITSs) *or* conceptual knowledge with exploratory tasks (within ELEs). One exception is Holmes (2013) who investigated a games-based environment that provided both opportunities for children to discover numeracy concepts for themselves, using authentic problems that could only be solved using mathematics, and opportunities for them to practice related procedures. While Holmes did not explicitly test this hypothesis, combining exploratory learning and structured practice tasks should be more effective for learning because it promotes the iterative development of conceptual and procedural knowledge: conceptual understanding that students can directly apply to problem-solving should in turn deepen the conceptual understanding.

Some indirect evidence for this *combination effect* comes, for example, from research on productive failure in physics education (Kapur, 2008) and iterative lesson sequencing in mathematics education (Rittle-Johnson & Koedinger, 2009). Kapur (2008) investigated whether attempting to solve ill-structured problems before solving well-structured problems can be more productive than solving problems in the reverse order. The problems shared similarities with exploratory learning activities and structured tasks as defined above: ill-structured problems required students to discover how to structure and solve them by abstracting concrete information, while the well-structured problems required the application of predictable rules and principles. Kapur found that students who had solved ill-structured tasks first outperformed their counterparts later on in solving both ill-structured and well-structured tasks. However, students worked on these problems collaboratively and without educational technology, or support. Furthermore, there was no control condition where students learned with only one type of task. Rittle-Johnson and Koedinger (2009) investigated iterative lesson sequencing (lessons that alternate in focusing concepts or procedures) with an ITS. They found that the iterative lesson sequence fostered procedural knowledge more than a concepts-before-procedures sequence and that there was no difference in conceptual knowledge. However, the lessons that focused on concepts were heavily structured and did not provide the affordances for discovery that ELEs offer. Taken these limitations in mind, these findings still suggest that a learning environment combining exploratory and structured tasks could foster conceptual and procedural knowledge, and more so than structured tasks alone which the majority of existing tutoring environments provide.

In summary, while there is theoretical ground and indirect empirical evidence for combining exploratory learning with structured practice to promote conceptual and procedural knowledge acquisition, this hypothesis has not yet been explicitly tested. We report on two studies which investigated this question in two countries (Germany and the UK), using a newly-developed learning platform for fractions learning: iTalk2Learn. It is the product of an interdisciplinary research project funded by the EC under the 7th FP (italk2learn.eu.). Fractions were chosen as the learning domain because this mathematical topic is particularly difficult and challenging for young students (Charalambous & Pitta-Pantazi, 2007), and because students' fractions ability is a predictor for future math performance (Siegler et al., 2012)

Methods

Experimental design

The two studies were part of a larger research design that also investigated the impact of using speech to adapt to learners' needs. This paper reports data from two experimental conditions:

- (Full Platform) The full iTalk2Learn platform incorporating exploratory learning and structured practice.
- (No ELE) The iTalk2Learn platform incorporating structured practice, but not exploratory learning.

Due to the readily observable differences in learning tasks between the conditions, it was not feasible to run multiple conditions in the same classroom. Therefore, the studies were run in a quasi-experimental design.

Participants

Participants in both countries were students who were just about to start or at the beginning of formal fractions instruction. Fractions are taught earlier in the curriculum in the UK than in Germany. Parental consent for their involvement in the study was obtained for all participating students.

Participants in the study in the UK were Year 4 and Year 5 primary school students aged between 8 and 10 years from three schools. The schools were from a rural, suburban, and inner-city area. Seven students did not complete the study. Students were roughly stratified, according to previous teacher assessments of the children's mathematical ability, in three groups per grade per school which were then randomly assigned to one of the conditions, resulting in the following distribution across conditions: $N_{\text{Full Platform}} = 61$ and $N_{\text{No ELE}} = 60$.

Participants in the study in Germany were fifth and sixth grade secondary school students aged between 10 and 12 years from four schools from suburban areas. Participating students could not be stratified due to timetable constraints of the participating schools, so students participated within their class, and classes within schools were randomly assigned to one of the conditions. Class sizes varied, and, due to a technical failure, data was lost for one class of 33 students assigned to the No ELE condition, resulting in the following distribution across conditions: $N_{\text{Full Platform}} = 100$, and $N_{\text{No ELE}} = 51$.

Dependent measures

Participants completed an online and an offline fractions test, a questionnaire on attitudes to learning, mathematics and fractions, a questionnaire on their experience using the platform, and questions on their experience of the task they had just completed. We recorded all participant interaction with the platform, including speech. For a subsample of participants, while they worked with the platform observers assessed their affect. This paper reports data from the online fractions test.

For the online fractions test, two isomorphic versions were designed. Students were randomly allocated one version at the first time of measurement and the other version at the second time of measurement. Two subscales with three items each were constructed to measure procedural knowledge (see questions 22, 24, and 25 in Figure 1) and conceptual knowledge (see questions 20, 21, and 23 in Figure 1). The students received one point for each correctly-answered item and consequently obtained two aggregated scores, one per subscale. Internal consistency for the procedural scores at pre-test was $\alpha_{\text{UK}} = .40$, $\alpha_{\text{Germany}} = .07$, and at post-test $\alpha_{\text{UK}} = .53$, $\alpha_{\text{Germany}} = .36$. Internal consistency for the conceptual scores at pre-test was $\alpha_{\text{UK}} = .40$, $\alpha_{\text{Germany}} = -.03$, and at post-test $\alpha_{\text{UK}} = .36$, $\alpha_{\text{Germany}} = -.06$. The low internal consistency is addressed in the discussion.

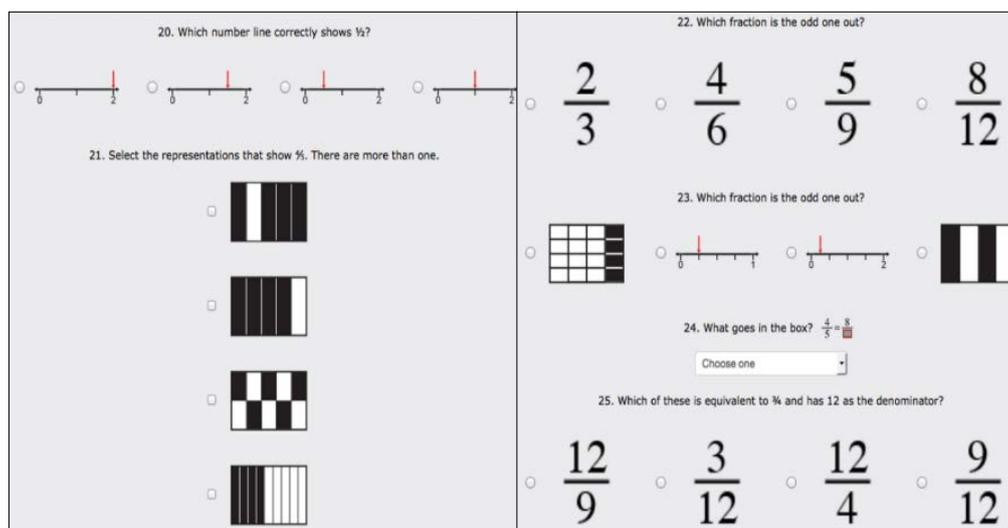


Figure 1. Online fractions test.

Procedure

Individual sessions were run with up to 15 students in the UK and up to 30 students in Germany. Each session lasted approximately 90 minutes including breaks. During the first 10 minutes, the students were introduced to the study and to the iTalk2Learn platform with the components being introduced depending on the experimental condition. To ensure that the introduction was as standardized as possible, it was scripted and was delivered by the same researchers in each session. The students were then asked to complete several instruments. The online fractions test was presented following the questionnaire on attitudes to learning, mathematics and fractions, both together in one browser window. Students were given 10 minutes total for these two instruments. Students then worked with the iTalk2Learn platform for approximately 40 minutes. During this main experimental period, the researchers adopted an intervention protocol that specified the allowable interactions and prompts. In the last 30 minutes of the session, the students were asked to complete the final instruments. The online fractions test was

presented following the user experience questionnaire, both together in one browser window. Students were given twenty minutes in total for these two instruments.

iTalk2Learn platform

The pedagogy of the iTalk2Learn platform is based on an intervention model for fostering robust knowledge described by Mazziotti et al. (2015). For the present studies, the model was instantiated for the topic of equivalent fractions. The platform combined an ELE delivering exploratory tasks, developed within the iTalk2Learn project, Fractions Lab (Hansen, Mavrikis, Holmes, & Geraniou, 2015; <http://fractionslab.lkl.ac.uk>), with one of two ITSs delivering structured tasks. In the UK, the ITS was a commercial system, Maths-Whizz (www.whizz.com); in Germany, it was an academic system, Fractions Tutor (e.g., Olsen, Belenky, Aleven, & Rummel, 2014; Rau, Aleven, & Rummel, 2013). The next section describes these learning environments and the tasks provided by them. Tasks were chosen by mathematics education experts based on an original coherent system for fractions learning that takes into account misconceptions and errors that are typical for learners at the beginning of formal fractions instruction (Hansen et al, 2014). Then, the adaptive support available to students is described. Finally, a section on the Student Needs Analysis (SNA) explains how tasks were sequenced within and switched between learning environments.

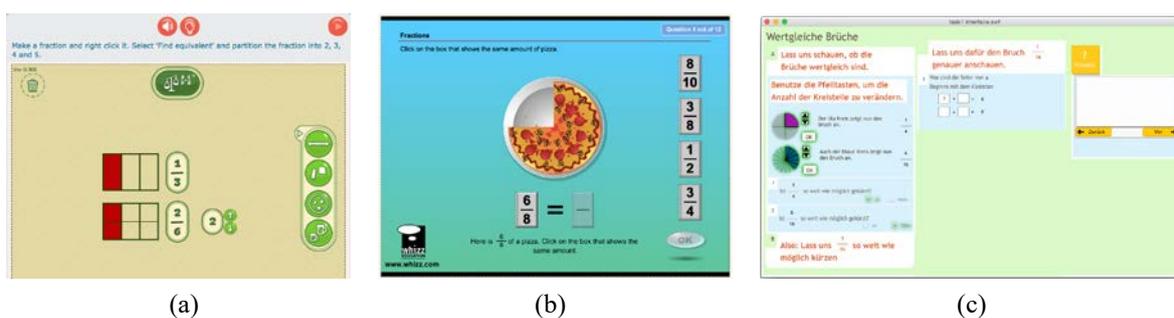


Figure 2. Exploratory Learning Environment (Fractions Lab; a), Structured Practice Environment used in the UK (Maths Whizz; b), and Structured Practice Environment used in Germany (Fractions Tutor; c).

Fractions Lab

Fractions Lab is an ELE that provides exploratory tasks that aim to help the student develop conceptual knowledge of fractions. In the Fractions Lab interface (see Figure 2a), a learning task is displayed at the top of the screen. Students can choose fraction representations (from the right-hand side menu) which they manipulate in order to solve the given task. For example, they can change the fraction's numerator or denominator, and find an equivalent fraction. An example task is shown in Figure 2a, which served both to introduce the student to available Fractions Lab functionality, and to introduce them to the idea and appearance of fraction equivalence with representations (Hansen et al., 2015).

Maths-Whizz

Maths-Whizz is a commercial system that provides structured practice content. This content is delivered in three stages: a teaching page which explains, procedurally, how to complete the following exercises successfully, interactive exercises with guided instruction and immediate feedback (see Figure 2b), and a short test. The exercises use a range of graphical representations such as circles, rectangles, number lines, liquid measures, symbols and sets of objects within contexts that the students may be familiar with.

Fractions Tutor

This web-based Cognitive Tutor for learning fractions (e.g., Olsen et al., 2014) enables students to solve fractions problems step-by-step, and receive immediate feedback or ask for on-demand hints. Content is presented on the same page and revealed step by step while students solve the problem (for an example, see Figure 2c). The exercises use a range of graphical representations such as circles, number lines, and symbols.

Adaptive support

While the students interacted with the ELE and with the ITS, they were given automatic task-independent support (TIS) and task-dependent support (TDS).

Within the ELE, the type of support provided (TIS or TDS) was based on a Bayesian Network which aims to change a negative affective state, for example frustration or boredom, into a positive affective state such

as enjoyment by adapting the feedback to the student’s affective state. Affective states were inferred from the student’s speech and from interaction data, that is, whether or not feedback had been followed. TIS support included affect boosts (e.g., “Well done. You’re working really hard!”), or talk-aloud prompts (e.g., “Please explain what are you doing.”). TDS support included instructive feedback (e.g., “Use the comparison box to compare your fractions.”, Holmes, Mavrikis, Hansen, & Grawemeyer, 2015), more open-ended feedback (e.g. “Good. What do you need to do now, to complete the fraction?”), reflective prompts (e.g., “What do you notice about the two fractions?”), affirmation prompts (e.g., “The way that you worked that out was excellent.”), or task sequence prompts (e.g. “Are you sure that you have answered the task fully? Please read the task again.”). The way how the feedback was provided to the student (high- or low- interruptive) was adapted according to their affective states (Grawemeyer, Holmes, Gutierrez-Santos, Hansen, Loibl & Mavrikis, 2015).

Within the ITS, TDS was provided based on students’ performance and consisted of highlighting mistakes and providing problem-solving instruction. TIS was provided as described above, based on the Bayesian network and adapted to students affect states deduced from their speech.

SNA: Sequencing within and switching between learning environments

In the Full Platform condition, students began their iTalk2Learn session in the ELE. While the student was engaged with the ELE, the Student Needs Analysis (SNA) component drew on various inputs (e.g., screen/mouse action, speech) to determine whether the student was under-, over-, or appropriately challenged by the task and thus to identify the next task appropriate for them. After each second task completed by the student, the SNA switched to the alternative type of task (i.e. when they had completed two exploratory tasks, they were switched to the ITS, and vice versa). If the student was switched to the ELE, the level of challenge that they had experienced on the previous task was taken into account when calculating the next task. The first task in the ITS was mapped to the fine-grain goal of the completed task in the ELE (e.g., partition a fraction to find its equivalent). The next task in the ITS stayed within the same fine-grain goal but increased the level of challenge based on a sequence determined by math education experts. Students continued in this fashion, alternating between exploratory learning and structured practice every second task until the 40 minutes were concluded. In the No ELE condition, students worked on the ITS only and received tasks based on the same sequence used in the Full Platform condition.

Findings

Table 1 presents scores on the online fractions knowledge test for the conceptual and procedural subscales. There was a medium correlation between these subscales on the post-test, $r(151) = .25$ in Germany and $r(121) = .26$ in UK, both $p < .01$.

Table 1: Scores on online fractions knowledge test

Subscale	Country	Condition	Pre-test		Post-test		Effect size	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>d</i>	95% CI
Conceptual	Germany	Full Platform	0.79	0.74	1.21	0.69	0.59	[0.30, 0.87]
		No ELE	0.73	0.63	0.53	0.64	-0.32	[-0.71, 0.07]
	UK	Full Platform	1.00	0.95	1.52	0.85	0.58	[0.22, 0.94]
		No ELE	0.88	0.92	0.70	0.77	-0.21	[-0.57, 0.15]
Procedural	Germany	Full Platform	0.95	0.80	1.42	0.94	0.54	[0.26, 0.82]
		No ELE	0.69	0.65	0.90	0.85	0.28	[-0.11, 0.67]
	UK	Full Platform	1.33	0.96	1.97	1.02	0.65	[0.28, 1.01]
		No ELE	1.47	1.02	1.87	1.07	0.38	[0.02, 0.74]

Note. Scores are summed across three items per subscale and can vary between 0 and 3.

Two-way 2 (condition: Full Platform or No ELE) x 2 (time of measurement: pre-test or post-test) multivariate ANOVAs with repeated measures on the time variable and conceptual and procedural subscale scores as the two dependent measures were conducted for each country separately. Using Pillai's trace, analyses showed significant effects of time of measurement for participants from both Germany, $V = .117$, $F(2,148) = 9.834$, $p < .001$, $\eta_p^2 = .117$, and UK, $V = .277$, $F(2,118) = 22.643$, $p < .001$, $\eta_p^2 = .277$. There were also significant effects of condition for participants from both Germany, $V = .132$, $F(2,148) = 11.274$, $p < .001$, $\eta_p^2 = .132$, and UK, $V = .133$, $F(2,118) = 9.025$, $p < .001$, $\eta_p^2 = .133$. Importantly, there were also significant interaction effects of time of measurement and condition for participants from both Germany, $V = .109$, $F(2,148) = 9.068$, $p < .001$, $\eta_p^2 = .109$, and UK, $V = .114$, $F(2,118) = 7.604$, $p < .001$, $\eta_p^2 = .114$. Results were similar in both countries: Students in both conditions showed learning gains, but these were stronger for the Full Platform condition. This interaction is now investigated further.

Follow-up univariate analyses showed significant effects of time of measurement on the procedural scores for participants from both Germany, $F(1,149) = 18.552$, $p < .001$, $\eta_p^2 = .111$ and UK, $F(1,119) = 16.337$, $p < .001$, $\eta_p^2 = .265$, but not on the conceptual scores for participants from neither Germany, $F(1,149) = 2.206$, $p > .05$, nor UK, $F(1,119) = 3.078$, $p > .05$. There were significant effects of condition on the procedural scores for participants from Germany, $F(1,149) = 10.618$, $p < .001$, $\eta_p^2 = .067$, but not from UK, $F(1,119) < 1$. For conceptual scores, there were significant effects of condition for participants from both Germany, $F(1,149) = 16.465$, $p < .001$, $\eta_p^2 = .100$, and UK, $F(1,119) = 13.999$, $p < .001$, $\eta_p^2 = .105$. Finally, there were no significant interaction effects of time and condition on the procedural scores for participants from neither Germany $F(1,149) = 2.552$, $p > .05$, nor UK, $F(1,119) = 2.279$, $p > .05$. But on the conceptual scores, there were significant interaction effects for participants from both Germany, $F(1,149) = 16.697$, $p < .001$, $\eta_p^2 = .101$, and UK, $F(1,119) = 13.245$, $p < .001$, $\eta_p^2 = .100$. Results were similar in both countries: Students in both conditions showed equal learning gains on procedural scores. The decrease of conceptual scores in the No ELE condition is not significant: the 95% confidence interval of the effect indicates that a small increase is similarly likely. Students in the Full Platform condition did show significant learning gains on conceptual scores.

Discussion

Robust knowledge consists of conceptual and procedural knowledge that need different types of instructional support. Yet, learning systems that have been developed for mathematics education are usually constrained either to exploratory tasks or to structured tasks and thus can promote learning only to a limited extent. We demonstrated in this paper our attempt to overcome this limitation by combining exploratory tasks from Fractions Lab, a newly-developed exploratory learning environment, and structured tasks from Maths-Whizz and Fractions Tutor, two proven intelligent tutoring systems.

Two studies provided clear evidence that the combination of exploratory tasks (to foster primarily conceptual knowledge) and structured tasks (to foster primarily procedural knowledge) in one learning environment promotes fractions knowledge more than state-of-the-art ITSs providing structured tasks only. More specifically, the combination of tasks led to stronger conceptual learning gains without hindering procedural learning. The latter is particularly remarkable given that learning time was split between exploratory and structured tasks in the Full Platform condition, and given that exploratory tasks primarily target conceptual knowledge. Interestingly, procedural learning gains were smaller in the No ELE than in the Full condition. Different from the contexts in which Maths-Whizz and Fractions Tutor are usually deployed, in the present studies students had very limited time to study very specific learning content. Moreover, participants had not worked with these learning environments before. Against this background, the clear learning gains observed in the Full Platform condition are even more impressive.

In spite of the overall results, there are some limitations worth discussing. The first limitation concerns the measurement of procedural versus conceptual knowledge. These constructs can hardly be measured independently (Schneider & Stern, 2010), but we emphasize that our measures are meant to primarily tap one versus the other type of knowledge. Moreover, in the German sample, the conceptual scores are not internally consistent which highlights the need to investigate their dimensional structure and further develop a valid measurement of procedural versus conceptual knowledge. That said, the clear result patterns overall does provide first evidence of an interesting effect. Following a multi-method approach, we have collected more data which, once analyzed, will shed further light on the validity of the results presented here. Another limitation is the short duration of the intervention. While conducting the studies in school classrooms increased external validity, it also placed constraints on the learning time available to us through the schools. An extended follow-up study may lead

to more robust knowledge gains, evidenced by larger internal consistencies of measures, and retention and transfer effects.

Despite these limitations, and the differences between the two studies conducted in Germany and in the UK, the results were remarkably similar between the two countries. This speaks to the generalizability of our findings and the external validity of the combination effect. Finding evidence for the combination effect underlines the need to foster both types of knowledge jointly, as Rittle-Johnson et al. (2001) highlighted with their iterative model of knowledge development.

The effect of combining task types prompts a series of follow-up questions. One of these questions asks for the component that makes the combination effect effective. For example, is the order of exploratory followed by structured tasks essential for realizing the combination effect? Or could the order be reversed? We based the order of exploratory and structured tasks implemented in our studies on prior research that showed conceptual learning should be fostered first (e.g., Kapur, 2008). This principle was not only realized within the first two tasks, but formed one of the rules of our intervention model employed throughout learning with iTalk2Learn. The iTalk2Learn system now provides an additional, proven research context in which the generalizability of the prior research findings can be tested.

References

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369–406. doi:10.1037/0033-295X.89.4.369
- Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological Review*, 94(2), 192–210. doi:10.1037/0033-295X.94.2.192
- Anderson, J. R., Boyle, C., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence*, 42(1), 7–49. doi:10.1016/0004-3702(90)90093-F
- Canobi, K. H., Reeve, R. A., & Pattison, P. E. (2003). Patterns of knowledge in children's addition. *Developmental Psychology*, 39(3), 521–534. doi:10.1037/0012-1649.39.3.521
- Charalambous, C. Y., & Pitta-Pantazi, D. (2007). Drawing on a theoretical model to study students' understandings of fractions. *Educational Studies in Mathematics*, 64(3), 293–316. doi:10.1007/s10649-006-9036-2
- Grawemeyer, B., Mavrikis, M., Holmes, W., & Gutiérrez-Santos, S. (2015). Adapting feedback types according to students' affective states. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Lecture notes in computer science. Artificial Intelligence in Education* (pp. 586–590). Springer International Publishing.
- Grawemeyer, B., Holmes, W., Gutiérrez-Santos, S., Hansen, A., Loibl, K., & Mavrikis, M. (2015). Light-bulb moment? Towards adaptive presentation of feedback based on students' affective state. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (pp. 400–404). Atlanta, Georgia, USA: ACM. doi:10.1145/2678025.2701377
- Hansen, A., Mavrikis, M., Holmes, W., Grawemeyer, B., Mazziotti, C., Mubeen, J., & Koshkarbayeva, A. (2014). *Report on learning tasks and cognitive models* (iTalk2Learn deliverable 1.2). Retrieved from <http://www.italk2learn.eu/deliverables-and-publications/deliverables/>
- Hansen, A., Mavrikis, M., Holmes, W. & Geraniou, E. (2015). *Designing interactive representations for learning fraction equivalence*. Paper presented at ICTMT. Faro, Portugal. 24-27 June.
- Holmes, W. (2013). *Level Up! A design - based investigation of a prototype digital game for children who are low-attaining in mathematics*. Doctoral thesis (D.Phil.) University of Oxford. Retrieved from <http://solo.bodleian.ox.ac.uk>.
- Holmes, W., Mavrikis, M., Hansen, A., & Grawemeyer, B. (2015). Purpose and level of feedback in an exploratory learning environment for fractions. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial Intelligence in Education* (Vol. 9112, pp. 620–623). Cham: Springer.
- Hoyles, C. (1993). Microworlds/Schoolworlds: The transformation of an innovation. In C. Keitel & K. Ruthven (Eds.), *Learning from Computers: Mathematics Education and Technology* (pp. 1–17). Berlin, Heidelberg: Springer.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. doi:10.1080/07370000802212669
- Koedinger, K. R., & Corbett, A. T. (2006). Cognitive Tutors: Technology bringing learning sciences to the classroom. In K. R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–77). New York, NY, US: Cambridge University Press.

- Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, 36(5), 757–798. doi:10.1111/j.1551-6709.2012.01245.x
- LeFevre, J.-A., Smith-Chant, B. L., Fast, L., Skwarchuk, S.-L., Sargla, E., Arnup, J. S., . . . Kamawar, D. (2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through Grade 2. *Journal of Experimental Child Psychology*, 93(4), 285–303. doi:10.1016/j.jecp.2005.11.002
- Mavrikis, M., & Gutiérrez-Santos, S. (2009). Informing the design of intelligent support for ELE by communication capacity tapering. In U. Cress, V. Dimitrova, & M. Specht (Eds.), *Lecture notes in computer science. Learning in the synergy of multiple disciplines* (pp. 556–571). Springer Berlin.
- Mavrikis, M., Gutiérrez-Santos, S., Geraniou, E., & Noss, R. (2013). Design requirements, student perception indicators and validation metrics for intelligent exploratory learning environments. *Personal and Ubiquitous Computing*, 17(8), 1605–1620. doi:10.1007/s00779-012-0524-3
- Mazziotti, C., Holmes, W., Wiedmann, M., Loibl, K., Rummel, N., Mavrikis, M., . . . Grawemeyer, B. (2015). Robust student knowledge: Adapting to individual student needs as they explore the concepts and practice the procedures of fractions. In M. Mavrikis, G. Biswas, S. Gutiérrez-Santos, T. Dragon, R. Luckin, D. Spikol, & J. Segedy (Eds.), *Proceedings of the Workshops at the 17th International Conference on Artificial Intelligence in Education AIED 2015* (Volume 2; pp. 32–40).
- Olsen, J. K., Belenky, D. M., Alevan, V., & Rummel, N. (2014). Using an intelligent tutoring system to support collaborative as well as individual learning. In S. Trausan-Matu, K. Boyer, M. Crosby, & K. Panourgia (Eds.), *Lecture notes in computer science. Intelligent tutoring systems* (pp. 134–143). Springer International Publishing.
- Rau, M. A., Alevan, V., & Rummel, N. (2013). Interleaved practice in multi-dimensional learning tasks: Which dimension should we interleave? *Learning and Instruction*, 23, 98–114. doi:10.1016/j.learninstruc.2012.07.003
- Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, 91(1), 175–189. doi:10.1037/0022-0663.91.1.175
- Rittle-Johnson, B., & Koedinger, K. (2009). Iterating between lessons on concepts and procedures can improve mathematics knowledge. *The British Journal of Educational Psychology*, 79(Pt 3), 483–500. doi:10.1348/000709908X398106
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346–362. doi:10.1037/0022-0663.93.2.346
- Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: a multimethod approach. *Developmental Psychology*, 46(1), 178–192. doi:10.1037/a0016701
- Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., . . . Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23(7), 691–697. doi:10.1177/0956797612440101
- Thompson, P. W. (1987). Mathematical microworlds and intelligent computer-assisted instruction. In G. P. Kearsley (Ed.), *Artificial intelligence and instruction: Applications and methods* (pp. 83–109). Addison-Wesley Longman Publishing Co., Inc.
- VanLehn, K. (2006). The behavior of tutoring Systems. *International Journal of Artificial Intelligence in Education*, 16(3), 227–265.

Acknowledgments

We thank participating students and teachers, and the iTalk2learn consortium. We would like to extend special thanks to Sergio Gutiérrez-Santos, Beate Grawemeyer, and José Luis Fernandez-Gomez for the development of the iTalk2Learn platform and their support. The research leading to these results has received funding from the European Commission's Seventh Framework Program under grant agreement n°318051.

The Effects of Self-Regulated Learning on Students' Performance Trajectory in the Flipped Math Classroom

Zhiru Sun, University of Southern Denmark, zhiru@sam.sdu.dk

Lin Lu, The Ohio State University, lu.962@osu.edu

Kui Xie, The Ohio State University, xie.359@osu.edu

Abstract: The flipped classroom is a unique instructional model with in-class activities being heavily dependent on pre-class learning engagement. Practitioners and researchers have well recognized the critical role of the pre-class learning on students' performance in the flipped class. However, little research has investigated how students learn in the pre-class setting and what factors can help students improve their pre-class learning. Based upon the self-regulated learning theory, this study used both survey research and learning analytics methods to uncover students' performance trajectories of a semester in two flipped undergraduate math courses and examined whether the three key self-regulated learning factors – self-efficacy, the use of metacognitive strategy, and self-regulatory behaviors – were significantly different across different performance trajectories. The results revealed six different performance trajectories among students and found significant differences across these trajectories in terms of the three self-regulatory factors.

Keywords: self-regulated learning, flipped classroom, learning analytics, survey-based research

Introduction

The flipped classroom model has attracted ever-growth attention since its inception in 2007. It is an emerging instructional model that commonly consists of two parts: pre-class Internet-based individual learning and in-class group-based collaborative learning (Bishop & Verleger, 2013). In a flipped class, students normally learn content materials through online instructional videos and text readings on their own pace and schedule prior to the class. Then they gather together in person and apply learned knowledge through group-based collaborative activities in a classroom (Bergmann & Sams, 2012; Herreid & Schiller, 2013). The in-class activities are oftentimes designed to be highly associated with the pre-class learning materials. Thus, the success of in-class learning is heavily relied on students' preparation in pre-class. In the practice of flipped classroom model, teachers have observed the critical influence of the pre-class preparation on the in-class performance (Schell, 2013). Research also empirically demonstrated the significantly positive relationship between the pre-class and in-class achievement (Sun, 2015). Given the important role of the pre-class learning in the flipped classroom, how do students learn in the pre-class and what factors can help students improve their pre-class learning have been continuous questions among flipped classroom teachers.

The pre-class online learning emphasizes a student-centered learning environment, in which students take control of their own learning process (Flipped Learning Network, 2014). For example, students are provided with various learning resources, and they navigate through the materials and learn the content at flexible time, pace, and space (Fulton, 2012). To ensure that they are prepared for the in-class group activities, students are oftentimes required to watch the associated lecture videos and complete the assigned homework prior to class. However, this student-centered way of learning has been a well-known challenge for both students and teachers. Students complain the workload ahead, and teachers lack awareness and confidence that students would spend enough time and efforts on the lectures and be prepared for the in-class activities (Acedo, 2013). To overcome this challenge, researchers suggest that a new or advanced student-centered skill set is demanded to support such the student-centered way of learning (Estes, Ingram, & Liu, 2014). That means, students are expected to be proactive, inquire, and obtain foundational knowledge in a self-directed manner to succeed in the pre-class learning (Talbert, 2014). They are also desired to be active in the setting and pursuit of learning goals, use certain learning strategies to solve problems, monitor their behaviors, and reflect on the performance. In essence, the student-centered environment promotes self-regulated learning as a critical and helpful skill set for the enhancement of students' pre-class learning performance (Connor, Newman, & Deyoe, 2013).

Self-regulated learning is an integrated learning process that regulates students' motivation, behaviors, and metacognitive activities to pursue the planned and adapted personal goals (Schunk, 2001). The self-regulated learning research has suggested three factors that are significant in the self-regulation process, which are *self-efficacy* (Bandura, 1997; Pintrich & Zusho, 2007), *the use of metacognitive strategy* (Duncan & McKeachie, 2005; Winne, 2001), and *the regulation of behaviors* (Hadwin et al., 2007; Winne & Hadwin, 2008). Self-efficacy

denotes as “people’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances” (Bandura, 1986, p.391). Research has demonstrated its robust predicting effect on academic performance (Crede & Philips, 2011) and also suggested its fundamental role in the self-regulated learning process (Winne & Hadwin, 1998, 2008). According to the self-regulated learning model proposed by Winne and Hadwin (1998), while giving a learning task, students normally first define the task such as an easy or a hard task based on their own cognitive factors, and such the definition of the task will guide their following self-regulatory behaviors, for example the regulation of learning time and efforts. Self-efficacy, as one of the cognitive factors, helps students to first generate the task definition and then leads them to regulate their behaviors to complete the task. For example, if a student holds higher-level self-efficacy towards a task, he or she would feel more confident to perform the task, devote more time and efforts in the task, and persist longer while meeting obstacles during the task.

Metacognitive strategy refers to the self-evaluation strategy that guides students’ planning, monitoring, and modifying their cognition and behavior (Pintrich & De Groot, 1990). The use of this strategy is believed to be the hallmark of self-regulating one’s learning process because by using this strategy, students compare their current profiles of task performance with the standards they set previously for satisfactory task performances, create a list of matches and mismatches between the current profiles and the standards, and regulate their motivation, cognition, and behavior in the following learning activities (Winne, 2001). Research has revealed that the use of metacognitive strategy plays an important role in students’ learning and, in general, is positively related with academic performance (Pintrich, 2004). In particular, the use of metacognitive strategy has been quite critical in the hypermedia or computer-supported learning environment (Azevedo, 2005). Because such the learning environment provides students with dynamic and nonlinear access to a wide range of information represented as texts, graphs, and videos, without deploying the metacognitive strategy, students are unable to monitor their performances and easily lost in the great amount of information (Azevedo & Cromely, 2004). The pre-class learning setting is the environment that supports students’ learning mainly by computers, which shares the same feature in the students’ self-directed learning with the hypermedia learning environment. Therefore, the use of metacognitive strategy is believed to be a significant part of the self-regulated learning in the pre-class setting.

The regulation of behaviors in general means the management one’s behaviors in ways to achieve set goals (Carlson & Moses, 2001). Pintrich and his colleagues (1993) attempted to summarize the essential self-regulatory learning behaviors into four categories in the Motivated Strategies for Learning Questionnaire, which are time/study environmental management, effort regulation, peer learning, and help seeking. They believed that to achieve certain learning goals, self-regulated learners tend to find a specific time and place that they can concentrate on the learning task, spend more time on doing the task and persist longer while meeting obstacles, identify peers in the class who are capable of answering questions, and feel unafraid to seek help from others if needed. Both research and theory suggest that self-regulatory behaviors students enacted during learning have a positive effect on their achievement (Winne & Hadwin, 2008).

The majority of research explored the effects of these three self-regulatory factors on students’ one-time achievement such as a post-test score or a final exam score (Azevedo & Cromely, 2004; Crede & Philips, 2011). Since the self-regulated learning is a process, it is valuable to explore the effects of such the self-regulatory process on a performance trajectory or trend, for example a performance trend in an entire semester. Because compared to the one-time achievement, the analysis of the performance trend is able to provide more detailed look on students’ learning process during a semester and identify potential linkage between changes in the performance and the self-regulatory factors.

In this paper, we first identify students’ performance trajectories based on four major exam scores obtained during a semester in two flipped undergraduate math courses, and then investigate whether there are significant mean differences across different performance trajectory groups in terms of the three key self-regulated learning factors – math self-efficacy, the use of metacognitive strategy, and self-regulatory behaviors. We use both survey research and learning analytics methods to uncover students’ self-regulated learning process in the pre-class learning setting. These methods allow us to examine both students’ psychological perceptions and their detailed learning behaviors during the pre-class learning. Two major research questions guide the design of this study, which are as follows:

1. What are students’ performance trajectories in a semester in the flipped undergraduate math courses?
2. Are there any mean differences across different performance trajectories in terms of math self-efficacy, the use of metacognitive strategy, and self-regulatory behaviors?

Methods

Participants

A total of 151 undergraduate students from Calculus I and II courses in a large Midwestern university in the United States participated in the study. There were 50.3% female and 49% male. The majority of participants were White students (68.9%), followed by Asian students (25.2%), and the rest from other ethnical backgrounds. Over 84% of students were freshmen and sophomore.

Study context and procedure

The study included 16 flipped sessions of Calculus I and II. Every Tuesday and Thursday, students attended 55 minutes recitation sessions either in person or through online web-conferencing tool. During the recitation session, students applied newly learned knowledge from the pre-class learning through problem-based and group-based work. To effectively learn in the recitation session, students were required to complete the associated pre-class lectures before attending the recitation session. An online lecture commonly consisted of lecture videos, information slides, embedded quizzes, formative feedback for the quizzes, and associated homework at the end of the online lecture. To obtain the insight of students' perceptions of their self-regulated learning, two online surveys were sent to consented students at the 3rd and 10th week of the semester.

Measures

Measures of this study include math self-efficacy, the use of metacognitive strategy, behavioral data, and exam scores. Math self-efficacy and the use of metacognitive strategy were measured on seven-point Likert scales at the first and second survey, respectively. The behavioral log data was retrieved at the end of the semester from the learning management system that students used for the flipped courses. The four major exam scores were obtained from the course instructors. Means, standard deviations, Cronbach's alpha of survey scales, and correlation coefficients between all variables are summarized in Table 1.

Math Self-efficacy (MSE): Motivated Strategies for Learning Questionnaire (MSLQ: Pintrich et al., 1991) is a widely used measure in the areas of motivation and self-regulated learning (Anthony & Artino, 2005). Extensive research has provided evidences to show that MSLQ is a reliable and valid instrument (Pintrich et al., 1993; Duncan & McKeachie, 2005). The 5-item self-efficacy subscale of MSLQ was adapted to measure students' math self-efficacy of learning in the flipped math courses. One sample item is "*I'm confident I can understand the basic concepts taught in math.*"

Use of Metacognitive Strategy: The 12-item metacognitive self-regulation learning strategy subscale of MSLQ was adapted to measure students' use of metacognitive strategy while learning in the pre-class online setting of the flipped courses. One sample item is "*When studying the pre-class lecture materials, I try to determine which concepts I don't understand well.*"

Behavioral Data: Three types of behavioral data were retrieved from the learning management system, which include the number of attempts for each online homework, the progress of watching each online lecture (i.e. complete the lecture or not complete), and the time (minutes) spent on each online lecture. These behavioral data were log data that presents detailed information of how students work on each online lecture and homework during the semester.

Exam Scores: Students were tested four times during the semester in the paper-and-pencil exams. The four exams largely evenly spread out in the semester, and each exam was designed to test students' understanding and application of the learned knowledge prior to the exam. There were 16 weeks in the entire semester. Exam 1 was conducted in week 4; exam 2 was in week 9; exam 3 was in week 14; and exam 4 was in week 16. The first three exams were the midterms, and they, in total, accounted for 50% of the final grade. The last exam was the final exam, and it accounted for one third (approximately 33%) of the final grade.

Table 1: Means, standard deviations, Cronbach's alpha, and correlation coefficients of measures

Variable	Mean	SD	Alpha	Zero-order Correlations				
				1	2	3	4	5
1. Math Self-efficacy	4.85	1.48	.93	-				
2. Metacognitive Strategy	3.55	2.29	.89	.175**	-			
3. Attempt	.99	.96	-	-.001	.069**	-		
4. Progress	.39	.49	-	.004	.033**	.566**	-	
5. Time (minus)	90.28	231.27	-	-.048**	.006	.310**	.217**	-

Note: ** $p < .01$

Analysis

Hierarchical cluster analysis

Hierarchical clustering was employed in this study as a learning analytics method to answer the first research question. Hierarchical clustering, compared to other clustering methods, better helps researchers to explore the unknown numbers of latent groups in the dataset based on the similarity of interested variables (Vellido, Castro, & Nebot, 2011). In this study, students took four major exams at the 3rd, 9th, 14th, and 16th week of the semester that assessed their understanding of the content knowledge within a certain period of time. The scores of the four exams were standardized before the analysis. Ward's minimum variance method (Ward, 1963) was used as the criterion to evaluate the similarity by the squared Euclidean distances between groups. Each cluster or group was expected to describe a different performance trajectory of students in the semester.

ANOVA analysis

ANOVA analysis was the primary method that used to answer the second research question. The dependent variables included the three self-regulatory factors – math self-efficacy, the use of metacognitive strategy, and self-regulatory behaviors. The independent variable was the performance trajectory groups generated by cluster analysis. In addition, Tukey's HSD test, a post hoc tests in ANOVA, was conducted to further explore the mean differences of the three self-regulatory factors across different performance trajectory groups at an adjusted alpha level .017 (i.e. .05/3).

Findings

Cluster analysis findings

Agglomeration schedule was examined to determine the number of clusters that could be kept for further analysis. Figure 1 demonstrated the grouping procedure with the distances between clusters in each combining stage. It was noticed that the increase of distance between clusters started to go up from the stage of cluster 7 to the stage of cluster 4. To see which stage caused the highest increase of distance after combining clusters, Figure 2 was plotted. It was shown that the change of distance between clusters increased dramatically after agglomerating cases from six clusters to five clusters. This indicated that six clusters could be an appropriate solution.

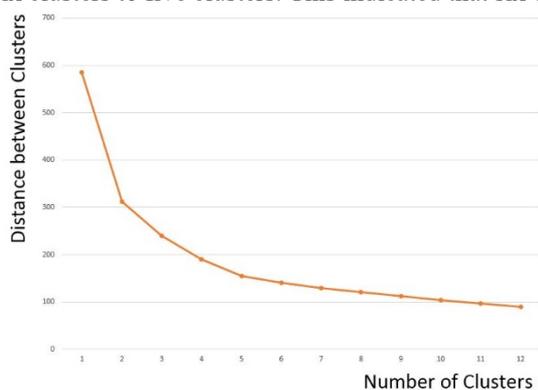


Figure 1. Distance between clusters.

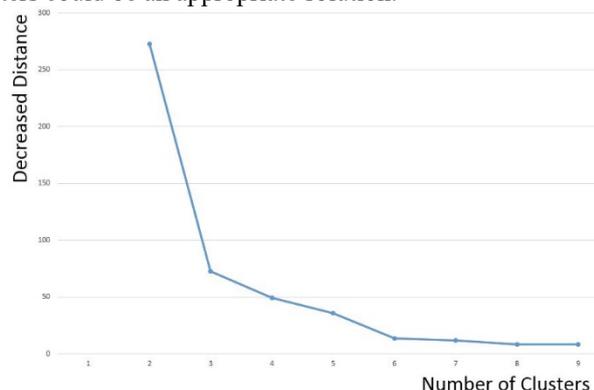


Figure 2. Distance change between clusters.

Figure 3 below demonstrates a six-group solution. Based on the performance trend, we interpret these six groups as follows:

1. *Group 1 – High performance group:* Students in group 1 started above the standard mean of the exam and ended also above the standard mean of the exam. There was a slightly increasing trend of the performance during the semester.
2. *Group 2 – Decreasing performance group:* Students in group 2 started at a high level, around one standard deviation above the standard mean of the exam, but ended around the standard mean of the exam. There was a clearly decreasing trend of the performance during the semester.
3. *Group 3 – Increasing performance group:* Students in group 3 started below the standard mean of the exam but ended above the standard mean of the exam. There was a clearly increasing trend of the performance during the semester.

4. *Group 4 – Decreasing performance group:* Students in group 4 started above the standard mean of the exam but ended below the standard mean of the exam. There was a clearly decreasing trend of the performance during the semester.
5. *Group 5 – Increasing performance group:* Although students in group 5 started below the standard mean of the exam and also ended below the standard mean, there was a slightly increasing trend of the performance during the semester.
6. *Group 6 – Low performance group:* Students in group 6 started below the standard mean of the exam and decreased to even lower grade at the end of the semester.

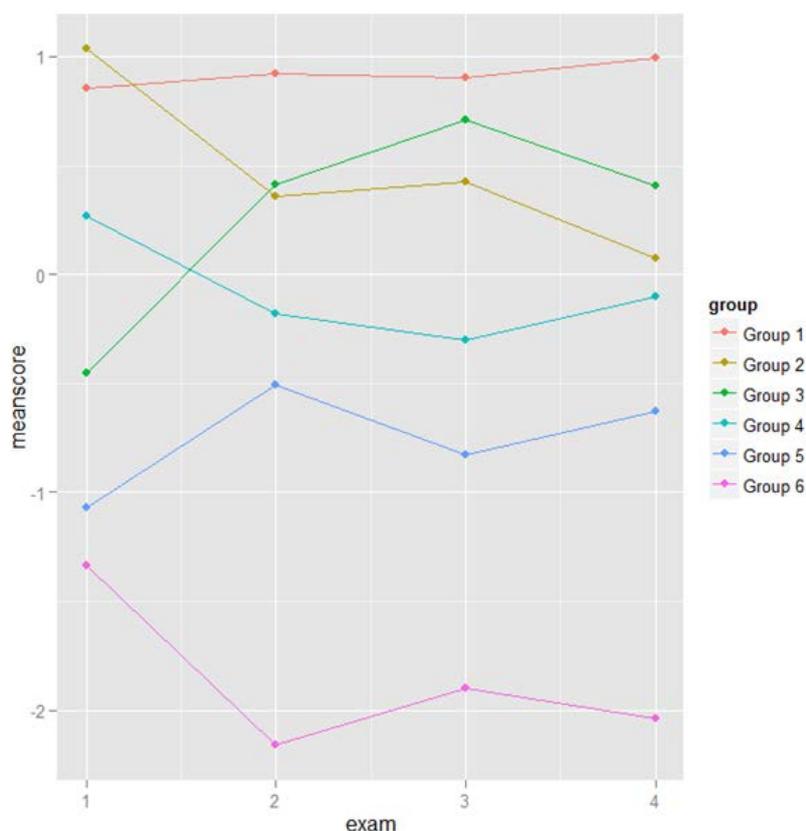


Figure 3. Students' performance trajectories in a semester.

ANOVA analysis findings

ANOVA analysis results showed that there were significant mean differences among the six performance trajectory groups in terms of math self-efficacy [$F(5,145)=10.493, p<.001$], the use of metacognitive strategy [$F(5,108)=3.959, p=.002$], the number of attempts for each online homework [$F(5,9666)=7.248, p<.001$], and the progress of watching each online lecture [$F(5,9676)=13.463, p<.001$]. No significant difference was found on the time spent on each online lecture across the six groups. The detailed descriptive statistics of each group including number of cases, mean, and standard deviation of each self-regulatory factor were summarized in Table 2.

In addition, Tukey's HSD post hoc test revealed more detailed information regarding the mean differences across six groups by pairwise comparison. Specifically, with respect to math self-efficacy, perceived math self-efficacy in group 1 was significantly higher than the rest groups and followed by the order: group 2, group 3 or 4, group 5, and group 6. No significant difference was detected between group 3 and 4. This result indicates that students in group 1 hold the highest level of confidence in their capabilities of learning math, while students in group 6 have the lowest level of confidence, and students in group 2 are more confident than those in group 3 and 4. Based on Figure 3, we can see the mean of math self-efficacy in each group roughly correlates with its first exam score, with a higher level math self-efficacy leading to a higher score in the first exam. But no such relationship is observed between the math self-efficacy and students' final exam score.

With respect to the use of metacognitive strategy, students in group 3, 1 and 5 were using significantly more metacognitive strategy than those in group 4, 2 and 6. Especially, students in group 3 reported to use the

metacognitive strategy most frequently, and students in group 2 reported to use at a very low frequency. It is interesting to notice in Figure 3 that the performance trajectory of group 3, 1 and 5 all presents an increasing trend, while the performance trajectory of group 4, 2, 6 all presents a decreasing trend. It seems that the performance trajectory may be positively related with the use of metacognitive strategy. The more frequency of the use of metacognitive strategy, the more likely the performance would be improved throughout the semester.

With respect to the two self-regulatory behaviors – attempts and progress, students in group 6 always ranked lowest among six groups on these two behaviors, indicating they enacted significantly less self-regulatory behaviors during the pre-class online learning than students in other groups. No significant differences were found among the rest groups in these two behaviors. Figure 4 demonstrates detailed weekly-self-regulatory behaviors of the six groups. As we can see in Figure 4, students in group 6 basically attempted least on the online homework in each week and also had the lowest complete view of the online lectures in each week. Based upon the performance trajectory in Figure 3, the results indicate that the lack of self-regulation on learning behaviors in the pre-class learning setting may lead to the failure of the course.

Table 2: Descriptive statistics of three self-regulatory factors in six performance trajectory groups

	Group 1 (n=38)		Group 2 (n=15)		Group 3 (n=25)		Group 4 (n=34)		Group 5 (n=24)		Group 6 (n=15)	
	Mean	SD										
MSE	5.87	.89	5.40	1.02	4.63	1.17	4.74	1.41	4.10	1.66	3.36	1.29
Metacog	4.35	2.24	2.05	2.24	4.38	1.87	3.23	2.10	3.69	2.10	1.96	2.10
Attempt	1.04	.94	.96	.91	.97	1.03	.99	.96	1.03	.96	.83	.89
Progress	.43	.50	.41	.49	.42	.49	.36	.48	.40	.49	.30	.46
Time	82.29	191.51	83.81	201.80	96.20	247.84	88.45	238.71	104.37	157.23	89.29	261.83

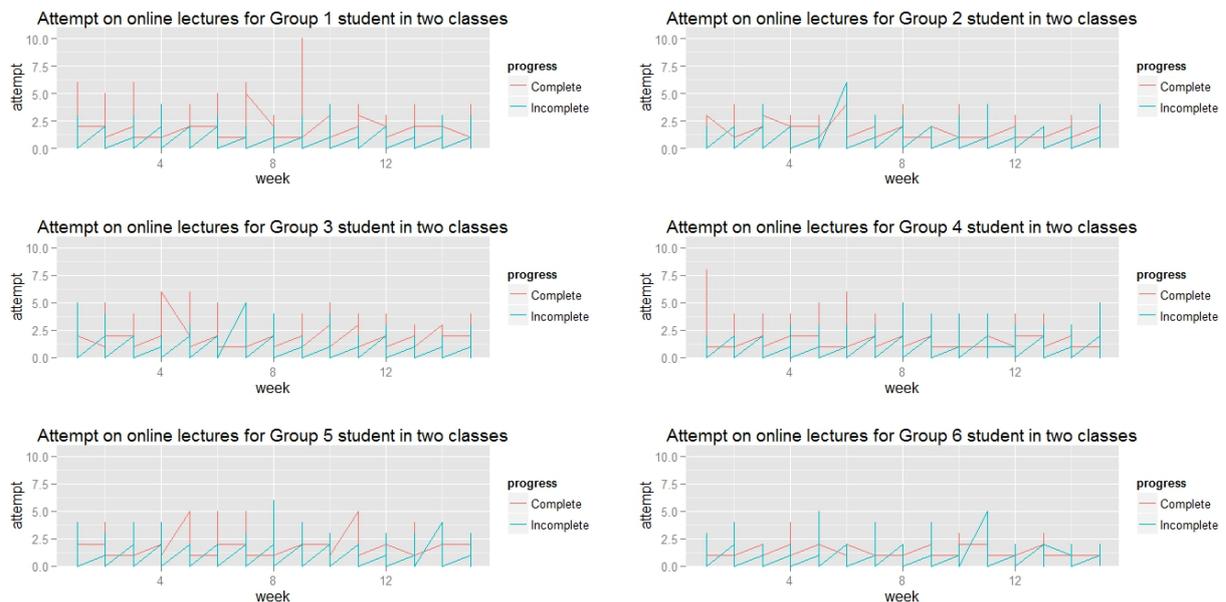


Figure 4. Students' weekly self-regulatory behaviors in a semester.

Conclusions and implications

The present study took an initial step to use both survey research and learning analytics methods to uncover the self-regulated learning process in the pre-class setting of the flipped classroom, and examined the relationship between the self-regulated learning process and students' performance trajectories during a semester in two flipped undergraduate math courses. The cluster analysis revealed six different performance trajectory groups among students. The analysis of survey and behavioral log data demonstrates that students in these six groups are significantly different in the self-regulation process during the pre-class learning. Specifically, math self-efficacy plays a foundational role in the influence of the performance trajectory. Students with higher level math self-efficacy tend to achieve higher in the first exam. However, such the positive relationship does not hold for the rest exams. For example, although students in group 2 perceive higher math self-efficacy than those in group 3, they,

however, achieve lower in all exams except the first one than students in group 3. With respect to the use of metacognitive strategy, students who reported as a frequent user of this strategy are more likely to have an increasing performance trajectory during the semester (e.g. students in group 3), while students who reported to rarely use this strategy tend to have a decreasing performance trajectory (e.g. students in group 2). With respect to the regulatory behaviors, students who attempt significantly less on the online homework and only complete a small proportion of the online lecture review (e.g. students in group 6) are very likely to fail in the flipped courses.

The findings from this study have several implications in the field of learning analytics and flipped classroom research. First, in this study, we have shown that learning analytics method offers a great potential to the discovery of behavioral and performance trend, such as the performance trajectory presented in this study. The examination of the trend, rather than a one-time achievement, provides a more detailed look of students' learning process during a semester and makes it possible to identify relationship between the change in the trend and the self-regulated learning process. Second, the findings in this study suggest that teachers in the flipped class should enact tactics to improve or sustain students' perceived math self-efficacy, such as complimenting their growth, attributing the poor performance to the lack of efforts, and encouraging them try harder. They should also pay more attention to students who rarely involve in the pre-class learning and may provide personal feedback to these students during in-class group work. Most importantly, teachers should implement strategies to enhance students' use of metacognitive strategy, such as providing metacognitive feedback to students when they got wrong in the embedded online quizzes, and helping students reflect on their pre-class learning while facilitating group work in class.

References

- Acedo, M. (2013, November 27). *10 pros and cons of a flipped classroom*. Retrieved from <http://www.teachthought.com/uncategorized/10-pros-cons-flipped-classroom/>
- Artino, Jr., Anthony, R. (2005). *Review of the Motivated Strategies for Learning Questionnaire*. Retrieved from http://www.sp.uconn.edu/~aja05001/comps/documents/MSLQ_Artino.pdf
- Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? the role of self-regulated learning. *Educational Psychologist, 40*(4), 199-209.
- Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*(3), 344-370.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. NJ: Prentice-Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. NY: Freeman/Times Books/Henry Holt & Co.
- Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. Eugene, OR: International Society for Technology in Education.
- Bishop, J. L., & Verleger, M. A. (2013, June 23). *The flipped classroom: A survey of the research*. Paper presented at 120th American Society for Engineering Education National Conference, Atlanta, GA. Retrieved from <http://www.studiesuccessho.nl/wp-content/uploads/2014/04/flipped-classroom-artikel.pdf>
- Carlson, S. M., & Moses, L. J. (2001). Individual difference in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032-1053.
- Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learning and Individual Differences, 21*(4), 337-346.
- Connor, P. K., Newman, D. L., & Deyoe, M. M. (2013). *Self-regulated learning and blended technology instruction in a flipped classroom*. Paper presented at 120th American Society for Engineering Education National Conference, Atlanta, GA.
- Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist, 40*(2), 117-128.
- Estes, M. D., Ingram, R., & Liu, J. C. (2014). A review of flipped classroom research, practice, and technologies. *International HETL Review, Volume 4*, Article 7,
- Flipped Learning Network (2015, April 29). *The four pillars of F-L-I-P*. Retrieved from <http://www.flippedlearning.org/definition>
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition Learning, 2*, 107-124.
- Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching, 42*(5), 62-66.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385-407.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

- Pintrich, P. R., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 731-810). Dordrecht, Netherlands: Springer.
- Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.
- Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement*, 53(3), 801-813.
- Schunk, D. H. (2001). *Self-regulation through goal setting*. Retrieved from <http://www.ericdigests.org/2002-4/goal.html>
- Shell, J. (2013, April 22). *What is a flipped classroom? (in 60 seconds)*. Retrieved from <http://blog.peerinstruction.net/2013/04/22/what-is-a-flipped-classroom-in-60-seconds/>
- Sun, Z. (2015). *The role of self-regulation on students' learning in an undergraduate flipped math class*. Retrieved from ProQuest Digital Dissertations.
- Talbert, R. (2014). Inverting the linear algebra classroom. *Primus*, 24(5), 361-374.
- Vellido, A., Castro, F., & Nebot, A. (2011). Clustering educational data In C. Romero, S. Ventura, M. Pechenizkiy, & R. S. J. d. Baker (Eds.), *Handbook of educational data mining* (pp. 75-92). Boca Raton, FL: CRC Press.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301), 236-244.
- Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*, 2, 153-189.
- Winne, P. H., & Hadwin, A. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (pp. 297-314). New York, NY: Lawrence Erlbaum Associates, Inc.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. *Metacognition in Educational Theory and Practice*, 93, 27-30.

Supporting Inquiry Learning as a Practice: A Practice Perspective on the Challenges of IBL Design, Implementation and Research Methodology

Fleur R. Prinsen, Welten Institute, Open University of the Netherlands, fleur.prinsen@ou.nl

Abstract: Inquiry Based Learning (IBL) provides learners with an authentic learning challenge and valuable opportunities to improve their scientific domain knowledge, while expanding their repertoire of creative skills and attitudes. But mastering the “what”, “why” and “how” of the cultural tools that science employs, and appropriating the structures and systems designed to enculturate them into communities of inquiry is no easy feat. Taking a practice perspective on IBL can illuminate some challenges that teachers, students and educational designers encounter, and may help researchers operationalize important variables, capturing the complexity of this learning practice. A new definition of IBL will be unpacked showing the added value of regarding it as a practice. This will inform the pedagogical design approach, but will also hold implications for methods of research. The feasibility of the ambition to enculturate learners into an authentic practice, providing them with transferrable skills, will be discussed.

Keywords: Inquiry Based Learning, practice, authentic, design, implementation

Introduction

Inquiry experiences can provide valuable opportunities for learners to improve their scientific domain knowledge and expand their repertoire of creative learning skills and attitudes (Saunders-Stewart, Gyles, & Shore, 2012). Mastering scientific inquiry means mastering the “what,” “why,” and “how” of the cultural tools that scientists employ (Lehrer and Palincsar, 2014). These tools (e.g. systematic ways of thinking, research instruments, and conventions like peer-review) are embedded within the activities through which scientists build their knowledge. When introducing Inquiry Based Learning (IBL) into the educational context, this represents considerable complexities and uncertainties for both teachers and learners. Still, IBL is increasingly valued and recognized as a powerful form of learning, fit for the 21st century (Hmelo-Silver, Duncan and Chinn, 2007).

According to Anderson’s (2002) definition “Inquiry science is a hands-on constructivist approach to science education. Students address teachers’ and students’ questions about natural phenomena or events by conducting scientific investigations in which they collaboratively develop plans, collect and explain evidence, connect the explanations to existing scientific knowledge, and communicate and justify the explanations” (Anderson, 2002). This definition specifies some explicit activities that students and teachers engage in. Students conduct collaborative scientific investigations, that mimic to some degree the practices of actual scientists who are co-creating new knowledge. Through their collaborative activities they practice more systematic ways of thinking (applying e.g. critical thinking, analytic thinking and information literacy), learn to use instruments (computer/technical) and practice their collaboration and communication skills. Through engaging in inquiry practices, students can become aware of the process of producing, testing, and revising knowledge and the criteria of evaluating knowledge claims (Smith, Maclin, Houghton, & Hennessey, 2000).

Still, the implementation of IBL often runs into problems. The richness in social meanings, embedded in the practices of scientific communities of inquiry, may not be easily grasped in the context of schools, since activities in that context draw on implicit or explicit understandings of how learning takes place that are not in direct accordance, or are discontinuous with the epistemology underlying the practice of IBL. But in the 21st century we should be preparing students to engage in a multiplicity of communities, and help them be aware how learning can happen across contexts (e.g. Kim, Hung, Jamaludin & Lim, 2014). Consequently, both teachers and students need to learn to balance and manage the tensions that arise when ‘stepping into’ a different context of learning (see also Hung, Ng, Koh, & Lim, 2009).

But how do we best explain the difficulties that learners have in mastering and appropriating (Wertsch & Rozin, 1998) the cultural tools of science inquiry? How do we get learners engaged in this complex process? What kind of conceptual and technological tools can we apply? An often cited publication by Edelson, Gordin, & Pea (1999) lists several challenges of IBL and how they might be addressed through technology and curriculum design. These challenges (motivation, accessibility of investigation techniques, background knowledge & skills, and practical constraints in the learning context) are still prevalent today. The current paper

problematizes the way in which these challenges are framed, and provides some thoughts towards explaining why students may not respond to pedagogical and technological IBL designs the way we anticipate. The aim is to introduce a relational epistemology of practice, and a new definition of IBL, which will help those involved (teachers, designers, researchers and learners) better understand and appropriate the structures and systems that are designed to help build their own communities of inquiry. Figure 1 provides an overview of the considerations which emanate from taking a practice perspective on IBL.

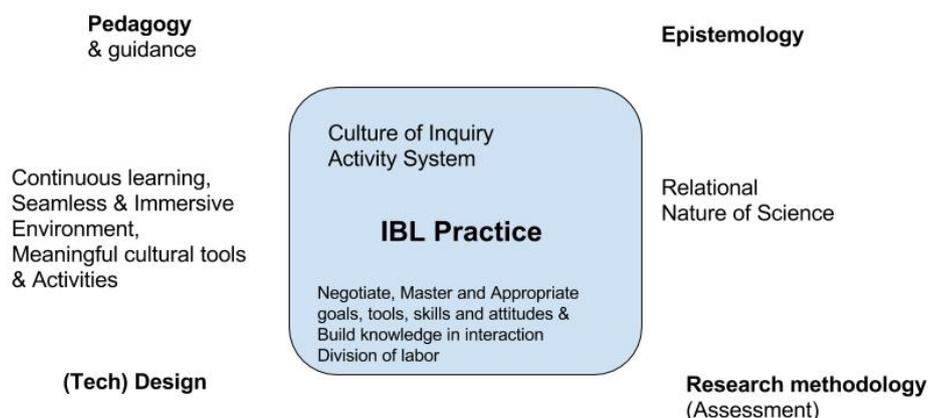


Figure 1. Epistemology, Design, Research and Pedagogy of IBL as Practice.

A relational epistemology of practice

IBL draws on several theoretical frameworks to explain how learning happens; for instance experiential learning theory (Healey, 2005), situated learning (Crawford, 2007), constructivist theories of learning (Hmelo-Silver et al., 2007), social learning theory (Bandura, 1986; West, 2009) and socio-cultural theory (Wortham, 2008; Kyza, 3012). These theories all consider aspects of epistemology. Epistemologies specify the nature of knowledge (e.g. certainty/diversity), how knowledge develops (ways of knowing, e.g. the relationship between the knowers and the known), and how it is justified (quality control). Our epistemologies influence how we, as teachers (Lunenburg, Ponte & van de Ven, 2007) researchers (Fenstermacher, 1994; Allert, 2010), designers (Chee, 2011) or learners (Hogan & Maglienti, 2001; Baek, & Schwarz, 2015), view the world and act in it. They influence how we see learning, how we study and support it, and how we (chose to) learn and assess.

If we want to support deep learning (involving critical thinking, analysis, and evaluation skills) and meaning making, our epistemological beliefs and those held by students matter. Said simply, if we think that all knowledge is already 'out there', that google is an authoritative source for true and certain knowledge, and that we can just look up THE answer to our questions, why would we even engage in inquiry as learners? Students' epistemologies are dependent on the contexts in which they encounter knowledge (Hammer and Elby, 2002; Sandoval and Morrison, 2003). Maybe we pay too little attention to the view of knowledge which is held in practical educational contexts, and to the implications this view has for our understanding of IBL and its practical design. IBL activities, when not properly embedded in (or as an expansion on) the primary context (Dohn, 2013) of learning, are at risk of not being experienced as fully meaningful. They are introduced as side-projects, after which we return to what is really at stake.

IBL most often involves collaborative arrangements, and many educational IBL designs involve technological support. Suthers (2006) discussed several epistemologies applicable to collaborative learning arrangements, supported by technological designs. He lays bare a variety of assumptions we may hold of what it means to learn in collaborative settings. A common account of learning seems to position itself somewhere between purely individual epistemologies (learning remains a process within individual minds) and completely intersubjective epistemologies (learning consists of interactions), stating that learning is a group activity that results in individual changes. In IBL approaches learning is often characterised as a form of collaborative knowledge construction (Stahl, 2000), implying an interactional constructivist epistemology.

So interaction between people leads to learning, but does this provide us with a complete picture of what is going on when we introduce IBL into classrooms? An intersubjective epistemology assumes a participatory process within which beliefs are enacted, without being necessarily mutually accepted. Intersubjectivity is a process of mutual constitution (Matusov, 1996, Wegerif, 2006). When students are negotiating their understandings of how they are supposed to learn (and which beliefs to hold in the process), this represents learning at an interpersonal level. To better support learning as becoming embedded within the

practice of a community, learning might additionally have to be addressed at the community level, where learning is a process of legitimate peripheral participation (Lave & Wenger, 1999). When IBL is first introduced we cannot assume that a community of scientific practice can be instantly created and that it will pursue externally driven goals; we should not ignore the critical issues of negotiated meaning, personal history and trajectories of learning (e.g. Henderson, 2015). Shared goals and expectations will have to be negotiated and cultural tools appropriated. Scientific epistemology refers to beliefs and views about how scientific knowledge is developed and justified, and involves a set of ideas and assumptions about the nature of science (Hogan and Maglienti 2001; Sandoval 2005). Even though it is difficult to attain the levels of epistemological understanding that professional scientists attain (Wu & Wu, 2011; Chuy et al., 2010), we could introduce students as legitimate peripheral participants (Lave & Wenger, 1999) in a community that develops scientific practice.

The technological tools we, as educational designers, introduce also need to be considered. These tools become similarly embedded in the practice of IBL, and they hold representations that externalise the designers' beliefs about what is important in IBL practice. The (broader) pedagogical design shapes, and to some extent may constrain the activities that students develop. Technology design can provide possibilities for certain (intersubjective) relations, but may not for others. Thus, the consequence is that an applicable epistemology includes the tools available and considers how the students relate to them. This epistemology is not only intersubjective, it is relational; a relational epistemology of practice. The epistemology of IBL is in essence that knowledge emerges, or is constructed, in relational interaction, between learners, more knowledgeable others, and the cultural resources of a community of practice (CoP). Based on the ideas presented about regarding IBL from a practice perspective, a new definition of IBL Practice is introduced:

Inquiry-Based Learning Practice: Engagement in continuous, immersive and meaningful inquiry activities, organised around recognised and shared goals, applying a repertoire of skills and attitudes necessary for participation in a community of inquiry, while negotiating, mastering and appropriating its cultural tools.

This expansive definition achieves a couple of things, which will now be demonstrated by unpacking and discussing its elements. First, it explicitly includes the cultural tools of a broad community of inquiry. When doing inquiry, students enter into a tradition of knowledge creation evolving for centuries. Being aware that we, as learners, are standing on the shoulders of giants, that we can take an active role in the co-generation of knowledge, science being collective in nature (Schibeci, 1986; Fleer, 2013) both humbles and empowers us. The definition further reminds us that knowledge is continuously being (re)developed. To speak of a practice, the participants need to hold some things in common, and the definition claims a clearly recognized common goal is necessary. The activities that we design for IBL may not initially be perceived as meaningful. Students are often task-oriented and are not often invited to consider if these tasks trigger meaningful activity for them. Meaning is created and negotiated in activity, and activity is formed by the meaning it instantiates and perpetuates (Dohn, 2014). The use of the term immersive reminds us that we should help students cross the boundaries that science's cultural tools might raise for them, not only by providing appropriate designs that support seamless learning interactions, but by examining with them the relations they are forming with each other and with the cultural tools of science. The referral to the 'community of inquiry' aims to evoke the concept of 'community of practice', not simply as an aspirational state of harmonious collaboration, an instrumentalist strategy (Henderson, 2015), or a way to achieve the adoption of (mobile) IBL technology, but to ensure we acknowledge the complex relationship between technology, learners, the collective and a 'given' socio-cultural context.

Design issues

So given the complex relationship between these elements, how do we design for IBL practice? According to Wenger (1998), practice is not a result of design but a response to design. Even if a set of procedures are imposed on the learners, the practices surrounding them will be a result of negotiated meaning by the participants. The activities students engage in are emergent from the design of tasks, the affordances of the learning environment (Suthers, 2006), and the division of labor. Participants are bound by socially constructed webs of belief (essential to understanding what they do), and are not simply connected through the tasks they are given (Brown, Collins, & Duguid 1989). This means that, even though the teacher and the mediating technology can provide cues for activities, how they will emerge remains to a certain degree uncertain. The consequence of this uncertainty is that it may be wise to apply broad principles providing teachers and educational designers with useful guidance and helpful foci, without constraining creativity. Educational design generally needs to be based on a pedagogical framework (e.g. Protopsaltis, Bedek, Kopeinik, Prinsen, & Parodi, 2015), but when implementing and adapting such a design, we should consider taking the practice of IBL as a starting point. This may prevent us from designing systems that do not amply (seamlessly) support the developing practice. A

learning environment, instead of being viewed as an application, might more productively be conceived of as an approach (Downes, 2005), that encourages the development of communities of inquiry. For instance, personal learning environments (PLE's) allow for a certain degree of flexibility, providing contextually appropriate toolsets that can be adjusted (e.g. by being modular) according to current needs and circumstances. Even if we cannot design learning itself, from a CoP stance we can design an environment that will either facilitate or frustrate emergent practices and identity (Wenger, 1998). The fact that practices are by nature in constant developmental flux means that IBL design will not necessarily progress toward an ideal or 'always successful' design, but that the possibility of adaptation will be key. User-driven development and formative evaluations (Bedek et al., 2015) can be adopted as part of the ongoing design process, allowing to pinpoint and help resolve arising issues in a process of collaboration between technology development and pedagogy.

IBL entails the application of a repertoire of skills and attitudes, and mastery of cultural tools. Extensive scaffolding is employed in inquiry learning approaches (Hmelo-Silver et al., 2007), reducing cognitive load through cognitive apprenticeship (Quintana et al., 2004), and allowing students to learn in complex domains. The main assumption behind such designs is that given structure and guidance, students will become increasingly accomplished problem-solvers. Scaffolding is provided through task structuring, coaching and the provision of hints. A literature review on potential types of guidance for supporting student inquiry when using for instance virtual and remote labs (Zacharia et al., 2015) identifies six types of guidance; constraining the process, using a performance dashboard, providing prompts, providing heuristics, scaffolding, and direct presentation of information. Care should be taken, though, to not forget that we are designing an environment that engages students in an ongoing social process and that their systematic ways of thinking are not just mediated by artefacts but also through dialogue about the reasons behind provided structure and guidance. Guidance thus also means that teachers provide time to pause and identify the rules and norms (conventions) governing the collective understanding in the moment, placing the perceptions and views of the members of the culture under study.

Since IBL is a social process, many current designs include feedback (e.g. dashboards) and reflection tools (e.g. digital notebooks, or reflection widgets) with which social factors and student perceptions can be examined. Additionally, the affordances of social media are increasingly integrated to support collaboration across time and space, and to provision more just in time support. The ways in which the inclusion of such media actually supports continuous, immersive and meaningful IBL activities is still under study.

Students' responses to technological designs are most often examined through usability measures, and the examination of cognitive load and student motivation. Sometimes engagement is examined through log-file data. Usability tests often does not include pedagogical criteria (Nokelainen, 2006), even though pedagogical usability should be central to any design effort. This brings us to further consider some issues relating to how the reconceptualization of IBL impacts on the ways in which we research the success of our designs.

(Research) Methodological issues

How do we research and capture inquiry learning as a practice? It is a challenge to move abstract conceptualisations into the realm of actual data. When we assert that "students' employment and understanding of provided cultural tools (elements of our designs) are embedded in, and interdependent with, complex social and cultural contexts (e.g. Rasmussen & Ludvigsen, 2010; Säljö 2010), it makes sense to follow students' activities in context and over time, to focus on students' interaction trajectories (Ludvigsen et al., 2011) and on the process of knowledge construction and meaning making. Situated learning provides us with a messy reality, but CoP (social practice theories) can have significant value as an analytical framework to understand it.

The CoP approach has been criticised as being too abstract, and consequently too difficult to operationalize (Storberg-Walker, 2008). It is clear that practice is not only the thing that members do (participate/activity), but also the ways in which they understand what they do (attributed meaning). Still, there is uncertainty as to the appropriate level of analysis (Storberg-Walker, 2008); are we analysing at the level of collective meaning (examining mutual engagement; interactions) or individual meaning (examining participation)? This is a multi-level issue, and a challenge to understand how an event at one level can shape an event at another.

By capturing interaction patterns (e.g. through network analysis, actor network analysis, and/or semantic analysis) and how they develop over time, we can focus on the relations being formed with provided (and emergent) cultural resources. This means combining quantitative and qualitative measures. While collecting evidence of students using the tools (e.g. through logs) and applying inquiry skills, we should also consider to what level they appropriate them; if they recognise their use and value.

Storberg-Walker (2008) provides, as one solution to the problem of capturing collective meaning making, the establishment of research to practice partnerships, in which practitioners take a leading role in

identifying, assessing, critiquing, and problematizing aspects of practice. It seems that not only teachers and students should engage in the ongoing social process of inquiry, but researchers should immerse themselves in the local practices more often and co-engage in the negotiation of the tools supporting inquiry practice. After all, they are themselves practical experts.

Implications for the educational context

The aforementioned issues have clear indications for the educational context. Introducing a new practice in this case means introducing a new view of knowledge and of how learning takes place. This requires explicit attention, since learners' epistemologies influence how they see and approach learning. Making time to explore with students their views on the Nature of Science can help bring to light their current conceptions (Wu & Wu, 2011) and introduce novel conceptions, more in line with how the source, the nature and the justification of knowledge is perceived in communities of inquiry.

Working from a relational epistemology will place the focus on the group as a learning agent. The application of social learning analytics (Buckingham Shum & Ferguson, 2012) can help shift the focus towards properties of learning which come into being through social activity. They can help examine how students are interdependent of each other in the IBL process.

Since IBL practice can only be emergent, as a response to educational design, it seems design can only provide scaffolds and cues for IBL activities, but cannot determine this response. The ways in which students respond needs additional monitoring, e.g. by providing them with opportunities to reflect on their experience. In this way they can explore how their personal viewpoints, feelings and values fit with the practices of science.

It is clear that especially the 'why' of the cultural tools needs explicit negotiation (discovering the reasoning behind the practice). It may help to provide students with exemplary cases and to discuss with them the role of activity structures, supporting systems and artefacts. Learning technologies form only one component of the IBL 'orchestration', that is, "the process of productively coordinating supportive interventions across multiple learning activities, occurring at multiple social levels" (Dillenbourg et al. 2009, 12).

Conclusion and discussion

The exposition above attempts to make the cases that all the uncertainty, discomfort and hard work that go into supporting IBL as a practice is worth it too make IBL a meaningful and productive experience for students. Of course the perspective laid out in this article will be disputed, and there are many remaining questions, for instance relating to the level of authenticity we can provide in and outside the classroom. Knight et al. (2014) endorse that students should be given the opportunities to engage in authentic learning challenges, wrestling with problems and engaging in practices increasingly close to the complexity they will confront when they move on from school. According to Knight et al. (2014) the focus on practice reflects this growing recognition in educational thinking. But it might be that there is no such thing as real authenticity in the context of our classrooms; are they by their nature artificial? In a draft version of a paper on 'Defining authenticity', Brown and Menasche (2006) argue for degrees of authenticity, rather than positing it as a binary concept (authentic or not authentic). They define three types of task authenticity: 'genuine', 'simulated' and 'pedagogical' and state that genuine authenticity "exists when learners engage in tasks in ways and for reasons they would in the real world" (Brown & Menasche, 2006, p.3).

But if IBL practice is considered as being in constant developmental flux, and can only approach authenticity, then what do we mean by taking 'IBL practice' as a starting point when implementing and adapting IBL designs in practice? We should realise that mimicking the actual practices of professionals is an unobtainable ideal in the context of e.g. primary and secondary schools. As Wu & Wu (2011) rightfully point out, when we aim to examine whether students apply the epistemological beliefs of professional scientists (i.e., formal epistemology), rather than the views and ideas they generated during inquiry practices in school (i.e., practical epistemology; Sandoval 2005), we will find that students still hold naïve epistemological views. Clearly we should take into account their proximal knowledge of the nature of science (Hogan 2000; Sandoval 2005). Their practical epistemology (involving ideas of the nature of knowledge, the approach of producing knowledge, and the criteria of evaluating knowledge) will more accurately reflect their decisions while constructing scientific knowledge and the criteria they apply in the evaluation of the knowledge which they are developing. Studies of how professionals learn in practice (Cheetham & Chivers, 2001) show that the point of becoming a fully competent professional is often not reached until long after professional training has ended, and reaching it involved a range of experiences in the authentic CoP context.

An additional concern that touches on learner participation in the practice of IBL is a concern over equity. Learners should be provided with learning opportunities that are equitable in access and quality. The

social and cultural dynamics of the collaborative practices within learning communities may have differential effects on learner participation and outcomes (e.g. Prinsen, Terwel, Zijlstra, & Volman, 2013). For instance, not all learners develop similar interests in science (Prinsen, 2012). According to Schreiner & Sjøberg (2007), interest in science can be conceptualized as a process of identification that develops within personally relevant social practices. The development of interest is a process that is mediated in various social contexts (Costa, 1995) and with various cultural artefacts, through participation in practices in which particular values are enacted (e.g., Bøe, Henriksen, Lyons & Schreiner, 2011). Those who, through practice, identify successfully with science persevere and re-engage with it over time. Hung and Chen (2007) develop a similar argument when they propose approaches, which emphasize the identity enculturation aspect in diverse communities.

Finally, design should take into account the developmental stages at which the IBL practice currently manifests itself. Schools, at any one point, have their own form of authentic practice which will have to be fit into a continuum of approaches that move towards the practices of actual professional scientific communities. Authentic learning challenges, according to Knight et al. (2014), provide the opportunity for developing transferable skills and competencies, and the qualities needed to thrive in complexity and uncertainty. An often heard question from teachers is “Should or should we not decrease the complexity of the IBL process?” Even though there are many challenges presented in this paper, we may have to accept the complexity of IBL and progressively support students in grappling with this complexity, without expecting them to ‘get it right’ too soon. The newly proposed definition would suggest that it is a continuous process, and if we want learning to become a life-long endeavour, we owe it to our students to let them work through the struggles they will inevitably encounter when they continue to inquire the phenomena of life.

References

- Allert, H. (2010). Coherent social systems for learning: An approach for contextualized and community-centred metadata. *Journal of interactive Media in Education*, 2004(1), Art-7.
- Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education*, 13(1), 1–12.
- Baek, H., & Schwarz, C. V. (2015). The influence of curriculum, instruction, technology, and social interactions on two fifth-grade students’ epistemologies in modeling throughout a model-based curriculum unit. *Journal of Science Education and Technology*, 24(2-3), 216-233.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bedeck, M. A., Firssova, O., Stefanova, E. P., Prinsen, F., & Chaimala, F. (2014). User-driven Development of an Inquiry-Based Learning Platform: Qualitative Formative Evaluations in weSPOT. *Interaction Design & Architecture (s)*, (23), 122-139.
- Bøe, M. V., Henriksen, E. K., Lyons, T., & Schreiner, C. (2011). Participation in science and technology: Young people’s achievement-related choices in late modern societies. *Studies in Science Education*, 47(1), 37-71.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, 18 (1), 32-42.
- Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In *Technology-enhanced learning*(pp. 3-19). Springer Netherlands.
- Chee, Y. S. (2011). Learning as becoming through performance, play and dialogue: A model of game-based learning with the game Legends of Alkhimia.
- Cheatham, G., & Chivers, G. (2001). How professionals learn in practice: an investigation of informal learning amongst people working in professions. *Journal of European Industrial Training*, 25(5), 247-292.
- Chuy, M., Scardamalia, M., Bereiter, C., Prinsen, F., Resendes, M., Messina, R., ... & Chow, A. (2010). Understanding the nature of science and scientific progress: A theory-building approach. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie*, 36(1).
- Costa, V. B. (1995). When science is “another world”: Relationships between worlds of family, friends, school, and science. *Science education*, 79(3), 313-333.
- Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of research in science teaching*, 44(4), 613-642.
- Dohn, N. B. (2014). Implications for Networked Learning of the ‘Practice’ Side of Social Practice Theories: A Tacit-Knowledge Perspective. In *The design, experience and practice of networked learning* (pp. 29-49). Springer International Publishing.
- Downes, S. (2005). Feature: E-learning 2.0. *Elearn magazine*, 1.

- Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the learning sciences*, 8(3-4), 391-450.
- Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. *Review of research in education*, 3-56.
- Fleer, M. (2013). Affective imagination in science education: Determining the emotional nature of scientific and technological learning of young children. *Research in Science Education*, 43(5), 2085-2106.
- Hammer D, Elby A (2002) On the form of a personal epistemology. In: Hofer BK, Pintrich PR (eds) Personal epistemology the psychology of beliefs about knowledge and knowing. L. Erlbaum, Mahwah, pp 171–192.
- Healey, M. (2005). Linking research and teaching exploring disciplinary spaces and the role of inquiry-based learning. *Reshaping the university: new relationships between research, scholarship and teaching*, 67-78.
- Henderson, M. (2015). The (mis)use of community of practice: Delusion, confusion and instrumentalism in educational technology research. In Scott Bulfin, Nicola F. Johnson and Chris Bigum (Eds). *Critical perspectives on education and technology*. New York: Palgrave Macmillan. pp.127-140.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99-107.
- Hogan, K. (2000). Exploring a process view of students' knowledge about the nature of science. *Science Education*, 84(1), 51–70.
- Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions. *Journal of Research in Science Teaching*, 38(6), 663-687.
- Hung, D., & Chen, D. T. V. (2007). Context–process authenticity in learning: implications for identity enculturation and boundary crossing. *Educational Technology Research and Development*, 55(2), 147-167.
- Hung, D., Ng, P.T., Koh, T.S., & Lim, S.H. (2009). The social practice of learning: A craft for the 21st century. *Asia Pacific Education Review*, 10, 205–214.
- Kim, M. S., Hung, W. L. D., Jamaludin, A. B., & Lim, S. H. (2014). Expanding “within context” to “across contexts” learning: a case study of informal and formal activities. *Interactive Learning Environments*, 22(6), 704-720., DOI: 10.1080/10494820.2012.745424
- Knight, S., Wise, A. F., Arastoopour, G., Williamson Shaffer, D., Buckingham Shum, S., & Kirschner, P. A. (2014). Analytics for learning and becoming in practice. In: *International Conference of the Learning Sciences (ICLS 2014)*, 23-27 June 2014, Boulder, Colorado, pp. 1680–1683.
- Kyza, E. A. (2013). Networked Technologies to Foster Students' Collaboration and Reflection. In *Emerging Technologies for the Classroom* (pp. 113-126). Springer New York.
- Lave, J., & Wenger, E. (1999). Legitimate peripheral participation. *Learners, learning and assessment, London: The Open University*, 83-89.
- Lehrer, R., & Palincsar, A. S. (Eds.). (2014). Investigating Participant Structures in the Context of Science Instruction: A Special Issue of Cognition and Instruction (Vol. 22). Routledge.
- Ludvigsen, S. R., Rasmussen, I., Kränge, I., Moen, A., & Middleton, D. (2011). Temporalities of learning in intersecting trajectories of participation. In S. R. Ludvigsen, A. Lund, I. Rasmussen, & R. Säljö (Eds.), *Learning across sites: New tools, infrastructures and practices*. London: Routledge.
- Lunenberg, M., Ponte, P., & Van De Ven, P. H. (2007). Why shouldn't teachers and teacher educators conduct research on their own practices? An epistemological exploration. *European educational research journal*, 6(1), 13-24.
- Matusov, E. (1996). Intersubjectivity without agreement. *Mind, Culture, and Activity*, 3(1), 25-45.
- Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Journal of Educational Technology & Society*, 9(2), 178-197.
- Prinsen, F.R. (2012). *Connective possibilities: Supporting the development of interest in science; inclusion of girls in (online) knowledge networks*. Approved ELS Starting grant application. Utrecht University. DOI 10.13140/RG.2.1.4563.4320.
- Prinsen, F. R., Terwel, J., Zijlstra, B. J., & Volman, M. M. (2013). The effects of guided elaboration in a CSCL programme on the learning outcomes of primary school students from Dutch and immigrant families. *Educational Research and Evaluation*, 19(1), 39-57.
- Protopsaltis, A., Bedek, M., Kopeinik, S., Prinsen, F., & Parodi, E. (2015). *DELIVERABLE D2.3.2: Revised Pedagogical and Diagnostic Frameworks*. Collaborative Grant Agreement No.: 318499 Last accessed: <http://portal.ou.nl/documents/7822028/5a4e60fd-8698-4e3c-ab1d-b9ecf4413e15>.

- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., ... & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The journal of the learning sciences*, 13(3), 337-386.
- Rasmussen, I., & Ludvigsen, S. (2010). Learning with computer tools and environments: A sociocultural perspective. *International handbook of psychology in education*, 399-435.
- Säljö, R. (2010). Digital tools and challenges to institutional traditions of learning: technologies, social memory and the performative nature of learning. *Journal of Computer Assisted Learning*, 26(1), 53-64.
- Schibeci, R. A. (1986). Images of science and scientists and science education. *Science Education*, 70(2), 139-149.
- Schreiner, C. & Sjøberg, S. (2007). Science education and youth's identity construction - two incompatible projects? In D. Corrigan J., Dillon, & R. Gunstone (Eds.), *The Re-emergence of Values in the Science Curriculum*. Rotterdam: Sense Publishers.
- Shum, S. B., & Ferguson, R. (2012). Social learning analytics. *Journal of educational technology & society*, 15(3), 3-26.
- Rogers, Y., & Price, S. (2008). The role of mobile devices in facilitating collaborative inquiry in situ. *Research and Practice in Technology Enhanced Learning*, 3(03), 209-229.
- Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, 89, 634-656.
- Sandoval WA, & Morrison K (2003) High school students' ideas about theories and theory change after a biological inquiry unit. *J Res Science Teach* 40(4):369-392. doi:10.1002/tea.10081
- Saunders-Stewart, K. S., Gyles, P. D., & Shore, B. M. (2012). Student Outcomes in Inquiry Instruction A Literature-Derived Inventory. *Journal of advanced academics*, 23(1), 5-31.
- Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction*, 18(3), 349-422.
- Stahl, G. (2000). A model of collaborative knowledge-building. In *Fourth international conference of the learning sciences* (Vol. 10, pp. 70-77). Mahwah, NJ: Erlbaum, 2000a.
- Storberg-Walker, J. (2008). Wenger's communities of practice revisited: A (failed?) exercise in applied communities of practice theory-building research. *Advances in Developing Human Resources*, 10(4), 555-577.
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 315-337.
- Suthers, D. D. (2005, May). Technology affordances for intersubjective learning: A thematic agenda for CSCL. In *Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!* (pp. 662-671). International Society of the Learning Sciences.
- Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: a literature review. *Educational technology research and development*, 63(2), 257-302.
- Wegerif, R. (2006). A dialogic understanding of the relationship between CSCL and teaching thinking skills. *International Journal of Computer-Supported Collaborative Learning*, 1(1), 143-157.
- Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems thinker*, 9(5), 2-3.
- Wertsch, J. & Rozin, M. (1998). *Mind as Action*. New York: Oxford University Press.
- West, R. E. (2009). What is shared? A framework for understanding shared innovation within communities. *Educational Technology Research and Development*, 57(3), 315-332.
- Wortham, D. W. (2008). *Activity systems in the inquiry classroom*. ProQuest.
- Wu, H. K., & Wu, C. L. (2011). Exploring the development of fifth graders' practical epistemologies and explanation skills in inquiry-based learning classrooms. *Research in Science Education*, 41(3), 319-340.
- Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S. A., Kamp, E.T., Mäeots, M., Siiman, L., & Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: a literature review. *Educational technology research and development*, 63(2), 257-302.

Introducing Academically Low-Performing Young Science Students to Practices of Science

Toi Sin Arvidsson, Columbia University, tya2102@tc.columbia.edu
Deanna Kuhn, Columbia University, dk100@columbia.edu

Abstract: Our objective was to engage low-achieving young adolescents in activities introducing them to practices of science. These extended beyond control of variables to include attribution and prediction in coordination of multiple variables affecting an outcome, as well as argument and counterargument in advancing and challenging claims. Social science content was used to help students see the purpose and value of scientific practices. The objective was largely met, as evidenced by two 6th grade classes ($n = 49$), both outperforming a control group ($n = 23$). Although students engaged successfully in argument and counterargument, less successful was meta-level reasoning about argumentation and nature of science. Importantly in its practical implications, one intervention group showed less gain in 10 45-min whole-class sessions than did the other group who engaged as pairs in the same sequence of activities over an average of six 24-min individualized sessions, suggesting the greater efficacy of individualized engagement.

Keywords: scientific thinking, scientific practice, low-achieving populations, multivariable reasoning, argumentation

Introduction

Attracting a broad range of students to the study of science has become an objective of increasing concern. To become educated in the practice of science, students need to engage in communities of scientific inquiry in which they develop shared goals, methods, norms, and, ultimately, it is hoped, values. In so doing they become participants in shared practices that are part of larger disciplinary norms and traditions they come to recognize and appreciate (Sandoval, 2014).

Implementing these objectives of science education are a tall order. *Science Education* editor John Rudolph notes that such objectives in fact go back as far as Dewey and that the problem lies not with these educational goals but with schools' seeming inability to devise workable school experiences to achieve them. Adding to the challenge is the restricted way in which scientific practice has been defined, both in the K-12 classroom and in educational research. Science practice typically has been taken to mean use of the scientific method, which in turn has been regarded as the design and analysis of a controlled experiment. Moreover, the experiment is a univariable one, its essence being the control (by holding constant) of variables (COV) in order to identify the effect of a single variable on an outcome. In the real world, in contrast, outcomes are most often caused not by a single cause but by multiple factors acting in concert, a fact that practicing scientists are well aware of and take into account in both their theoretical models and empirical investigations. The univariable logic and execution of COV represents at most one narrow slice of authentic scientific inquiry, and the most recent writing on developing children's competency in science emphasizes involving students not in acquisition of a tool kit of discrete skills such as COV but rather in the practice of science as an authentic, integrated whole.

This is the approach we have sought to implement in the work described here. Our approach is focused on getting students to experience that the methods of science have purpose that makes sense to them and hence are of value. Most important to the approach implemented here, then, is that students' activity be situated in the context of what students will see as a meaningful purpose and goal. In two initial studies, urban middle-school students addressed, for example, the topic of juvenile delinquency, among other social science topics. The reason for employing social science topics such as teen crime goes beyond the fact that such topics are ones our student population are familiar with. Although students will feel they already know something about them, they likely will not know that such topics are the stuff of science. What better way, then, to get them to see its power and relevance? In the course of such activities, students come to see how their (and others') beliefs about a phenomenon like teen crime are subject to influence by means of application of a scientific method.

In this study, we worked with students from a similar population over an extended period and hence were able to engage them in a way that captured and integrated multiple strands of scientific practice. We included COV but went beyond it, making it clear that multiple factors were likely contributors to the outcomes of concern and hence needed to be examined, taken into account, and their effects coordinated. A data analysis tool for K-12 students, InspireData, was integral to this objective as it allows students to visually represent the effects of multiple

factors. Students could then use this multivariable understanding to predict outcomes based on evidence across multiple variables and thereby achieve the larger goal of the activity – drawing on evidence, rather than only their own beliefs, as a source in seeking to understand the relationships being examined. Finally, we engaged them in scientific writing in the form of reports to the sponsoring foundation. This activity included addressing challenges to their claims, thus exercising skills of both argument and counterargument.

Our pedagogical method can be characterized as one of guided inquiry. The phases of the investigation were segmented for students into a sequence of component tasks, with care taken to make clear the purpose and goal of each one and its purpose within the larger task. Students are not given direct instruction as to strategies to apply to the component tasks; rather, attention is focused on the task goal and on their coming to recognize the weaknesses of inferior strategies they use in not achieving this goal. As a culminating activity, they reflect on how their final conclusions differ from their initially solicited beliefs about the roles of each of the factors. Doing so leads to reflection on the task as a whole and on how their evidence-based conclusions provide knowledge central to achieving the best task solution.

Participants came from the low SES, low-achieving middle-school population in which several researchers have reported it difficult to develop rudimentary scientific thinking skills, compared to success in doing so among more privileged groups (Siler et al., 2010; Lorch et al., 2010). Some of these low-performing students, for example, in seeking to design an experiment fail to manipulate the focal variable. This failure can be attributed at least in part to absence of more basic understanding of the purpose of scientific investigation as a) seeking to answer questions whose answers are not already known, and b) engaging in causal analysis, rather than seeking only to optimize outcomes. Furthermore, it is widely observed that students in such populations typically show little interest in or disposition to study science. This population, then, seemed to us an especially important one to reach and achieve success with, remaining mindful of the practicality of the methods examined for large-scale classroom use. We therefore included here a comparison of two parallel methods, identical except that one is administered to pairs of students by a researcher while the other is administered to a whole class by the classroom teacher.

Methods

Student and school sample

Participants were 72 students (38 females) from three 6th-grade and one 7th-grade science classes, all taught by the same teacher. Participating in pairs in a pair intervention condition were 25 students (12 females) from one 6th-grade classroom. An equivalent 6th-grade class of 24 students (12 females) participated in the same intervention as a whole class. Twenty-three students (14 females) drawn equally from another 6th-grade class and a 7th-grade class served as a control group and received only the post-intervention assessments.

A 10-item written multiple-choice task of a type commonly used for this purpose and administered at the beginning of the school year confirmed that students showed little mastery of the control of variables (COV) strategy, with a majority of students scoring no more than 50% correct and scores of 100% correct rare, a finding consistent with others for this population. Group comparisons on this test were non-significant.

Intervention procedure

The content of the intervention was identical across conditions except that the classroom group participated as a whole class led by the classroom teacher and assisted by the first author (whose presence enabled her to confirm fidelity of implementation) and assistant, while in the pair condition the pair worked in a corner of the classroom with a facilitator (the first author) present throughout the intervention. In the classroom condition, students worked with a partner for most activities. In both conditions, the adult scaffolded all activities using a planned protocol of prompts that did not provide direct instruction (unlike tutoring) or hints, but drew attention to the task goal and challenged the weaknesses of inferior strategies. The intervention was administered to the classroom group over 10 45-min class sessions over a period of 16 days. The number of sessions in the pair condition varied as it was tailored to students' progress. Of 13 pairs, one completed the intervention in four sessions and one completed it in five, while most of the rest took six sessions and four pairs required seven sessions. Sessions averaged about 24 minutes (allowing two per class period). These sessions took place over an average of 32 days, with a range from 14 to 59.

At the first session, the activity was introduced about an organization trying to recruit astronauts and therefore want to know what factors (fitness, parents' health, family size, and education, with family size being the only non-effective factor) matter to applicants' performance. Following this introduction, pairs were asked to record on a form which of the four factors they thought would and would not matter. In the classroom condition,

a tally across the class was shown, and in both groups it was noted that opinions differ. Students were then told to work as a team and that they must come to an agreement for everything they do.

Control of variables phase

In the first phase, students were given a set of records of a list of applicants. Each record showed the levels of a list of factors being investigated for each applicant. Students were told that if they studied the records carefully they could determine which factors make a difference to performance and which don't. Students may select any records and request the performance level of the applicant by first explaining what they planned to find out by evaluating the records. Students were then provided with the performance levels of the records.

Once a pair was certain they had reached a conclusion about a factor's status, they could enter it on a "Draft Memo" to the foundation director. In the pair condition, if the conclusion was a valid one based on a controlled comparison, the pair proceeded to choosing another factor to examine. If no controlled comparison existed allowing a valid inference, the adult embarked on a sequence of probe questions whose purpose was to support recognition of the limitation of the students' investigative approach in not yielding a definitive conclusion (e.g., "Couldn't it also be the difference in education that's leading to the different outcomes?"). No superior approaches were suggested, and scaffolding did not go beyond highlighting failure to achieve the goal (a definitive conclusion). In the case of valid conclusions, challenging probes were introduced, e.g., "Suppose someone disagrees with you and doesn't think that this factor makes a difference; what could you tell them to convince them?" In the classroom condition, this probing could not be conducted individually, but once per class session (typically at the end of the session) a whole-class discussion occurred that followed this model, using one pair's work as an example.

Once a pair in the pair condition and the majority of pairs in the classroom condition had achieved at least three controlled comparisons (showing fitness, a two-level factor, and education, the only three-level factor, effective and family size, a two-level factor, ineffective), pairs completed their final memo to the foundation director, indicating which factors applicants should be asked about and which they should not and justifying their recommendations with results from their investigations.

Multivariable coordination phase

Students at this point were ready to transition to the next phase in which they represented and reasoned about the influence of all of the factors operating at once. Students were then introduced to the representation of their data using charts generated by the program InspireData and told, "All of the cases that you and your classmates have looked at before are here." It was explained that a chart shows the performance levels of all applicants with each diamond representing a case and its identity seen by hovering over the diamond (Figure 1).

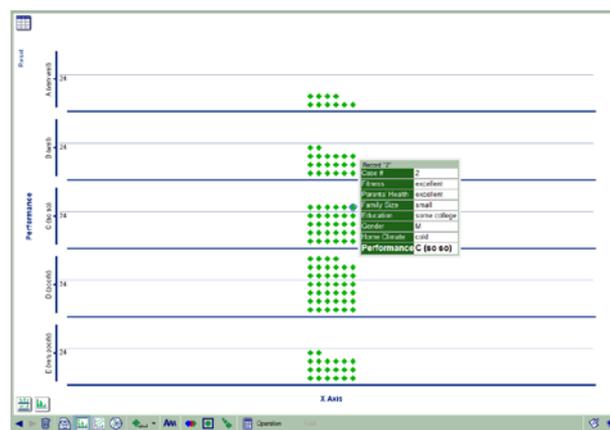


Figure 1. InspireData chart showing all cases.

It was then illustrated that charts can be generated to separate cases into different categories. A new chart with only those cases in which the applicant's fitness was average rather than excellent was shown. To highlight that other factors also contribute to the outcomes, students were then asked why it was that these applicants all of the same fitness level showed a range of performance outcomes.

Next, students were shown a third display in which all levels of the fitness variable are included. They were asked to draw conclusions about whether the factor makes a difference to the applicants' performance. Given the ability to see more data at once, students were asked to see if they reached the same conclusions as they did earlier when comparing individual cases presented on cards. Students were then provided InspireData charts for

each of five factors, four introduced previously and one new one (home climate, a non-effective factor), each of the same form, showing outcomes for all levels of the factor. In their pairs, they did this and then wrote memos to the foundation director confirming their earlier conclusions based on a larger sample or revising their conclusions if they thought necessary.

As in the previous phase of the intervention, prompts were introduced in the case of both correct and incorrect conclusions, e.g., “Suppose someone disagrees with you and doesn’t think that this factor makes a difference; what could you tell them to convince them?” In the classroom condition, this probing could not be conducted individually, but once per class session (typically at the end of the session) a whole-class discussion occurred that followed this model, using one pair’s work as an example.

Prediction phase

Students were told that now that they had reached final conclusions, they could try using them to evaluate a new set of applicants. They would then be able to select five to be chosen for the program and compare their choices to those of their classmates. Students were told that they could select up to four factors about the new applicants that they could receive information on. As students selected the factors, the adult reminded them to review the InspireData chart and consider whether knowledge of status on this factor would be informative as to outcome. In the classroom condition, a similar process took place at the whole-class level. Information about 10 new applicants on four factors (including one non-effective one, whether or not it was asked for) was presented one at a time. In addition to making each prediction, they were asked for each one, “Which of the four factors you have data on mattered to your prediction?” Students were encouraged to review the InspireData charts to double check their decisions or when there were disagreements between the student pair. In a final discussion pairs made selections of the five top-rated candidates and, in the classroom condition, shared these with the whole class. This discussion included remembering the beliefs they had initially held about the factors and noting that they would not have chosen the same applicants before and after the analysis they had conducted.

Post-intervention assessment

All post-intervention assessments except one were conducted individually and all were delayed in order to assess maintenance of achievements. Among students in the pair condition these assessments occurred an average of 26 days following completion of the intervention (range 14 to 42 days). Among students in the classroom condition they occurred an average of 32 days following completion of the intervention (range 18 to 46 days). Assessments for students in the control condition occurred during this same time period.

Findings

Designing experiments and making inferences

Final post-intervention achievement, maintenance and near transfer

The first component of the delayed post-intervention assessment was administered individually to students in both intervention conditions to assess the extent to which they maintained and transferred experimental design and inference skills (and specifically COV). The two items, each introduced a new variable (height and then strength) and asked students to select up to two cases to test its effect. Performance is summarized in Table 1. The large majority of both groups consistently designed comparisons that varied the focal variable. A majority in the pair condition also consistently designed controlled comparisons, as did half of those in the classroom condition, with the remainder doing so only inconsistently. The percentage of students who consistently designed controlled comparisons was higher in the pair than the classroom condition, $X^2(1) = 4.86, p = .027$.

Table 1: Comparison of intervention groups on maintenance and near transfer of design and inference skills

Skill levels	Pair condition	Classroom condition
Never varied focal variable	1 (4%)	3 (13%)
Varied focal variable only sometimes	2 (8%)	3 (12%)
Consistently varied focal variable but inconsistent control of other variables	2 (8%)	6 (25%)
Consistent controlled comparison	20 (80%)	12 (50%)

Note. $n = 25$ for the pair condition and 24 for the classroom condition.

Far transfer

The far transfer task, consisting of three written items of two levels of complexity, involving new content but assessing the same skills as the near transfer task. Performance by condition is summarized in Table 2. Thus, comparing tables 1 and 2, to an approximately equal extent across the two intervention conditions, fewer students maintained controlled comparison consistently in the far transfer context (table 2), when content was new and represented only in a traditional paper-and-pencil format. Results differed, however, by item complexity. For the two-variable items, the percentage of students who consistently chose controlled comparisons was significantly higher in the pair group than in the control group, $X^2(1) = 5.60, p = .018$, and marginally higher in the pair group than the classroom group, $X^2(1) = 3.50, p = .062$. The difference between classroom and control groups was non-significant for the two-variable items. However, for the more complex three-variable item, significantly higher percentages of students in both pair and classroom groups showed controlled comparison compared to the control group, $X^2(1) = 16.33, p < .001$ and $X^2(1) = 3.85, p = .050$, respectively. In addition, the percentage was higher in the pair than the classroom group, $X^2(1) = 5.13, p = .024$.

Table 2: Comparison of groups on far transfer of design and inference skills

Skill levels	Pair condition	Classroom condition	Control condition
Two-variable items			
Did not consistently vary focal variable	1 (4%)	12 (50%)	10 (43%)
Consistently varied focal variable but inconsistent control of other variables	9 (36%)	4 (17%)	7 (30%)
Consistent controlled comparison	15 (60%)	8 (33%)	6 (26%)
Three-variable item			
Did not vary focal variable	3 (12%)	4 (17%)	7 (30%)
Varied focal variable but inconsistent control of other variables	1 (4%)	7 (29%)	10 (43%)
Controlled comparison	21 (84%)	13 (54%)	6 (26%)

Note. $n = 25$ for the pair condition, 24 for the classroom condition and 23 for the control condition.

Multivariable analysis and prediction

This component, administered to all students, has previously been reported on (Kuhn, Ramsey, & Arvidsson, 2015). It presents (authentic although simplified) data about factors (Employment, Family size, Education, Home Climate) having an effect on average life expectancy across different countries and one non-contributing factor (Country size). Students were asked to predict life expectancy of additional countries based on information on their status on the identified factors. The task also asks respondents to indicate which factors they considered in their prediction.

Students overall did well on this task, indicating they understood the task and were capable of performing it, yet there were significant differences in performance across groups. Eighty-three percent of students in the pair condition, 63% of those in the classroom condition, and 22% of those in the control condition had modal performance of zero error (a correct prediction). Of remaining students who did not attain a modal performance level of zero error, all but one student showed a modal level of 1, with the remaining student showing a modal level of 2. When the mean prediction error scores over the six items were compared, the pair condition showed significantly less error than the classroom condition, $t = 2.78, df = 37.70, p = .009$. The classroom condition showed significantly less error than the control group, $t = 2.52, df = 42.05, p = .016$.

In attributing factors as having influenced their prediction, students again showed good performance overall but significant group differences. The pair group most often attributed influence to the four effective factors and least often to the ineffective factor. The control group were less likely than the pair group to attribute influence to each effective factor and more likely to attribute influence to the ineffective factor, with the classroom group intermediate but closer in performance to the pair group than to the control group.

Students were scored based on the consistency of their attributions as follows:

- 1-Chose only one but inconsistent factor across 6 countries
- 2-Chose only one consistent causal factor across 6 countries
- 3-Chose multiple but inconsistent factors across 6 countries
- 4-Chose multiple consistent (but not all four effective) factors across 6 countries
- 5-Chose four effective factors completely consistently across 6 countries

The mean difference in score between pair and classroom groups was significant, $t=3.32$, $df=42.38$, $p=.002$. The classroom group significantly outperformed the control group, $t=2.86$, $df=37.87$, $p=.007$, as did the pair group, $t=7.73$, $df=42.12$, $p<.001$. With respect to individual patterns of consistency in attributions, the same pattern of group differences appears. Only one student in the pair group attributed influence to the non-effective factor one time. In the classroom group, one student did so one time, while seven (29%) did so multiple times. In the control group, six (26%) students did so one time while 11 (47%) did so multiple times.

In the pair group, 18 students (72%) consistently attributed influence to all four effective factors and never to the ineffective factor. In the classroom group, only nine students (38%) showed this pattern, while 13 (54%) showed inconsistency in attribution across cases (i.e., a factor is in some cases claimed to have influenced a prediction and in other cases not). In the control group only two students (9%) showed the consistent pattern with the remaining 21 (91%) showing inconsistency.

With regard to number of factors to which influence was attributed, 83% (19) of the pair group, 54% (13) of the classroom group and 17% (4) of the control group most frequently correctly attributed influence to four factors. Only students in the control group—22% (5)—most frequently identified only a single factor as responsible for the outcome. (Remaining students most often chose 2 or 3 factors.). Pair and classroom groups performed significantly better than the control group in most frequently attributing to four factors with $p<.001$ and $p=.015$, respectively (Fisher's Exact test).

Argumentation

The final component of the delayed post-intervention assessment is an elaboration of the cancer task used by Kuhn et al. (2015). One of its parts pertains to types of counterargument and the other to reconciling divergent claims. Both are far transfer tasks as they bear no surface similarity to and make no reference to the intervention content. Furthermore, both ask students to reason about argument rather than only engage in it.

Evidence and counterargument

In the first part of the argumentation task, students were told:

The Public Health department of Portland, Ohio has noticed that the percentage of residents diagnosed with cancer is much higher in the inner city than in the outlying neighborhoods. The department is undertaking a study to find out why there are more people getting sick with cancer in the inner city than the outlying area.

The student was then asked to choose among four options that would constitute the strongest evidence to show someone was wrong who claimed that the difference was due to the fact that city people more often go to tanning salons. Results for students in the three conditions appear in Table 3. Shown are percentages of respondents choosing each option as providing the strongest counterargument to the claim that tanning salon use was a causal factor with respect to cancer rates. As seen in Table 3, most respondents chose B or C with a smaller proportion favoring A. All students appeared to recognize that D was irrelevant to the claim and none of them chose that option. Among the three relevant options, A, B, and C, Chi-squared goodness-of-fit tests showed that distribution of choices of students in the classroom and control conditions did not differ significantly from chance, $X^2(2) = 1.75$, $p = .417$ for the classroom condition, and $X^2(2) = 0.61$, $p = .738$ for the control condition. For the pair condition, significance was borderline, $X^2(2) = 5.83$, $p = .054$.

Table 3: Percentages of students making different response choices regarding counterargument

	Pair condition	Classroom condition	Control condition
A. Air pollution is a more likely cause of cancer in the city	4 (17%)	5 (21%)	6 (26%)
B. Many people outside the city also go to tanning salons and don't get cancer	6 (26%)	10 (42%)	9 (39%)
C. Many people who don't go to tanning salons also get cancer.	13 (57%)	9 (38%)	8 (35%)
D. There are more tanning salons outside the city than in the city.	0 (0%)	0 (0%)	0 (0%)

Note. $n = 23$ for the pair condition, 24 for the classroom condition and 23 for the control condition.

Reconciling claims

The final task continued the topic theme but was presented in writing in the classroom on an occasion about one and a half months after the other posttests, reducing possible influence of students' particular responses to the

previous task having the same theme. Two potentially conflicting causal claims are now explicitly presented and the question asks the participant how to interpret this discrepancy:

You were hired by the Health Department to find out why people living in the city of Logan, Georgia are getting cancer more often than people who live outside the city. You tested and found out that air pollution was worse inside the city than outside. You wrote a report of your findings to the Health Department director, telling her that air pollution was a likely cause of the increase in cancer.

She also got a report from another person she hired. This report said that a likely cause of the cancer increase was not enough stores in the city for people to buy healthy fruit and vegetables that lower risk of cancer.

The director isn't sure what to conclude and she has written you asking for advice. What would you write back? Give her the best advice you can.

This question, in contrast to the previous one, solicits reasoning not about the claims themselves but rather meta-level reasoning about their status in relation to one another and how the discrepancy between them is to be understood – a form of reasoning that is epistemological in nature and central to scientific practice. We expected that answers to this question would give us insight into students' epistemological understanding regarding the nature of science, more specifically the extent to which they understood it as an enterprise involving competing claims whose status evolves as more evidence is brought to bear on them.

Contrary to our expectations, a very large majority of participants did not address the question as one of how the divergence in claims is best understood and reconciled. Their answers thus did not bear directly on their understandings regarding the nature of the scientific enterprise, except in the negative sense of their not seeing the divergence as warranting attention or interpretation. Instead, students' dominant stance was one of how to use toward practical ends the information that had become available. Their understanding of the nature and objectives of science, in other words, remained one common among this age group – producing good outcomes rather than analysis of causes and effect. Thus, the objective students identified in the context of this question was to reduce cancer. None of the responses raise questions about whether the validity of the causal claims should be evaluated, rather than their being simply put into action.

Conclusions and implications

In light of a history of difficulty in effecting advances in higher-order thinking skills in the disadvantaged, low performing population studied here, the goal of this study was largely met. With 80% of students in the pair condition and 50% in the classroom condition consistently showing controlling across multiple tasks, these results compare favorably to previous efforts with similar populations devoted only to the COV strategy (Lorch et al., 2014; Siler et al., 2010). With unfamiliar material, the majority (60%) of the pair group maintained consistent control, while a third (33%) of the classroom group did so. These findings are consistent with the view that continued and varied experience across a succession of domains is necessary if consolidation of higher-level cognitive strategies (as assessed in tests of maintenance and far transfer) is to be achieved.

With respect to the less studied skills entailed in coordination of effects of multiple variables, both intervention groups displayed considerable mastery, maintenance, and transfer, especially in relation to the far from optimal performance shown by adults (Kuhn et al., 2015) as well as the present control group. Playing a critical role in this success, we believe, is the InspireData tool, as it affords students a visual representation of relations among variables (vanAmelsvoort, Andriessen, & Kanselaar, 2007). The charts allowed students to see and interpret a representation of the more common and realistic case of multiple variables in action together – a portrayal fundamental to scientific practice. A further contributing factor, we believe, is the social science frame of both intervention and transfer tasks that attached a readily understandable purpose and goal to the reasoning students were asked to do. Also, the frame of the present intervention task – optimizing astronaut selection – has the engineering focus of producing good outcomes, but it suggests a way that the two orientations, engineering and analysis, can be coordinated. The typical conception has been that the orientation of science students needs to progress away from the engineering focus and toward an analysis focus. Conceived differently, science students can come to appreciate that analysis is an essential tool in the achievement of engineering goals.

Results are more mixed with regard to argumentation and to developing meta-level thinking about argument and related understandings of nature of science. The intervention activities involved not only a repeated requirement to justify one's claims with appropriate evidence, but also repeated engagement with the probe, "Suppose someone disagreed with you..." which we expected to afford exercise in defending and supporting alternative claims using evidence-based arguments. This expectation was largely met within the intervention.

Students with practice became successful in meeting these demands and most often did so confidently by means of direct appeal to evidence, e.g., “I’d show them the results on the chart.”

Where students demonstrated less success was in extending these competencies to meta-level reflection about argumentation in new, far transfer contexts unrelated to the intervention. No group differences clearly appeared, despite the intervention students’ strong performance in taking into account the effects of multiple variables on an outcome, both within the intervention and in a transfer task with new content. However, they commonly did not recognize absence of an antecedent in the presence of the outcome (option C) as similarly weak in being consistent with an alternative factor having produced the outcome, choosing this option as often as the correct option B. Students’ understanding of the strength of various kinds of evidence in weakening (as opposed to supporting) a claim is thus fragile and in need of development, a finding consistent with research showing the use of evidence to weaken claims more difficult than its use to support claims (Kuhn et al., 2015).

Finally, despite their mastery of counterargument within the intervention, students’ performance was weak in the assessment of meta-level reasoning about argumentation in the final task, asking them to account for divergent claims. The large majority of students did not treat divergent causal claims as a cause for attention, examination, or attempted reconciliation. Instead, without acknowledging the divergence, they focused on one or the other or both imputed factors as causes worthy of action, without further investigation. This conception stands in stark contrast to recognizing diverging claims as needing to be examined and evaluated as an undertaking central to the practice of science.

It is now commonly emphasized how critical it is that students develop their understandings of the nature of science by engaging in it. The nature of science, most science educators now agree, cannot be taught in a deep way through passive direct instruction and rather must be experienced through students’ own activities. As students engage in a larger number of purposeful, goal-directed activities that involve science practices, over a range of content, they are in the best position to extract general attributes of these practices and to appreciate their value. This experience can only accrue gradually in a facilitative context. As well as the skill components involved in coordinating evidence and claims in the service of argument, scientific practice encompasses values and norms that come from participation with others in a community that upholds shared standards of knowing.

We turn finally to the comparison of our two intervention groups. Overall it was the individually instructed group who consistently showed superior performance, despite the lesser time (less than a third as many sessions on average) invested. This outcome, we believe, is one that has important practical implications, in speaking to the value of individualized instruction, in particular for the population we studied. Constant interaction with each other and the adult and required to think about and justify whatever they said at just the time they said it may have explained some of the differences.

The conclusion to be drawn, we believe, is that chronically disadvantaged, underserved, and underperforming students like the ones we worked with have the potential to be successful in developing higher-order intellectual skills, given sustained purposeful activity in an advantageous setting that engages students in dense exercise of reasoning. In current work, we are therefore exploring ways to make the protocol used here automated in a way that could make it practical for large-scale use. This is of course a sizeable step from personal conversation between peers and a more knowledgeable facilitator, but we believe it may be one worth pursuing, especially in seeking to reverse the persistent lack of success among low achieving students.

References

- Kuhn, D., Ramsey, S., & Arvidsson, T. S. (2015). Developing multivariable thinkers. *Cognitive Development, 35*, 92-110.
- Lorch Jr, R. F., Lorch, E. P., Calderhead, W. J., Dunlap, E. E., Hodell, E. C., & Freer, B. D. (2010). Learning the control of variables strategy in higher and lower achieving classrooms: Contributions of explicit instruction and experimentation. *Journal of Educational Psychology, 102*(1), 90-101.
- Lorch Jr, R. F., Lorch, E. P., Freer, B. D., Dunlap, E. E., Hodell, E. C., & Calderhead, W. J. (2014). Using valid and invalid experimental designs to teach the control of variables strategy in higher and lower achieving classrooms. *Journal of Educational Psychology, 106*(1), 18-35.
- Sandoval, W. (2014). Science education’s need for a theory of epistemological development. *Science Education, 98*, 383-387.
- Siler, S., Klahr, D., Magaro, C., Willows, K., & Mowery, D. (2010, January). Predictors of transfer of experimental design skills in elementary and middle school children. In *Intelligent Tutoring Systems* (pp. 198-208). Springer Berlin Heidelberg.
- van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (2007) Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *Journal of the Learning Sciences, 16*, 485-521.

Idea Identification and Analysis (I²A): A Search for Sustainable Promising Ideas Within Knowledge-Building Discourse

Vwen Yen Alwyn Lee, Nanyang Technological University, alwynlee@ntu.edu.sg
Seng Chee Tan, Nanyang Technological University, sengchee.tan@ntu.edu.sg
Joon Kit Kelvin Chee, Ministry of Education, chee_joon_kit_kelvin@moe.edu.sg

Abstract: This is a study on the identification and analysis of promising ideas, which examines the knowledge building discourse of 20 students within Knowledge Forum, through the I²A system in a mixed method approach. It is a challenge in providing support to maintain efforts in identifying and improving ideas, which this study addresses through visualization and step-wise analysis of discourse network measures that are calculated using the Knowledge Building Discourse Explorer. By using Social Network Analysis with chosen keywords and analyzing quality of notes, we are able to identify and trace the evolvement of ideas that become promising talking points within the discourse space. Findings show that ideas can be classified and further analyzed to have an impact on subsequent discourse. Teachers can focus their resources on promising ideas that help the community maintain engagement in creative work to improve ideas, so as to learn more effectively and attain knowledge creation goals.

Introduction

In educational institutions, learning approaches have long been operating within the acquisition and participation metaphors (Sfard, 1998), but there is also a new approach which is seen as an integration of both metaphors of learning (Paavola, Lipponen & Hakkarainen, 2004). Sfard (1998) examined the acquisition and participatory metaphors of learning and suggested that both cognitive and situative perspectives are required. Paavola et al. (2004) further suggested that the combination of both metaphors could be problematic, but an integration of both would require a third approach: the knowledge-creation metaphor of learning. Scardamalia and Bereiter's model of knowledge creation, termed Knowledge Building (Scardamalia & Bereiter, 2003), is one of the three influential models reviewed by Paavola et al. (2004), besides transformation of activity system (Engeström, 1999) and creation of new products at organizational level (Nonaka & Takeuchi, 1995). Knowledge building focuses on K-12 learning environment that leverages learner's natural idea generation capability to work on continual improvement of their ideas, and for teachers to maintain student engagement in idea improvement. Scardamalia and Bereiter (2003) distinguished between two modes of learning: namely, the belief mode and the design mode. The belief mode engages learners with what they or others believe or ought to believe, whereas the design mode engages learners with the usefulness, adequacy, improvability and developmental potential of ideas. Knowledge building as a social process makes use of both learning modes, where learners engage in collaborative inquiry to enable creation, contribution and advancement of community knowledge. The concept of idea improvement (Scardamalia, 2002) is crucial as students are required to have motivation in their search for solutions to handle significant challenges, be aware of emergent themes of inquiry from multiple sources of inputs, acknowledge knowledge gaps, and participate in collaborative idea improvement (Zhang, Scardamalia, Reeve & Messina, 2009).

In order to encourage and promote continuous efforts in idea improvement, we need a process of sharing and improvement of ideas within the community, and also to identify potential relevant ideas that show relative significance and promise to the community, such that these promising ideas could be brought forward and be continuously improved upon. Isolated suggestions and ideas from individuals are often unable to frame the problem accurately and cannot provide a complete picture of the ideal solution in most problems. However, students might not have the capability to continuously maintain engagement in the process of idea improvement over long periods of time. In such situations, teachers need to provide relevant support and scaffolds, but reading and monitoring student discourse can be challenging, especially when students' posts cumulate and become increasingly unmanageable over time. With appropriate tools and technology, such processes can be facilitated for greater efficiency and effectiveness. For example, technology and tools can be used for visualization and analysis, to automate the process of identifying promising ideas to allow efficient learning of content, and to also provide the ability to track the impact after the introduction of promising ideas into the discourse community. If this can be achieved, students would be able to develop ideas to attain knowledge creation goals within the classroom, and they would not be engaged in other peripheral and less important ideas, which could sideline and affect their immersion in progressive problem solving (Bereiter & Scardamalia, 1993). Identifying promising ideas within a discourse is a reasonable and natural approach towards efficient learning, as discourse often plays a creative role in encouraging improvements on ideas (Lakatos, 1970). Platforms such as Knowledge Forum

(Scardamalia, 2004) and software tools like Knowledge Building Discourse Explorer (KBDeX) (Oshima, Oshima, & Matsuzawa, 2012) are used for supporting knowledge building pedagogies, practices and research. These platforms and tools help to characterize constructive interaction between students, not limited to just social interactions but also allowing discourse to play a part in determining the nature of ideas that surfaced during discourse. By doing step-wise analysis of idea occurrences and improvements brought about by different discourse units, the flow and content of ideas can be predicted to improve over the discourse, which can be used as feedback to inform future classes and for contributing back to the community.

In this paper, we seek to identify relevant and promising ideas which may not be immediately apparent to teachers and students, and propose an automated method of tagging promising ideas, which was previously done through manual consensus by students (Chen, Chuy, Resendes, & Scardamalia, 2010). The proposed method of tagging promising ideas and analysis of subsequent discourse would help in deciding whether an idea has surfaced within the discourse, and if it has sufficient promise and potential to warrant timely interventions so as to sustain creative work in improving emergent promising ideas that eventually lead to community “rise-above” and more effective learning. Rise-above refers to the integration of ideas or elevation of explanation at a higher theoretical or principle level. The method is conducted in the following phases: (a) identification of non-obvious ideas among the large number of threads and chatter within online discourse, which might be missed by teachers or students; (b) determination of promising ideas that can sustain engagement of students and contribute to the community discourse; (c) evaluation of the impact and contribution of promising ideas to the community such that timely interventions could provide sustained students’ engagement in creative work to improve ideas.

The focus on learning process, discourse units and epistemic words

In the world of knowledge acquisition, designers construct methods to allow learners to attain pre-determined learning goals; whereas in the world of learning, designers focus on methods that support learning processes, rather than the acquisition of knowledge and pre-defined learning goals (Strijbos, Martens & Jochems, 2004). It is crucial to focus on the learning process in which students seem to benefit from, even though pre-planned outcomes might deviate. The learning process within online discourse is relevant to the concept of continuously improving and sharing of students’ ideas for the community to rise above, and this process is not just limited to the social interaction of students, but also applicable to the usage of discourse units (known as notes within Knowledge Forum) and words within the learning environment. The ability to identify promising ideas, ideas that become consequential when worked on, is essential for creative work with ideas at all levels (Chen, Scardamalia, & Resendes, 2013). This leads to the need to analyze text within online discourse, which are representations of opinions and views of students in the digital format, comparable to speech and presentations within traditional classroom discourse in the physical and audible form. There are various forms of discourse analysis, including automated content analysis that uses Natural Language Processing and Machine Learning techniques (Sun, Zhang, Jin, & Lyu, 2014) to focus on the appealing properties of unsupervised learning in topical models. The objective of probabilistic topic models is to automatically discover topics within a corpus and topic assignments of documents based on frequency of unique words. As compared to common choices of topic models such as latent semantic indexing (LSI) (Hofmann, 2001) and latent Dirichlet allocation (LDA) (Blei, 2012), this research focuses on unique words similar to Sun et al. (2014). In addition, we explore deeper than the concept of topics and themes by identifying the specific ideas that the writer is trying to portray through the linkage and usage of words at different temporal positions within the discourse.

Further, the analysis of discourse units and words reflect individuals’ contributions to community knowledge, which cannot be analyzed through pre- and post-tests (Oshima et al., 2012), while the knowledge levels that are attained by students are commonly investigated through activity coding such as the coding scheme conducted by van Aalst (2009) on secondary students’ written discourse. Apart from the common practice of using Social network Analysis (SNA) in analyzing social patterns of learners, this study focuses on emergent development patterns of ideas. By using SNA based on words that learners use in their discourse, ideas can be identified and the level of community knowledge obtained by learners could be represented in a novel way to reflect promising potential of emergent ideas at a point in time within the discourse. SNA could be used as a new representation of community knowledge building by learners (Oshima et al., 2012), which researchers can use to further investigate different models associated with the knowledge building community. KBDeX as a software tool was utilized in this research as a graphical visualization for identified ideas, and through our analysis, was used to support interpretations of promising ideas within a community discourse.

Methods

The proposed I²A system in this paper utilized knowledge building platform Knowledge Forum (Scardamalia & Bereiter, 2006) for conducting online discourse. KBDeX (Oshima et al., 2012) was used for analyzing knowledge

building discourse using social network measures, so as to identify ideas with their degree of promisingness which have a subsequent impact on community discourse. The discovery of promising ideas and their effects through social network analysis were subsequently qualitatively validated by analyzing the quality of notes.

Discourse platform and context

Using the Knowledge Forum as a collaborative knowledge building environment, this research analyzes the online discourse of 20 eighth-graders who were being taught the topic of “Human Transport System” over a period of two weeks. We worked closely with a middle school science teacher, who was able to plan and execute the knowledge building lesson for the students in a computer-aided environment. The teacher involved within the study also had prior knowledge building training and was able to facilitate the knowledge building lessons effectively, having previously used physical “idea cards” to trigger interactions and present authentic problems. Students had experience using Knowledge Forum and were keen to continue using it, as they found that the social interactions could help advance community knowledge. Each student was provided with a booklet on a fictional character “Uncle Yong”, who has an impending heart attack. This problem was authentic to some students as their family members had previously encountered such an issue, thus the problem acted as a trigger aimed at eliciting students’ ideas about heart problems. A motivated student (Student S2) posed four thought-invoking questions that were closely related to issues that his peers would comprehend. These questions were: *What heart problem is he facing? How can he improve on his diet and his lifestyle to get healthier? How will the operation be done? What will be the consequences if he continues to eat unhealthily?* Responses were posted as notes within Knowledge Forum on the “Human Transport System” view, which served as a workspace for all written notes within the two weeks learning session. The questions encouraged discussions and debates among students, with the usage of scaffolds such as “I need to understand” when posting notes of query, “My Theory” and “New Information” for statements of information that they have garnered online.

Archiving ideas and offline discussions into online notes on Knowledge Forum allow teachers and researchers to retrieve, assess and analyze the discourse and idea generation process. The discourse consist of 101 notes shared by 20 students, where all students contributed at least one note of significant content. Some students were very active within Knowledge Forum, constantly replying to other peers’ contributions and were responsible for the majority of build-on and rise-above notes. The text within each note in Knowledge Forum were imported into KBDEx for visualization and analysis, and the majority of these notes were individual expressions of one or more ideas, apart from the rise-above notes which were co-authored by multiple students and represented collective ideas. Every note that represented expression of ideas and used within KBDEx, would be termed as a “discourse unit”, or DU for short, represented visually as nodes in a network. Examples of such discourse units containing their respective authors and text are shown in Figure 1.

Name	Text
9	my theory - his blood vessels in the heart were damaged due to the blockage of fatty deposits, he finds it hard to breathe and he is having chest
10	i need to understand - is the plaque blocking the coronary artery the same as the plaque in your teeth?-

Figure 1. Excerpts of discourse units (DUs) posted by students 9 and 10.

Knowledge building discourse explorer (KBDEx) and BC trends

KBDEx was developed as a cohesive and easy-to-use platform that affords visualization of student interaction networks, discourse units network and words network (Oshima et al., 2012). Social Network Analysis was used to generate visualization of students’ contributions within the discourse and to transform the data into a graphical representation, where measurement indicators such as “density” and “centrality” are used to determine the level of interaction between students (Wortham, 1999). The commonly used SNA has been found to be insufficient in examining community knowledge advancement through students’ collaboration and interaction network (Oshima & Oshima, 2007). Interactions between words that learners use in their discourse are important but often neglected, as these words tend to reflect the learners’ knowledge and ideas during discourse of a topic. During our analysis using SNA, we used a conventional network measure, betweenness centrality (BC) coefficient, to indicate important connections between words within discourse units. KBDEx graphically displays BC values that allows us to identify and differentiate promising ideas. Our interpretation of the BC coefficient can be viewed as ideas’ degree of importance to multiple stakeholders within the discourse at different junctures of the discourse. When viewed over the entire discourse, the BC coefficients form a BC trend line (Figure 2), which suggests the ideas within a DU contain a certain degree of promise towards discourse, and has varying levels of significance and impact on the community discourse after the DU was introduced into the discourse space.

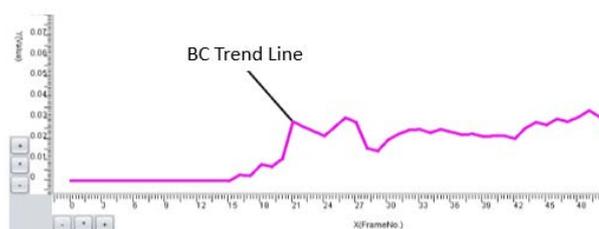


Figure 2. Example of a betweenness centrality (BC) trend line over the entire discourse.

Idea representation through keywords

Ideas are represented mainly through notes on Knowledge Forum, as students seek to query, inform or explain their ideas with their choice of known words. Keywords that are important for learning subject matter, managing learning, and gaining deeper understanding are known as epistemic words. We decided to focus on using objects and nouns to represent tangible items which students could relate easily to, given that the topic on human transport system was pegged at an introductory level. The process of choosing keywords was manually conducted by two experts, who were engaged to determine the list of keywords that reflect the level of knowledge, and are representative of key ideas regarding the topic of human transport system. The final list of keywords were identified with an inter-rater agreement of 82.4%, with differences between the experts resolved after further discussions. The final list of keywords contain 18 unique keywords with their respective 13 plural forms constituting a total of 31 keywords. The keywords are as follows: *red blood cells(s), white blood cell(s), platelet(s), oxygen, carbon dioxide, heart(s), lung(s), blood, blood vessel(s), artery/arteries, vein(s), capillary/capillaries, heart attack(s), heart disease(s), clot(s), antibody/antibodies, oxygenated, deoxygenated.*

Identifying relevant ideas using I²A

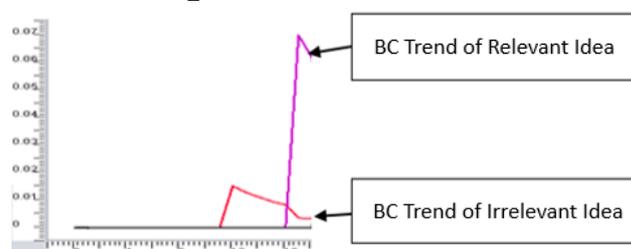


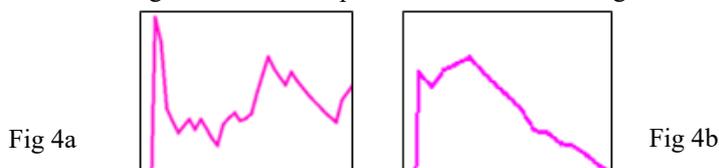
Figure 3. Example of BC trends for relevant and irrelevant idea.

The process of identifying ideas is dependent on the relevancy of contents within the discourse unit. The content of notes in Knowledge Forum form the basis of ideas within the discourse, and the usage of BC trends as a measurement index allows us to compare the relevancy of a discourse unit against other parts of the discourse. High betweenness centrality of a certain discourse unit at a certain juncture of discourse suggests that the selected discourse unit works as a key mediator that links other parts of the discourse through itself, and the idea contributed by the participant might represent an emergent idea that is worth exploring. When there are high values of BC with abnormal behavior within the BC plot, it suggests that a relevant idea could be present and is central at the particular point of the discourse, which learners should pay attention to. On the other hand, discourse units that possess smaller or insignificant values of BC could also represent the existence of less significant ideas that might not be relevant, since there is a lower extent or lack of community talk and discussion regarding the idea and the contents, hence indicating the presence of irrelevant ideas. However, if an idea has promising potential but does not possess sufficient relevancy to the community at the point in discourse, validation through qualitative analysis would then be required to prove and validate the idea's promisingness.

Detecting sustainability of interest in ideas using I²A

As ideas surface and are identified during discourse, it is likely one of two different scenarios could occur to it. If the idea comes across as a major eureka moment of novelty that makes sense and is also interesting to the community, continued discussion among students would most likely occur, with frequent queries and interests to keep the discussion alive with new perspectives, so as to continuously improve the idea and enhance community knowledge. As a result, after the initial spike in BC values, the continuous discovery of new ideas relevant to the original idea will cause the BC trend to decrease gradually with subsequent large BC fluctuations and increments across the remaining parts of the discourse as shown in Figure 4a. The other scenario depicts ideas that fall by the

wayside due to a lack of discussion and community disinterest. These ideas could either be deemed as common topics that were already thoroughly discussed, with convergence of opinions and consensus reached within the community; or it could also mean that potential promising ideas gained less traction than expected within the community. This can be similarly further verified using qualitative analysis. This situation where there is little to no community interest in continuing discussion of the idea at a specific discourse juncture, can be reflected in the BC plot (Figure 4b) with decreasing BC trend that tapers off or reduces to insignificant values.



Figures 4a and 4b. Examples of BC trends for promising interesting idea (left) and uninteresting idea (right).

Classification of ideas using I²A

Using the above idea attributes and combining the idea types mentioned in Figures 3, 4a and 4b, our observations in our analysis allow us to classify the pool of ideas within the discourse into four different classes, based on their relevancy and importance to discourse, sustainability of community interest, and impact to the community discourse (Figure 5). Different classes of ideas can surface within a discourse which affects the quality of talk and types of discussion. A relevant idea of sustainable community interest is classified as a promising idea, whereas if the idea has non-sustainable community interest, it is still a potential idea which the teacher can act upon, by engaging students and reviving interest in the topic. However, irrelevant ideas that retain continuous community interests often digress and affect the community’s direction in advancing knowledge, therefore requiring teachers to further examine affecting notes and their contents in detail, which can reveal possible distracting elements within the notes. A similarly irrelevant idea with minimal discussions and non-sustainable community interests would otherwise be considered a trivial idea. By classifying these ideas, the summation of the counts of different idea classes within the discourse could also provide a quantitative feel of the discourse’s quality.

		Relevancy to the Experts/Teachers	
		Relevant	Irrelevant
Interest to Community	Sustainable	Promising Ideas	Digressing Ideas
	Non-sustainable	Potential Ideas	Trivial Ideas

Figure 5. The four quadrants of idea classification

Findings and discussions

We ran the discourse through I²A using KBDeX for visualization, and were able to classify discourse units (DUs) and the respective contents into different idea types. This was done by using I²A to firstly identify ideas, followed by determination of relevant promising ideas, and lastly conducting an analysis of discourse belonging to different idea types for validation purposes. By tagging and classifying different ideas, teachers and students are able to learn more effectively and efficiently, by choosing to either prioritize improvement of relevant ideas to increase depth in knowledge, or to channel resources into broad based learning of all ideas.

Figure 6 displays three different examples of ideas that were identified throughout the discourse using I²A, along with our qualitative interpretations of how the idea was represented within the discourse unit. These ideas originated from notes extracted from the Knowledge Forum discourse and are an indication of different idea classes and types which can arise from a knowledge building discourse. The DUs were listed in chronological order and need not necessarily be in a conversation with the preceding or subsequent DU, since parallel discussions took place simultaneously across the community, with students responding to different threads over two weeks. The x-axis lists all the DUs within the discourse and are considered as turns in KBDeX, so that step-

wise analysis could be done to track the BC values for each DU, and later used for determining the impact of a certain DU or keyword on the community’s discourse. The y-axis indicates the normalized BC value throughout the whole discourse, which reflects the importance of the DU as a key mediator among other DUs, and is also representative of the DU’s centeredness within the network. The normalized BC values range from 0 to 1, with a higher score indicating higher centeredness of the DU and its greater role within the discourse. Based on the exemplified DUs and their respective classifications, such a semblance in other discourse can be similarly classified using I²A for most idea types, and allows the process to be reasonably scalable in longitudinal studies.

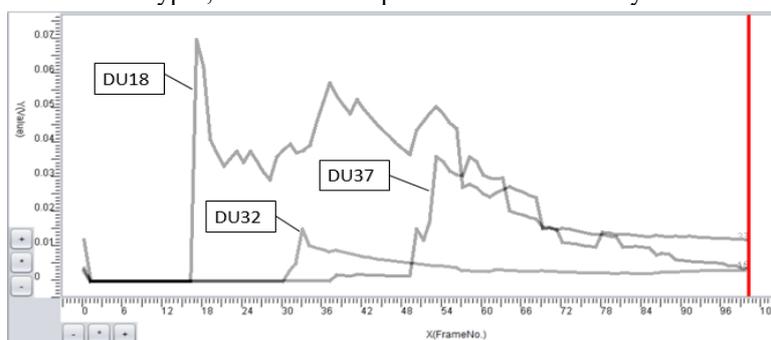


Figure 6. BC graph plot of DU18 (Promising Idea), DU32 (Trivial Idea), DU37 (Potential Idea).

DU32 - Largely irrelevant note with little sustained community interest – Trivial Ideas

S16: my theory - if the flow of oxygen-rich blood to your heart muscle is reduced or blocked, angina (an-ji-nuh or an-juh-nuh) or a heart attack can occur. angina is chest pain or discomfort. it may feel like pressure or squeezing in your chest. the pain also can occur in your shoulders, arms, neck, jaw, or back. angina pain may even feel like indigestion.

Explanation: The idea of DU32 was about physical effects due to blocked passages to the heart, which results in a heart attack. A new term “angina” was introduced, but was never picked up by other students for the remainder of the discourse. Even idea improvements by other students resulted in only a slight increase of BC, but the trend eventually subsided when no other references were made over the rest of the discourse. It was deduced that students reading the contents of DU32 were not spurred to further explore the term “angina” and the underlying reasons for the resulting physical body’s reactions. There was little to no interest regarding this idea as DU32 felt more like an informational statement rather than a trigger that stimulates motivation to improve on the idea of angina and its physical effects. Nonetheless, DU32 presented an idea that was slightly interesting to a small portion of the community who have no desire to improve it, and it was verified to be a trivial idea by the teacher.

DU37 - Relevant note with limited interests to community – Potential Ideas

S15: my theory - the components of blood include red blood cells, white blood cells, platelets, and plasma. some blood cells carry oxygen (necessary for metabolic reactions), some blood cells fight off invading substances that could destroy your cells, and other blood cells help to form clots, which keep your body from losing too much blood. the fluid portion of the blood carries nutrients needed to fuel each cell in the body. it also shuttles wastes that need to be transported to the excretory system to be passed out of the body and carbon dioxide that needs to be transported to the lungs to be exhaled. <http://www.dummies.com/how-to/content/what-are-the-components-of-human-blood.html> ~S15 & S12 –

Explanation: The notes leading up to DU37 provided the community with sufficient materials to craft a general idea (roles of blood components in the human transport system). The idea was improved with extra information about waste movements in DU37 itself and was further substantiated with a web link for verification and exploratory work. Curious students investigated the web link to gather more knowledge about the idea, evidenced by DU51’s addition of “blood cell ratios” as new information, and DU54’s web link with detailed explanations on “chances of infection”. The idea was continuously improved throughout the discourse before a final consensus was reached by the community in the form of a rise above note. The difference between this potential idea and other ideas was the additional work that students put into collating and interpreting of information from previous notes to form connections between keywords. Readers were able to easily link knowledge acquired from other DUs within the current context, and be motivated to continue seeking answers to queries that they generate from reading other students’ work. The web link’s presence served as an added incentive to search for more evidence

to verify theory proposed by other students. DU37 was considered a culmination of work from rise-above between two students that can provide a more relevant and idea-centric discussion as part of potentially promising ideas.

DU18 – Relevant promising note with high community interest – Promising Ideas

S11: clogged arteries can lead to multiple medical conditions. 1. coronary artery disease. when plaque accumulates in the arteries carrying blood to the heart, it results in coronary artery disease, which could lead to heart attacks. 2. carotid artery disease. the carotid arteries run up either side of your neck and supply oxygen to your brain. when arterial plaque accumulates in carotid arteries, it can lead to stroke.

Explanation: DU18 played an important role in the early part of the discourse as it was the key mediator in linking up two different groups of students that were having separate discussions, but talking about similar ideas such as the definition of arteries and atherosclerosis, but there was a missing aspect of the human transport system that was not discussed, i.e., the blood as a medium that transports resources to other parts of the body and the consequences due to deprivation of such resources. DU18 acted as a trigger that invoked student interests and allowed convergence of ideas into a single note, where students can pool their resources into improving the promising idea. Other than containing part of the idea regarding the circulatory transport system, the causes and effects were also linked together within the same note to prove a causal relationship that made sense to the students. As a result, subsequent discussions were used to describe and explain observations from a different viewpoint, where students were able to start looking at the big picture and strived to improve on the relevant idea instead of focusing on insignificant details. The subsequent contributions were mostly motivated inputs from students that started thinking deeply about the roles of blood constituents and how it affects the human body. The previously mentioned relevant idea in DU37 was directly influenced and traced back to DU18, where the idea had since improved and evolved to become the main idea in DU37.

This study of applying analytics showcases the ability to identify promising ideas currently present within the discourse and early identification of ideas which could evolve into promising ideas in later stages of the discourse. It provides an explanation of an idea's impact on subsequent discourse, since an earlier detected promising idea could have potential influence and ramifications on other ideas later in the discourse. A common observation of the discourse analysis shows that most of the BC values eventually taper off to lower BC values at the end of the discourse, as more notes were introduced into the discourse, and students within the community gradually reached a consensus on key ideas of the topic, resulting in lesser debates or conflicting perspectives regarding ideas, especially in situations where proposed ideas are substantiated with authoritative sources such as teachers or verified with multiple web resources. To measure the accuracy of automatic tagging of promising ideas, the identified ideas using I²A were compared against the teacher's list of ideas and keywords that should be acquired by students after the two week learning session. Most identified relevant ideas (88.2%), whether promising or not, were in concordance with the teacher's intention and pre-lesson learning objectives.

Conclusions and future directions

This current study introduces an initial effort in identifying and analyzing promising ideas within the knowledge building environment. By using SNA and measurement indicators such as betweenness centrality, we show that shifting analysis away from the student social interaction network to the idea network of discourse units allows us to identify and assess ideas along with their degree of promise. This work demonstrates the potential for automated tagging of promising ideas and the measurement of subsequent discourse impact on the community, which could significantly affect the community's rise-above. We believe that this research allows students and teachers to focus more on improving identified promising ideas beyond the current status quo, and also provides communities with the chance to advance knowledge by learning more efficiently, rather than spending time trying to identify possible relevant and promising ideas from a large pool of community ideas.

We acknowledge that there would be challenging work in the future, as we seek to expand our studies to students of different levels of expertise such as tertiary level students. We could track the level of understanding and types of ideas within students so as to address the evolution of ideas and learning processes from the learning journey of a novice into an expert, together with the exploration of possible uncovered idea types. The process of choosing keywords can be enhanced in subsequent studies by including verbs to the list of nouns, as verbs could represent actions and processes that link nouns together, similar to how mind maps function. Another potential direction of research is to further analyze the temporal dimension of the discourse data known as turns in KBDeX, which was briefly mentioned in the paper, so we can understand how SNA measures can affect the discourse with timely interventions. Having analyzed a knowledge building class in practice, future studies can also be conducted not limited to knowledge building communities, but also on other online discourse platforms.

References

- Bereiter, C. (2002). *Education and mind in the knowledge age*. Hillsdale, NJ: Lawrence Erlbaum.
- Bereiter, C., & Scardamalia, M. (1993). *Surpassing ourselves: an inquiry into the nature and implications of expertise*. Chicago: Open Court.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Chen, B., Chuy, M., Resendes, M., & Scardamalia, M. (2010). "Big Ideas Tool" as a new feature of knowledge forum. In *2010 Knowledge Building Summer Institute (KBSI)*. Toronto, Ontario, Canada.
- Chen, B., Scardamalia, M., & Resendes, M. (2013). *Dynamics of Promisingness Judgments in Knowledge Building Work of 8-to 10-Years-Olds*. Poster presented at the 2013 AERA Annual Meeting, San Francisco, CA.
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen & R.-L. Punamäki (Eds.). *Perspectives on activity theory* (pp. 19-38). Cambridge: Cambridge University Press.
- Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine learning*, 42(1-2), 177-196
- Lakatos, I. (1970). The methodology of scientific research programmes. In I. Lakatos and A. Musgrave (Eds.), *Criticism and the Growth of Knowledge* (pp. 91-195). Cambridge University Press, Cambridge.
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press.
- Oshima, J., Oshima, R., & Knowledge Forum Japan Research Group. (2007). Complex network theory approach to the assessment on collective knowledge advancement through scientific discourse in CSCL. In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of CSCL 2007* (pp. 563-565). Mahwah, NJ: Lawrence Erlbaum.
- Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. *Educational technology research and development*, 60(5), 903-921.
- Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of educational research*, 74(4), 557-576.
- Scardamalia, M. (2002). Collective Cognitive Responsibility for the Advancement of Knowledge. In B. Smith (Eds.), *Liberal Education in a Knowledge Society* (pp. 67-98). Chicago: Open Court.
- Scardamalia, M., Bereiter, C. (2003). Knowledge building. *Encyclopedia of Education*. New York: Macmillan, 1370-1373.
- Scardamalia, M (2004). CSILE/Knowledge Forum. In *Education and technology: An encyclopedia* (pp. 183-192). Santa Barbara: ABC-CLIO.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Eds.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York, NY: Cambridge University Press.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4-13.
- Strijbos, J. W., Martens, R. L., & Jochems, W. M. (2004). Designing for interaction: Six steps to designing computer-supported group-based learning. *Computers & Education*, 42(4), 403-424.
- Sun, W., Zhang, J., Jin, H., Lyu, S. (2014). Analyzing Online Knowledge-Building Discourse Using Probabilistic Topic Models. *International Conference of Learning Sciences*, 2, 823-830.
- van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 259-287.
- Wortham, D.W. (1999). Nodal and Matrix Analyses of Communication Patterns in Small Groups. In C. M. Hoadley, & J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference* (pp.681-686). Palo Alto, CA: Stanford University.
- Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge building communities. *Journal of the Learning Sciences*, 18(1), 7-44.

Acknowledgments

The research reported here was supported by the Centre for Research and Development in Learning, Nanyang Technological University, and we want to also thank Dr Teo Chew Lee and her team for the research funded by MOE Academies Fund NRF2012-EDU001-EL008, together with participating teachers and students in this study.

Training Learners to Self-Explain: Designing Instructions and Examples to Improve Problem Solving

Lauren E. Margulieux, Briana B. Morrison, Mark Guzdial, and Richard Catrambone
l.marg@gatech.edu, bmorrison@gatech.edu, guzdial@cc.gatech.edu, richard.catrambone@psych.gatech.edu
Georgia Institute of Technology

Abstract: In this experiment, we integrated two learning methods – subgoal learning and constructive learning – to explore their interactions and effects on solving computer programming problems. We taught learners to solve problems using worked example and practice problem pairs with one of three kinds of instructional design that either did not highlight the subgoals, described the subgoals, or prompted participants to describe the subgoals for themselves. In addition, we varied the distance of transfer between the worked example and practice problem pairs. We found that instructions that highlighted subgoals improved performance on later problem solving tasks. The groups that performed best were those that received subgoal descriptions with farther transfer between examples and practice problems and those that described subgoals for themselves with nearer transfer.

Keywords: worked examples, constructive learning, subgoal learning, self-explanation

Introduction

An important instructional tool for teaching problem solving in programming, and other science, technology, engineering, and math (STEM) domains, is the worked example. In this pedagogical approach, learners receive an example problem with the solution worked out (Renkl, Stark, Gruber, & Mandl, 1998). Ideally, students will use the worked examples to develop declarative rules or schemas that guide them in future problem solving. Empirical evidence has shown that learning with worked examples is more effective for acquiring problem solving skills than solving problems. This is called the worked example effect (Sweller, van Merriënboer, & Paas, 1998). However, research on the worked example effect has shown that merely presenting worked examples is not enough to promote student schema construction (Wittwer & Renkl, 2010). When studying examples, learners tend to focus on superficial features rather than the structural features because superficial features are easier to grasp and novices do not have the necessary domain knowledge to recognize the structural features of examples (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). For instance, when studying physics worked examples, learners are more likely to remember that the example has a ramp than that the example uses Newton's second law (Chi et al., 1989). A focus on superficial features leads to ineffective organization of information that, in turn, leads to ineffective recall and transfer (Bransford, Brown, & Cocking, 2000).

Subgoal learning

To promote deeper processing of worked examples, and thereby improve transfer, worked examples have been formatted to encourage subgoal learning by emphasizing the subgoals, or functional parts, of problem solving procedures to highlight the structural components of the problem solving process (Catrambone, 1998). Subgoals are the building blocks of procedural problem solving, and they are inherent in procedures. Each subgoal contains one or more steps. For the example in Figure 1, initializing the variables is a subgoal of the procedure used to solve problems with loops.

Research suggests that when instructions help students learn the subgoals of a procedure, students are better able to transfer knowledge to solve novel problems. To promote subgoal learning from worked examples, subgoal labeling has been used (e.g., Catrambone, 1998; Margulieux, Guzdial, & Catrambone, 2012). Subgoal labels are functional explanations that describe the purpose of a subgoal. For instance, in Figure 1 for the subgoal that initializes variables, the subgoal label might read "Initialize variables."

Subgoal labeled worked examples have improved problem solving performance in multiple STEM domains including statistics (Catrambone, 1998) and programming (Margulieux et al., 2012). Subgoal labels are believed to be effective because they visually group the steps of worked examples into subgoals and meaningfully label those groups (Atkinson, Catrambone, & Merrill, 2003). This format highlights the structure of examples, helping students focus on structural features and more effectively organize information (Atkinson, Derry, Renkl, & Wortham, 2000; Catrambone, 1998). Giving students subgoal labels, however, is a passive form of learning, and learning is generally more effective when students learn constructively (Chi, 2009).

Self-explanation

A common and effective type of constructive learning that might help learners understand subgoals is self-explanation. Self-explanation is a learning strategy in which students use prior knowledge and logical reasoning to make sense of information and gain knowledge. A review of self-explanation studies found it is effective across a range of domains as long as the domain has logical rules with few exceptions (Wylie & Chi, 2014).

Self-explanation of a worked example's solution identifies structural features and reasons about the function of the steps (Bielaczyc, Pirolli, & Brown, 1995). This purpose is similar to that of subgoal learning. By self-explaining worked examples, learners are more likely to recognize which features are structural and which are superficial. Learners, however, do not often engage in self-explanation on their own. Many studies (e.g., Chi et al., 1989) found that 10% or less of learners self-explained examples without external prompting. Much of the time, however, learners can self-explain if they devote additional resources to the task (Wylie & Chi, 2014) if they are reminded and guided to do so. Research has found little difference in the learning outcomes of students who are internally or externally prompted to self-explain, suggesting that self-explanation itself is the cause of learning benefits rather than learner characteristics (e.g., Bielaczyc et al., 1995).

Current research

The current research explores the effect of supporting learners to constructively develop their own subgoal labels through the process of self-explanation. We taught learners to solve problems using `while` loops with instructions that either did not have subgoal labels, had subgoal labels created by an instructional designer, or had placeholders for the student to generate their own subgoal labels (see Figure 1 for an example). The instructions included three worked examples and three practice problems.

No labels	Given labels (passive)	Placeholder for label (constructive)
sum = 0 lcv = 1	<u>Initialize Variables</u> sum = 0 lcv = 1	<u>Label 1:</u> _____ sum = 0 lcv = 1
WHILE lcv <= 100 DO	<u>Determine Loop Condition</u> WHILE lcv <= 100 DO	<u>Label 2:</u> _____ WHILE lcv <= 100 DO
lcv = lcv + 1	<u>Update Loop Variable</u> lcv = lcv + 1	<u>Label 3:</u> _____ lcv = lcv + 1
ENDWHILE	ENDWHILE	ENDWHILE

Figure 1. Partial worked example formatted with no labels, given labels, or placeholders for generated labels.

The worked examples and practice problems were interleaved so each worked example was paired with a similar practice problem. The practice problems either had isomorphic or contextual transfer from the worked examples. Isomorphic transfer meant that the worked example and practice problem were the same except for the values in each problem. For example, one worked example showed a program that would find the average tip amount for a restaurant server with the values \$15, \$5.50, \$6.75, etc. The paired practice problem with isomorphic transfer asked participants to find the average tip amount with the values \$20, \$8.25, \$9.75, etc. Alternatively, contextual transfer meant that the worked example and practice problem followed the same procedural steps but had different contexts. For example, for a worked example that found the average tip amount, the paired practice problem with contextual transfer asked participants to find average rainfall.

Giving learners practice problems to practice applying the procedure, even if the problems have minimal transfer from examples, allows students to monitor their learning and identify concepts that they superficially understand (Trafton & Reiser, 1993). The contextual transfer was intended to be harder for participants to map concepts from the worked example to the practice problem. More difficult mapping can improve learning by reducing illusions of understanding caused by shallow processing thus inducing deeper processing of information (Bjork, 1994; Eiriksdottir & Catrambone, 2011; Palmiter, Elkerton, & Baggett, 1991). However it can also increase cognitive load and potentially hinder learning by overloading cognitive resources (Sweller, 2010).

We hypothesized that students who generated subgoal labels would solve novel problems better than those who were given the subgoal labels, and both groups would solve problems better than those who had no subgoals at all. We also hypothesized that learners whose practice problems required contextual transfer would solve problems better than learners whose practice problems required only isomorphic transfer.

Methods

Materials

All participants received three worked examples and three practice problems. The examples demonstrated using `while` loops to solve problems that found the average amount of tips a restaurant server received (from an array of tip amounts), counted the number times a pair of dice rolled a 7 (from an array of dice rolls), and counted the number of prime numbers between 1 and 100. The isomorphic-transfer practice problems were in the same contexts, but they asked for the average tip amount (from a different array), the number of times a 2 was rolled (from the same array), and the number of prime numbers between 100 and 200, respectively. The contextual-transfer practice problems asked for the average amount of rainfall (from an array), the number of restaurants within 3 miles (from an array), and the number of unique phone numbers in a contact list (from an array), respectively.

Each participant received one of three formats for the worked examples. The first format did not highlight the subgoals of the procedure. The second format grouped individual steps of the example into subgoals and provided meaningful labels that described the function of each subgoal. This format is typical in subgoal label research (e.g., Catrambone, 1998; Margulieux et al., 2012). The third format grouped steps of the example into subgoals and provided a placeholder for participants to write their own labels. For this condition, each of the groups of steps was numbered as “label 1,” “label 2,” etc., and groups of steps that represented the same subgoal had the same number. For instance, groups that represented the “initialize variables” subgoal were called “label 1” regardless of where in the example they appeared. At the beginning of the session, participants who generated subgoals were told that each of the worked examples would have the same subgoals, and they were encouraged to update and improve upon their generated labels as they learned more about the procedure.

Mimicking the format of the worked examples, participants who received subgoal-oriented examples also received subgoal-oriented practice problems. If participants were given or generated subgoal labels in the examples, then the area in which participants solved practice problems was also structured with the given or generated subgoal labels, respectively. Instead of having a completely blank space to write the practice problem’s solution, like in the non-subgoal-oriented conditions, the subgoal-oriented conditions had several small blank spaces headed by subgoal labels or placeholders for labels. This design is typical of subgoal label research that uses practice problems (e.g., Margulieux et al., 2012) and was intended to support learners in initial problem solving and highlight connections between the examples and practice problems.

Participants assigned to generate their own subgoal labels received training on how to create subgoal labels. The training included expository instructions about generating subgoal labels, an example of a subgoal labeled worked example, and activities in which participants practiced generating subgoal labels and received feedback on their labels. The feedback was the same for all participants and asked them to compare the labels that they made to the labels that an instructional designer made. Participants who were not assigned to generate their own subgoal labels did not receive this training because it might have prompted them to generate their own labels, which would confound the results. Instead, these participants received training to complete verbal analogies (e.g., water : thirst :: food : hunger). Verbal analogies were considered a comparable task to subgoal label training because they both require analyzing text to determine an underlying structure.

After finishing the instructions (i.e., training, worked examples, and practice problems), participants completed novel programming tasks. The tasks asked participants to solve four novel problems using loops. Two of these problems required contextual transfer, meaning that they followed the same steps found in the instructions but had a different context (i.e., the same type of transfer as in contextual-transfer practice problems). The other two problems required both contextual and structural transfer, which is farther transfer than contextual. In these problems the context was new *and* the solution to the problem required a different structure than presented in the instructions. For example, the instructions included problems for averaging values, and the assessment included problems for averaging the first and second half of a list separately.

Design

The experiment was a 3x2, between-subjects, factorial design. Format of examples and practice problems (unlabeled vs. given subgoal labels vs. generate subgoal labels) was crossed with transfer distance between worked examples and practice problems (isomorphic vs. contextual transfer). The dependent variables were problem solving performance, quality of generated labels when applicable, and time on problem solving tasks.

Participants

Participants included in the final analyses were 120 students, 20 in each condition, from introductory programming courses in two technical universities in the Southeast United States (see Table 1 for demographics). Students were

offered credit for completing a lab activity or extra credit as compensation for participation. All students from these courses were allowed to participate, regardless of prior experience. To account for prior experience, participants were asked about their prior programming experience in high school and college and whether they had experience using `while` loops. Other demographic information collected included gender, age, academic major, high school GPA, college GPA, number of years in college, reported comfort with computers, expected difficulty of the programming task, and primary spoken language. There were no statistical differences among the groups for demographic data, which is expected because participants were randomly assigned to treatment groups. Participants also took a multiple-choice pre-test to measure problem solving performance for using `while` loops. Average scores on the pre-test were low, 1.6 out of 5 points, with 23% of participants scoring zero points. There were no correlations between pre-test score and format of worked examples and practice problems, $\rho = .07$, $p = .45$, or the transfer distance between them, $\rho = .01$, $p = .96$.

Table 1. Participant demographics

Age	Gender	GPA	Major
$M = 21.6$ years	71% male	$M = 3.2/4$	52% CS major

Many participants did not complete all tasks of the experiment. Participants received compensation regardless of the amount of time or effort that they devoted to the experiment, which might have caused low motivation in some participants. Participants who did not attempt all tasks ($n = 43$) were excluded from analysis. Participants who answered more than two questions correctly out of the five on the pre-test ($n = 12$) were also excluded from analysis because the instructions were designed for novices. Of the 175 students that participated in the experiment, 120 were included in final analyses.

Procedure

At the beginning of the session, participants completed the demographic questionnaire and pre-test. The pre-test had multiple choice questions about `while` loops from previous Advanced Placement Computer Science exams. Next, participants began the instructional period, which started with the subgoal label or analogy training. After the training, participants received the three worked example and practice problem pairs to help them learn to use `while` loops. When participants finished the instructions, they were asked to complete a 10 item survey designed to measure cognitive load while learning programming skills (Morrison, Dorn, & Guzdial, 2014). The placement of the survey at this point was to ensure measurement of cognitive load during the learning process and not during the assessments. Participants next completed the assessments. The assessments included four types of tasks, but the results of only the problem solving tasks, which were administered first, are discussed in this paper. Throughout the procedure, time on task was measured. Performance on activities in the subgoal label or analogy training and on practice problems was collected to ensure participants were completing tasks. The subgoal labels that participants generated were also recorded.

Results and discussion

Problem solving performance

Participants received a problem solving score based on the accuracy of their solutions. Participants earned one point for each correct line of code that they wrote, allowing for more sensitivity than scoring solutions as wholly right or wrong. If participants wrote lines that were conceptually correct but contained syntax errors (e.g., missing a parenthesis), they still received points. We scored logic errors (having `<` rather than `<=`) as incorrect. We considered scoring for conceptual and logical accuracy as more valuable than scoring for absolute accuracy because participants were in the early stages of learning. Participants could earn a maximum score of 44.

For problem solving performance among conditions, see Figure 2. We found a main effect of format of examples and practice problems, $F(2, 114) = 5.07$, $MSE = 176.5$, $p = .008$, est. $\omega^2 = .08$, $f = .21$. To explore this result, we conducted a post-hoc analysis with the LSD test because it is the most powerful for comparing three groups. We found that both subgoal-oriented formats (i.e., given or generate subgoal labels) performed better than the unlabeled group, mean difference = 7.8, $p = .01$, and mean difference = 8.6, $p = .005$, respectively. Both subgoal-oriented formats performed equally, mean difference = .78, $p = .80$. For transfer distance between examples and practice problems, we found no main effect, $F(2, 114) = 0.42$, $MSE = 176.5$, $p = .52$, est. $\omega^2 = .004$. These findings are tempered by an interaction between the two interventions.

We found a small, but interesting, interaction between the format of worked examples and practice problems and the transfer distance between them, $F(2, 114) = 2.71$, $MSE = 176.5$, $p = .071$, est. $\omega^2 = .05$, $f = .15$.

Though this interaction does not pass the threshold for statistical significance in the null hypothesis significance testing framework, the size of the effect makes it worth discussing. We found three levels of performance, as can be seen in Figure 2. The best performing groups were those that were given subgoal labels with contextual transfer ($M = 25.3$) and generated subgoal labels with isomorphic transfer ($M = 25.8$). The middle groups were those that received no subgoal labels with isomorphic transfer ($M = 16.9$), received labels with isomorphic transfer ($M = 18.9$), or generated subgoal labels with contextual transfer ($M = 19.9$). The worst performing group received no subgoal labels with contextual transfer ($M = 11.7$).

Each level of performance is separated by about seven points, or 16% of the total score. The difference between the middle and best level of performance was not statistically significant but had a medium effect size, as shown by the t-test comparing groups that were given subgoal labels (middle bars in Figure 2), $t(38) = 1.45$, $p = .15$, $d = .46$. Similarly, the difference between the middle and worst level of performance was not statistically significant but had a medium effect size, as shown by the t-test comparing groups that received labels with isomorphic transfer and that did not receive labels with contextual transfer (second and third bars from the left in Figure 2), $t(38) = 1.73$, $p = .09$, $d = .65$. Given these effect sizes, we would expect these differences to be statistically different with a sample size that was larger than 20 participants per group.

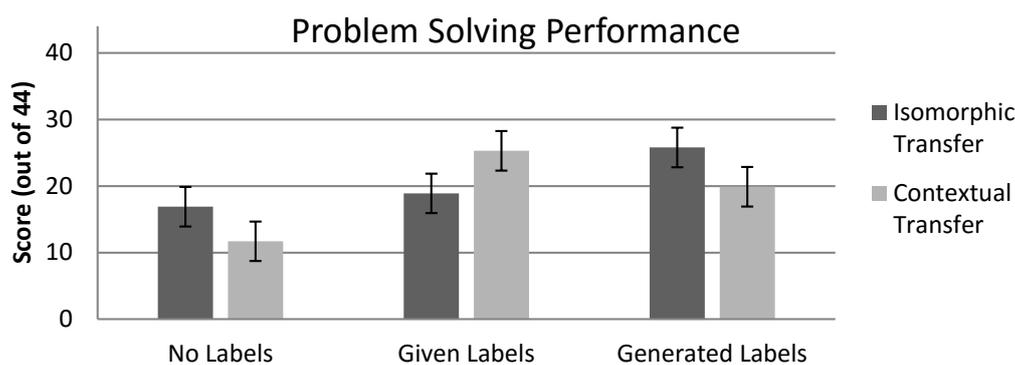


Figure 2. Performance for each group on problem solving tasks.

In summary, participants who received isomorphic transfer practice problems performed better than those who received contextual transfer unless they were given subgoal labels created by an instructional designer. This finding might be due to participants' mapping between worked examples and practice problems. In the isomorphic transfer conditions, it was obvious how the practice problems resembled the worked examples, allowing learners to easily apply the procedure from the example to solving the practice problem. Participants who received contextual transfer might have had difficulty mapping the example to the practice problem, which ultimately hindered learning unless they received subgoal labels that guided this transfer.

In general, participants who received subgoal-oriented instructions performed better than those who did not, suggesting that highlighting the subgoals of the procedure supported student learning. Which type of transfer was better for subgoal-oriented instructions depended on whether learners received or generated subgoal labels. Participants who received subgoal labels performed better with contextual transfer than with isomorphic transfer. This result might be due to contextual transfer allowing participants to build a more context independent understanding of the procedure, and receiving subgoal labels allowed participants to more easily map between the example and practice problems. In contrast, participants who generated subgoal labels performed better with isomorphic transfer than with contextual transfer. This result might be due to isomorphic transfer allowing participants to understand the connections between the examples and practice problems, making it easier to self-explain the subgoals of the procedure. Generating labels with contextual transfer might have overload cognitive resources enough to hinder learning.

Quality of learner-generated labels

We examined the subgoal labels that learners generated to explore the quality of labels that they produced. We used an iterative qualitative analysis in which we read a sample of participant responses to identify common themes then coded the data based on those themes (Braun & Clarke, 2006). We found there were two general types of labels: those including details that were specific to the worked examples and practice problems and those independent from the context. For instance, for the subgoal that initialized variables, two labels that were specific to the example in which participants calculated the average tip for a restaurant server are, "Establish container to hold tips" and "Create variable of tip values." Labels for the same subgoal that were not specific are, "Create

variables,” and “Define function and variables.” We found that, overall, twice as many participants generated specific labels ($n = 27$) than general labels ($n = 13$). However, a larger percentage of participants who received contextual transfer (40%) generated general labels than those who had isomorphic transfer (25%).

We explored how the specificity of generated labels affected problem solving performance and interacted with the transfer distance manipulation. The following results include data from only the generate subgoal label groups (i.e., the rightmost groups on Figure 2). We found that participants who generated general labels ($M = 25.8$, $SD = 13.4$) performed better than those who generated specific labels ($M = 19.8$, $SD = 13.7$), $F(1, 36) = 5.23$, $MSE = 144.6$, $p = .028$, $\text{est. } \omega^2 = .13$, $f = .36$. Similar to the general problem solving performance results, no main effect of transfer distance was found, $F(1, 36) = .92$, $MSE = 144.6$, $p = .35$, $\text{est. } \omega^2 = .02$. Again, these results are tempered by an interaction, this time between specificity of labels and transfer distance.

We found an interaction between specificity of labels and transfer distance (see Figure 3), $F(1, 36) = 5.52$, $MSE = 144.6$, $p = .024$, $\text{est. } \omega^2 = .13$, $f = .37$. There is not a difference between participants in the isomorphic transfer groups based on specificity of labels, $t(18) = .04$, $p = .97$. For the contextual transfer groups, however, participants who generated specific labels ($M = 12.1$, $SD = 9.4$) performed much worse than those who generated general labels ($M = 31.4$, $SD = 10.7$), $t(18) = 4.22$, $p = .001$, $d = 1.9$. To put these results in context, if we compare the scores of these two contextual transfer groups to the general problem solving performance results, the group that made specific labels performs on par with the lowest performing group (i.e., the unlabeled, contextual transfer group). In contrast, the group that made general labels performs six points (or 14% of the total score) better than the highest performing groups (i.e., the given labels, contextual transfer group and the generate labels, isomorphic transfer group).

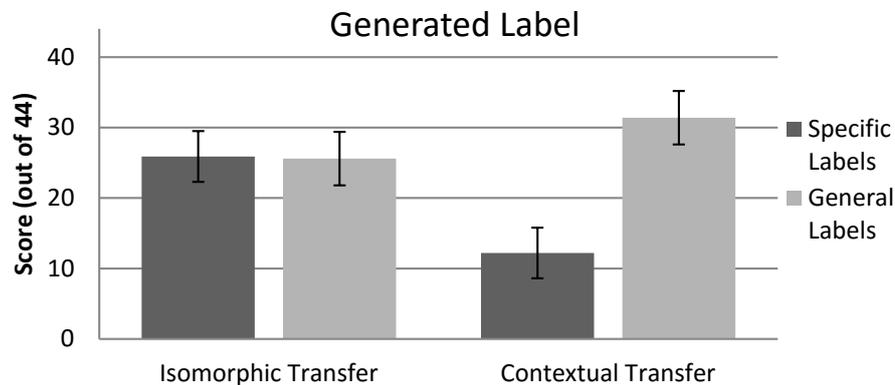


Figure 3. Performance for groups that generated subgoal labels on problem solving tasks split by transfer distance and specificity of generated labels.

In summary, participants who generated labels with isomorphic transfer performed relatively well, regardless of whether they created context-specific or general labels. For participants who generated labels with contextual transfer, however, their performance depends on whether they created specific or general labels. Those who created specific labels performed as poorly as the worst performing group, those who received no subgoal labels with contextual transfer. Participants in these groups were likely unable to discern the similarities between the examples and practice problems, which hindered their learning. On the other hand, participants who created general labels with contextual transfer performed better than any other group. This condition gave participants the most freedom to figure out the subgoals of the procedure for themselves, and if they were able to discover the context-independent subgoals, then they were better able to solve new problems.

We explored whether these higher achieving participants were simply those students who perform well regardless of the learning conditions. We found that 40% of participants in the generate subgoal labels with contextual transfer created general subgoal labels. This percentage is higher than the typical 10% of students who perform well in all learning conditions (e.g., Chi et al., 1989). We also found that students were more likely to create general labels if they had a high college GPA, $r_s = .44$, $p = .008$, or high school GPA, $r_s = .44$, $p = .01$. Based on these results, we concluded that higher achieving students were more likely to be successful in this condition but in higher numbers than would be expected if the instructional intervention did not affect learning.

Time on task

We measured the amount of time that participants spent completing the problem solving tasks during the assessment. Those who received contextual transfer in the instructions completed the tasks faster than those who received isomorphic transfer ($M = 16.9$ minutes, $SD = 10.8$), $F(2, 114) = 4.18$, $MSE = 78.9$, $p = .043$, $\text{est. } \omega^2 =$

.04, $f = .18$. There was no main effect for format of instructions, $F(2, 114) = 0.38$, $MSE = 78.9$, $p = .69$, est. $\omega^2 = .007$. There was an interaction between the two manipulations, $F(2, 114) = 4.11$, $MSE = 78.9$, $p = .019$, est. $\omega^2 = .07$, $f = .18$.

The pattern of results for time on task was almost identical to the pattern of results for problem solving performance. Participants who performed better took longer to complete the tasks. For example, participants who received contextual transfer finished the tasks more quickly and performed worse, except for those who received subgoal labels (see Table 2). The exception to this similar pattern was that participants who received no labels and isomorphic transfer took longer than other groups who performed better (i.e., groups that received labels with contextual transfer and that generated labels with isomorphic transfer). We examined the data for outliers that might skew the means, but we found no participants who spent a very short (i.e., less than 50% of the mean) or very long (i.e., more than 150% of the mean) amount of time on the tasks.

Based on these results, we conclude that completing the problem solving tasks correctly necessarily took longer than completing them incorrectly, but those who completed the tasks incorrectly devoted sufficient time attempting to achieve the correct answer. Alternatively, receiving contextual transfer during the instructions might have helped participants to apply their knowledge more quickly to the problem solving tasks, resulting in less time on task. Because these participants tended to perform worse on the tasks, we do not find this explanation likely.

Table 2. Time spent on problem solving tasks for each group (in minutes)

Group	No Labels	Given Labels	Generate Labels
Isomorphic Transfer	$M = 20.0, SD = 12.9$	$M = 13.3, SD = 8.3$	$M = 17.5, SD = 9.9$
Contextual Transfer	$M = 10.7, SD = 6.2$	$M = 15.4, SD = 5.9$	$M = 14.7, SD = 8.0$

Conclusions and implications

One of the biggest challenges in constructive learning is providing learners with enough support so they do not flounder but not so much support that they miss opportunities to construct knowledge. We found two types of instructional design that supported learning better than the others. Learners who received contextual transfer between worked examples and practice problems performed best when they were given meaningful subgoal labels. The labels likely guided the learners to recognize the similarities of the procedure between the two contexts. Learners who received isomorphic transfer performed just as well when they were guided to generate their own subgoal labels. This finding may be due to minimal transfer requiring less cognitive effort, allowing spare working memory capacity to be devoted to developing subgoal labels and improved learning.

Only a subgroup of another condition performed better than participants in these conditions: learners who received contextual transfer and generated subgoal labels that were not specific to the context of the worked examples and practice problems. This finding suggests that allowing learners who are capable of generating abstract labels with minimal guidance to do so is best for their learning. However, learners who were not capable of generating abstract labels performed as poorly as those in the lowest scoring condition. This result makes giving students this minimal amount of guidance risky, and until the factors that would predict success better are better understood, we do not recommend using this particular instructional design.

We recognize some limitations that affect the generalizability of our results. We did not measure learners' cognitive fatigue throughout the experiment, but we expect that it could be high, especially for those generating labels. Cognitively demanding tasks, such as constructing knowledge, could result in learners taking breaks during the learning process or becoming de-motivated, which might affect performance. Because we did not measure learner fatigue, break times, or related constructs, we do not know how they affected the results. In addition, for our analysis the quality of subgoal labels generated by participants, the sample size within each of those subgroups is small. For example, the number of people in the contextual transfer condition who generated context-independent labels was eight. Though the difference between these groups was large, it is possible that it is unreliable and that the actual effect size is smaller.

For future research, we plan to explore the factors that make students more or less successful in this paradigm that mixes subgoal learning and constructive learning. Perhaps we could improve the training for generating subgoals and, in turn, improve the learning of people who create their own labels. In addition, perhaps we could predict before learning begins which type of instruction will help the student to be most successful. For example, perhaps students with lower working memory capacity would perform the best when given subgoal labels and contextual transfer and students with higher working memory capacity would perform best when allowed to generate labels with isomorphic practice problems. Until we discover these predictive variables, we can conclude that learners can be successful when they generate their own subgoal labels, but only if they receive enough guidance from the instructions to support their constructive learning.

References

- Atkinson, R. K., Catrambone, R., & Merrill, M. M. (2003). Aiding transfer in statistics: Examining the use of conceptually oriented equations and elaborations during subgoal learning. *Journal of Educational Psychology, 95*(4), 762-773.
- Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of the Educational Research, 70*(2), 181.
- Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13*(2), 221-252. doi:10.1207/s1532690xci1302_3
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. J. Metcalfe & A. P. Shimamura (Eds.). *Metacognition: Knowing about Knowing*. MIT Press: Cambridge, MA.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (2000). How people learn: Brain, mind, experience, and school: Expanded edition. Retrieved from http://www.nap.edu/catalog.php?record_id=9853
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101.
- Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127*, 355-376. doi:10.1037/0096-3445.127.4.355
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73-105.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.
- Eiriksdottir, E., & Catrambone, R. (2011). Procedural instructions, principles, and examples: How to structure instructions for procedural tasks to enhance performance, learning, and transfer. *Human Factors, 53*(6), 749-770. doi:10.1177/0018720811419154
- Margulieux, L. E., Guzdial, M., & Catrambone, R. (2012). Subgoal-labeled instructional material improves performance and transfer in learning to develop mobile applications. *Proceedings of the Ninth Annual International Conference on ICER* (pp. 71-78). New York, NY: Association for Computing Machinery.
- Morrison, B. B., Dorn, B., & Guzdial, M. (2014). Measuring cognitive load in introductory CS: adaptation of an instrument. In *Proceedings of the Tenth Annual Conference on International Computing Education Research*, 131-138.
- Palmiter, S., Elkerton, J., & Baggett, P. (1991). Animated demonstrations versus written instructions for learning procedural tasks: A preliminary investigation. *International Journal of Man-Machine Studies, 34*, 687-701. doi:10.1016/0020-7373(91)90019-4
- Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*, 90-108.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123-138.
- Sweller, J., van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.
- Trafton, J. G., & Reiser, B. J. (1993). The contributions of studying examples and solving problems to skill acquisition. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 1017-1022). Boulder, CO.
- Wittwer, J., & Renkl, A. (2010). How effective are instructional explanations in example-based learning? A meta-analytic review. *Educational Psychology Review, 22*, 393-409.
- Wylie, R., & Chi, M. T. H. (2014). The self-explanation principle in multimedia learning. In R. Mayer (Ed.) *The Cambridge Handbook of Multimedia Learning, 2nd Edition* (pp.413-432). Cambridge University Press.

Acknowledgments

We would like to thank the students who participated in the study and their instructors who gave us their time. We also thank the reviewers who supplied comments which improved this paper. This work is funded by the National Science Foundation under grant 1138378. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

Validating a Model for Assessing Teacher's Adaptive Expertise With Computer Supported Complex Systems Curricula and Its Relationship to Student Learning Outcomes

Susan A. Yoon, Jessica Koehler-Yom, Emma Anderson, and Chad Evans
yoonasa@upenn.edu, jkoehl@gse.upenn.edu, emmaa@gse.upenn.edu, echad@sas.upenn.edu
University of Pennsylvania

Abstract: The success of the *Next Generation Science Standards* (NGSS) and similar reforms is contingent upon the quality of teaching, yet the shifts in teaching practice required are substantial. In this study, we propose and validate a model of adaptive expertise needed for teachers to successfully deliver NGSS-informed computer supported complex systems curricula in high school science classrooms. The model is comprised of three research-based qualities that we hypothesize teachers need to demonstrate: *flexibility*, *deeper level understanding*, and *deliberate practice*, in adapting interventions to their particular teaching contexts. We apply the model to understand whether it can, a) show variation between teachers, and b) predict student-learning outcomes. Results show that teacher enactments and interview responses assessed on the model reveal sufficient variation, and are also predictive of students' growth in complex systems understanding. The model has important implications for how to support teachers in adopting new science education reforms.

Keywords: teacher's adaptive expertise, computer supported curricula, complex systems, NGSS

Introduction

The recently released *Next Generation Sciences Standards* (NGSS) in the US provides a vision for science education reform that includes greater emphasis on constructing scientific models, using computational learning supports, and understanding systems and system interactions. The success of the NGSS and similar reforms is contingent upon the quality of teaching, yet the shifts in teaching practice required are substantial (Reiser, 2013; Wilson, 2013). It is also widely known that situational issues like lack of adequate resources, and few high quality professional development (PD) opportunities make adopting reforms challenging (Blandford, 2012; Calabrese Barton, 2007). Furthermore, given the complex nature of science teaching (Wilson, 2013), the extent to which teachers are able to adapt their practice and demonstrate expertise is likely to influence the success of curricular reform adoption and implementation.

With respect to teacher expertise, recent literature has broadened the stance on high quality teaching to include the effect of context on teaching and learning where previous conceptions of expertise, particularly in science, focused mainly on improving teacher's content and pedagogical content knowledge (Kennedy, 2010; National Research Council, 2011; National Science Board, 2010). This supports an expanded notion of expert teaching that includes not only teachers' human capital, but also how teachers navigate the complex and situated nature of teaching (Lin, Schwartz, & Hatano, 2005; Rozenszajn & Yarden, 2013) also known as adaptive expertise (Crawford et al., 2005; De Arment, Reed, & Wetzel, 2013; Hammerness et al., 2007).

While there is substantial theoretical literature on teachers' adaptive expertise, there is a need for empirical evidence illustrating specifically what adaptive expert teaching looks like and whether it promotes improved student-learning outcomes (Janssen et al., 2015; Soslau, 2012). In previous research, we conducted case studies of what adaptive expert teaching looks like when working with computer-supported complex systems curricula defining three essential characteristics: flexibility, deeper level understanding, and deliberate practice (Yoon et al., 2015b). The goal of this current study is to validate the proposed model of teacher adaptive expertise with a larger sample of 10 high school science teachers and to test whether the model is predictive of student learning of complex systems content with 351 students. If the model is robust, we should be able to see significant variation in teacher's adaptive expertise enactments. Thus, the first question guiding our research is, "Does the model differentiate between teachers' levels of adaptive expertise?" Likewise, if the hypothesis that teachers, who have higher levels of adaptive expertise, are better able to adopt and implement reforms, is correct, we should be able to see improved student learning in the classrooms of those teachers. Thus, the second question guiding our research is, "Do levels of teachers' adaptive expertise predict levels of student learning in the target area of reform?" In the following section we briefly review literature on teachers' adaptive expertise and outline the conceptual framework we used to measure it.

Teachers' adaptive expertise for complex systems teaching and learning

High school science teachers are expected to have extensive knowledge of their subject matter, but research has explored how the complex context in which teaching occurs influences student learning in addition to subject matter knowledge (Penuel et al., 2011). From this perspective, pedagogical content knowledge is considered to extend far beyond a teacher's grasp of the subject or training in pedagogy and classroom management. Recent literature defines teachers' pedagogical content knowledge as "teachers' understanding of how to help a group of students understand specific subject matter while using multiple instructional strategies, representations and assessments...working within the contextual, cultural and social limitations in the learning environment" (Nilsson, 2014; p. 1795). Some researchers assert that there is not a single definition of expert teaching, and that the value of teachers' knowledge is context dependent (e.g., Van Driel & Berry, 2012). However, in order to operationalize teaching expertise, we need a way to evaluate expert practice when we see it.

To investigate this, we turn to the study of expertise in the learning sciences (e.g., Bereiter & Scardamalia, 1993) and teaching (e.g., Berliner, 2001). In our previous research, we outlined what expertise might look like and how it can be developed with new or reform-oriented curricula and instruction (Yoon et al., 2015b) carefully examining the practice of three teachers. We constructed a model of important characteristics of teacher adaptive expertise based on literature briefly reviewed below.

How performance on novel problems can be enhanced has been characterized as adaptive expertise (Barnett & Koslowski, 2002; Hatano, 1982; Scardamalia & Bereiter, 1993). The complex nature of teaching requires teachers to be able to orchestrate myriad variables, see multiple perspectives, and identify problems and possibilities in existing and emergent situations (Bransford et al., 2005; Fairbanks et al., 2010). As adaptive experts, teachers deal with non-routine events, and seek to extend their capabilities always working at the edge of their competence. In this way, as Bereiter and Scardamalia (1993) suggest, expertise should be understood as a *process* rather than a *state*, which is by nature action oriented and therefore can be observed as it is enacted.

A review of the literature reveals three important characteristics with respect to actions associated with expertise, i.e., flexibility, ability to demonstrate deeper level understanding, and deliberate practice. Flexibility is characterized by the ability to opportunistically plan, change enactments faster than non-experts, and flexibly and critically apply their knowledge to new situations while constantly learning (Berliner, 2001; Bransford et al., 1999; Chi, 2011; Ferrari, 2002). Flexibility is also manifested in teachers' abilities to integrate aspects of teacher knowledge in relation to the teaching act while responding to their specific contexts (Tsui, 2009). Deeper level understanding is associated with one's ability to recognize meaningful patterns quickly, allowing one to attend to deeper level problem solving and in turn perform at a higher level (Berliner, 2001; Bransford, 1999; Ericsson et al., 2006; Ferrari, 2002; Hammerness et al., 2005; Levy & Murnane, 2004). Deliberate practice is manifested in one's ability to engage in reflection and conscious deliberation. Experts who demonstrate deliberate practice are highly motivated, self-regulated, and constantly seek to improve performance by identifying problems, addressing them, and finding new problems to work on (Bereiter & Scardamalia 1993; Berliner, 2001; Tsui, 2009).

In Yoon et al. (2015b), we found that the three teachers varied in these characteristics and that the variations appeared to influence the success of the implementation of project activities based in teaching and learning of complex systems. In this study, we aim to validate the model with a larger sample of teachers. We see this as an efficacy study testing the model with a different population, where the previous pilot work would be categorized as design and development research using the *Common Guidelines for Education Research and Development* (IES/NSF, 2013). Such progressions of research are necessary to ensure high quality evidence, meaningful findings, and actionable results through the systematic development of knowledge (IES/NSF, 2013). Before moving onto the study, we briefly describe complex systems and instruction that is best supported through computer simulations (Yoon et al., 2015a).

Complex systems are found in myriad natural and social phenomena. While definitions of complex systems vary based on discipline, central features include multiple interacting variables that form patterns discernable over time fueled by dynamic processes (Yoon, 2011). For example, ecological systems are made up of different trophic levels (e.g., producers and consumers) whereby through predator-prey interactions, energy flows through the system optimally existing in states of equilibrium (e.g., stable producer and consumer populations). System states fluctuate often by hidden variables (e.g., random mutations) at micro levels that are only observable over time at larger scales (a process known as emergence). Therefore, complex systems researchers use computer simulations to assist in modeling system evolution to investigate as well as to make predictions of future states (e.g., how drought patterns affect food webs).

Methods

Context

This study is part of a larger project aimed at building students' and teachers' understanding of biology content and instructional practices through exploring computational models of complex systems. For training on the intervention, teachers participated in a weeklong 30-hour summer PD workshop and follow-up Saturday workshops during the school year (approximately two workshops per semester totaling 10 hours) between 2012 and 2013. PD activities included hands-on training in five biology units on the topics of Genetics, Evolution, Ecology, the Human Body, and Animal Systems. The units entail working with the agent-based simulations that combine graphical blocks-based programming with a 3-D game-like interface. The units also include working through experiments that provide experiences in core scientific practices as outlined in the NGSS, such as analyzing and interpreting data, engaging in argument from evidence, and obtaining, evaluating and communicating information. There is no set sequence for the units; instead, teachers can implement the units in the order that suits their school curriculum. The curricular materials for each unit take 2 to 3 days to complete and include popular and academic literature about complex systems as well as short movies, PowerPoint presentations, and teacher and student activity guides.

PD activities also included training in complex systems structures and processes, such as feedback, interdependence, self-organization, and emergence, as well as activities that specified where the units fit into the high school biology curriculum. During the school year, teachers implemented the units in their biology classes.

Participants

We recruited 10 teachers—seven women and three men—from eight Boston-area public schools. The teachers came from a diverse set of schools. One school was as high as 75% ethnic/racial minorities, while another school was almost entirely White (3% minority). School-level percentages of low-income students ranged from approximately 14% to 80%. The percentage of students considered proficient or advanced on the state standardized science test ranged from 45 to 93. Teachers, on average, had 8.5 years of teaching experience, with a range of 3.5 to 19 years.

We collected student data in 10 classrooms drawn from 8 schools and a total of 351 students. The schools did not release individual student demographic and achievement data to us, so we cannot report accurate sample data in these areas. However, due to the range of classrooms, we believe that the students we worked with are a relatively representative sample of the population-level statistics that are reported. Table 1 provides further demographic details of the student population.

Table 1: School Level Demographic Data

School ID	School-Level Characteristics		
	% Low Income	% Above Proficient on State Assessment	% Minority*
1	36.0	71	NA
2	79.6	45	75.3
3	59.0	63	70.5
4	27.7	65	40.2
5	18.0	93	36.2
6	14.8	89	30.5
7	14.2	82	3.6
8	NA	78	25.5
Totals	35.6	73.3%	40.3

*Minority defined as non-white

Data sources

To see how teachers varied in their enactments of teachers' adaptive expertise, we analyzed classroom observations, year-end teacher interviews, teacher reflections of each unit implementation, and researcher focus group interviews. For the classroom observations, teachers and students were observed at least twice during the school year by research staff. The observation protocol required attention to details about how teachers and students worked with the simulations, how teachers facilitated learning with the curricular materials, and how teachers scaffolded student understanding of complex systems. Teacher interviews and reflections probed teacher's ideas about challenges in implementation and how they overcame them with particular focus on context variables such as student abilities, and resource availability. At the end of the year, a focus group interview was

also held with the research staff who observed in teacher’s classrooms to capture more formal and anecdotal information about the content of their communication with teachers between observations, and general impressions of teacher’s relative abilities to teach using intervention activities.

To determine student understanding of complex systems, two open-ended questions were administered before and after the intervention that provided scenarios on the topics of ecology and evolution. The scenarios are presented below.

Ecology Scenario

Imagine a flock of geese arriving in a park in Boston, where geese haven’t lived before. Describe how the addition of these geese to the park affects the ecosystem over time. Consider both the living and non-living parts of the ecosystem.

Evolution Scenario

Some mosquitoes no longer die when DDT (a powerful chemical used to kill insects) is sprayed on them, while others do. Many years ago, DDT worked to kill almost all mosquitoes. How would biologists explain why some mosquitoes do not die anymore when DDT is sprayed on them?

Analysis

To analyze teacher’s adaptive expertise enactments, the interview transcripts and recorded reflections and classroom observations were mined and coded for levels of enactments based on the categorization manual found in Table 2. The manual was first constructed and validated for the previous study (Yoon et al., 2015b) by researchers on the project team (authors 1–3) with only low and high enactments specified. For this study, we modified the manual to include enactments that were determined to fall in the middle range to provide a greater range. Two researchers external to the project were trained on the manual with 20% of the data yielding an interrater reliability alpha score of .75. The full data set was then scored by the authors. Discrepant or uncertain enactments were negotiated and given a single code. The highest score a teacher could receive in each category was 3 with a total combined adaptive expertise score of up to 9. We ran an analysis of variance (ANOVA) with the complete data set to see whether there was sufficient variation between teachers.

Table 2: Categorization Manual for Teachers’ Adaptive Expertise

Category and Definition	Examples
<p>Flexibility</p> <ul style="list-style-type: none"> • Ability to incorporate project activities into their daily practice. • Awareness of their student population and their needs • Awareness of the school context • Ability to respond to unexpected issues that arise • Able to adapt their practice by being able to incorporate project expectations in their situated context. 	<p>High: Teacher ran the simulation on the smartboard from his computer as students worked on their own so he could point things out and they could follow along.</p> <p>Medium: Teacher instructs the students that when they get to [argumentation], they need to stop and talk with another group and they should turn off their monitors when they are ready to discuss.</p> <p>Low: Teacher informed us that 17 of the 30 students in this class were failing. He had also NOT prepped the students at all in terms of introducing them to the simulation environment beforehand. I don’t think he reviewed the evolution concepts beforehand either.</p>
<p>Deeper level of understanding</p> <ul style="list-style-type: none"> • Actions of the teacher that demonstrate their ability to go beyond what they are required to do with the project. • Are able to implement extensions or make connections that build or address deeper level of knowledge construction or problem solving. • Bring in variation from outside the present system of activity 	<p>High: Teacher introduces complex systems. Talks about agents that follow certain rules. Talks about intestines and about phospholipid bilayer.</p> <p>Medium: Teacher told the students who finished early to “mess with what you’ve done.” The students were able to create their own mini experiments by changing one of the variables in the model. This kept the group engaged while waiting for slower students. This also allowed the students space to explore and ask new questions.</p> <p>Low: From here on, the teacher is completely hands-off. It is clear some students are very confused and don’t know where to start.</p>

<p>Deliberate practice</p> <ul style="list-style-type: none"> • Actions of the teacher that demonstrate their ability to show motivation, focus, and repeated effort to monitor their practice, • Devise and subsequently attempt new approaches to improve implementation • Teachers exhibit explicit evidence of reflecting on a problem and how to improve. 	<p>High: After some readjustment from Day 1 between teacher A & teacher B, teacher B determined they could split the laptop cart and she could use laptops in the room – this was a big advantage.</p> <p>Medium: Teacher was surprised in the first class that the students didn't "fly through it" and make it through the whole activity. He said that he could have prepared a bit better</p> <p>Low: In debrief with the teacher: He did not prepare his students at all before implementing the evolution activity. He seemed to think it 'went OK' and again reiterated that this group of students was a tough group</p>
--	---

Student responses to the open-ended question were scored on a three-point scale from (1) clockwork thinking (non-complex) to (3) complex thinking in four categories: predictability, processes, order, and emergence and scale. Table 3 shows a shortened version of the categorization manual which has been validated in previous studies, e.g., Yoon et al., 2015a; Yoon et al., forthcoming). We ran a regression analysis to see whether teachers' adaptive expertise scores were predictive of students' complex systems scores.

Table 3: Categorization Manual for Student's Complex Systems Understanding

Complex Systems Components	Descriptions
Predictability	The emphasis is on the <u>predictability of the effects caused by the agent</u> in question. According to the clockwork framework, the way in which a part or agent operates or affects other components of the system is predictable. In a complex framework, it is impossible to anticipate precisely the behavior of the system. This is because the actions of agents cannot be predicted (as random forces or chance factors can affect an agent's actions) even if we know the rules or characteristics of the agent.
Processes	Processes refer to the dynamism of the mechanisms that underlie the phenomena (i.e., how the system works or is thought to work). In a clockwork framework, there is a <u>beginning, middle, and end</u> in the system. The system is composed of <u>static events</u> . While perturbations (actions by/on parts) in the system may cause change to occur, the change terminates once an outcome is achieved. In a complex systems framework, there is <u>no definite beginning and end</u> to the activity. System processes are <u>ongoing and dynamic</u> .
Order	The focus is the organization of the system or phenomenon as <u>centralized or decentralized</u> . A clockwork framework assumes that all systems are controlled by a central agent (e.g., all action is dictated by a leader). Order is established top-down or determined with a specific purpose in mind. In a complex systems framework, control is decentralized and distributed to multiple parts or agents. Order in the system is self-organized or 'bottom-up' and emerges spontaneously.
Emergence and Scale	Emergence refers to the phenomenon where the complex entity manifests properties that exceed the summed traits and capacities of individual components (Davis & Sumara, 2006). In other words, these <u>complex patterns simply emerge from the simpler, interdependent interactions</u> among the components (Capra, 1996). In a clockwork framework, parts of the system are perceived to be isolated with little interdependency among them. This is because of the linear nature that characterizes these relationships. Thus, there are no large, global patterns that emerge from actions imposed on the system. Rather, these actions cause only localized changes (e.g., geese eat plants causes a decrease in grass). In a complex system, because parts or agents are interdependent in multiple ways, an action (small or large) that is imposed on the system may have large and far-reaching consequences on the numerous parts and agents of the system. This may in turn result in large-scale change and evolution.

Findings

Teacher adaptive expertise combined scores ranged from 3.67 to 7.65 out of a possible score of 9 (see Table 4). The mean adaptive expertise was 6.18 with a standard deviation of 1.01. To determine whether there were meaningful differences in adaptive expertise between teachers, we conducted an ANOVA of the teachers' adaptive expertise item scorings (See Table 5). From these results ($F(9, 362) = 15.07, p < 0.001$), we conclude that the teachers in our sample differed significantly in their adaptive expertise.

Table 4: Adaptive Expertise Scores by Category for Each Teacher

Teacher	Flexibility	Deeper Understanding	Deliberate Practice	Overall Expertise
1	2.25	2.17	2.00	6.42
2	1.15	1.18	1.33	3.67
3	2.65	2.59	2.42	7.65
4	2.06	1.79	2.20	6.04
5	1.67	1.43	2.43	5.53
6	2.06	1.79	2.36	6.21
7	2.50	2.13	2.00	6.63
8	2.21	2.18	2.62	7.02
9	1.92	1.10	2.67	5.69
10	2.22	2.04	2.67	6.93

Table 5: Analysis of Variance of Teacher Adaptive Expertise Item Scorings

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Teacher	9	60.53	6.73	15.07	<0.001
Residuals	362	143.30	0.40		

To test whether teachers' adaptive expertise scores were predictive of student understanding of complex systems, regressions were run holding student pre-scores constant. The first regression examined student complex system understanding in ecology and the second regression examined student complex system understanding in evolution. Table 6 shows results for ecology ($F(2, 350) = 22.13, p < 0.001$). For each one-unit increase in teacher combined AE rating, there was an increase of 0.94 points in student performance on the ecology question. Table 7 shows the results for evolution ($F(2, 295) = 28.64, p < 0.001$). Here, student scores were 1.24 points higher for each one-unit increase in their teacher's combined AE rating. Both of these analyses led us to the same conclusion: adaptive expertise is significantly related to student achievement in complex systems.

Table 6: Regression Assessing the Relationship Between Student Complex Systems Understanding of Ecology and Teacher Adaptive Expertise

	Estimate	Std. Error	t-value	p-value
(Intercept)	3.09	0.57	5.39	<0.001
Overall Combined AE Rating	0.94	0.27	3.46	<0.001
Pre-test	0.25	0.05	4.78	<0.001

F-statistic: 22.13 on 2 and 350 degrees of freedom (p-value <0.001)

Table 7: Regression Assessing the Relationship Between Student Complex Systems Understanding of Evolution and Teacher Adaptive Expertise

	Estimate	Std. Error	t-value	p-value
(Intercept)	1.55	0.75	2.07	0.039
Overall Combined AE Rating	1.24	0.36	3.44	<0.001
Pre-test	0.39	0.07	5.89	<0.001

F-statistic: 28.64 on 2 and 295 degrees of freedom (p-value <0.001)

Conclusions and implications

Our findings provide a model of the qualities important in defining teachers' adaptive expertise and what it looks like in the classroom with computer supported complex systems curricula. Furthermore, as we have found the model to demonstrate sufficient variation among teachers and to be predictive of student learning, this study provides important empirical evidence to the research on teachers' adaptive expertise (Janssen et al., 2015; Soslau,

2012). We believe that using an adaptive expertise model helps professional developers and researchers interested in learning how to train teachers to teach with complex systems resources and approaches by illustrating the range of contextualized classroom enactments. As Fairbanks et al. (2010) and other researchers suggest, the complex nature of teaching requires the orchestration of many variables and responses to emergent situations that often cannot be predicted a priori.

We are aware of the limitations of the study in that the sample was small and self-selected therefore we can make no generalizability of claims. We do believe, however, that this efficacy research provides evidence of the support of such an adaptive expertise model for teachers in line with developing systematic knowledge that can eventually be scaled (IES/NSF, 2013). Furthermore, as Wilson (2013) notes, it is rare to be able to link learning directly to teacher practices. These findings provide promising evidence in support of a larger scale effort to develop teachers' adaptive expertise to ensure adoption and high quality implementation of curriculum and instruction based in the NGSS.

References

- Yoon, S., Anderson, E., Koehler-Yom, Evans, C., Park, M., J., Sheldon, J., Schoenfeld, I., Wendel, D., Scheintaub, H., Klopfer, E. (forthcoming). Teaching about complex systems is not simple matter: Building effective professional development for computer-supported complex systems instruction. *Instructional Science*.
- Yoon, S., Anderson, E., Koehler-Yom, J., Sheldon, J., Schoenfeld, I., Wendel, D., Scheintaub, H., Klopfer, E., Oztok, M., & Evans, C. (2015a). Design features for computer-supported complex systems learning and teaching in high school science classrooms. *Journal of Research in STEM Education*, 1(1), 17-30.
- Yoon, S., Koehler-Yom, J., Anderson, E., Lin, J., & Klopfer, E. (2015b). Using an adaptive expertise lens to understand the quality of teachers' classroom implementation of computer-supported complex systems curricula in high school science. *Research in Science and Technology Education*, 33(2), 237-251.
- Yoon, S. A. (2011). Using social network graphs as visualization tools to influence peer selection decision-making strategies to access information about complex socioscientific issues. *Journal of the Learning Sciences*, 20(4), 549-588.
- Blandford, S. (2012). *Managing professional development in schools*. New York: Routledge.
- Bransford, J.D., Brown, A.L., & Cocking, R.R. (1999). How experts differ from novices. In Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.). *How People Learn: Brain, Mind, Experience, and School* (pp. 19-38). Washington DC: National Academy Press.
- Bransford, J., Darling-Hammond, L., & LePage, P. (2005). Introduction. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing Teachers for a Changing World: What Teachers Should Learn and be Able to do* (pp. 1-39). San Francisco: Jossey-Bass.
- Barnett, S. M. & Koslowski, B. (2002). Adaptive expertise: effects of type of experience and the level of theoretical understanding it generates. *Thinking & Reasoning*, 8(4), 237-267.
- Bereiter, C., & Scardamalia, M. (1993). *Surpassing ourselves: An inquiry into the nature and implications of expertise*. Illinois: Open Court.
- Berliner, D. C. (2001). Learning about and from expert teachers. *International Journal of Educational Research*, 35, 463-482.
- Calabrese Barton, A. (2007). Science learning's in urban settings. In Abell, A. K., & Lederman, N. (Eds.), *Handbook of Research on Science Education* (319-343). Mahwah, NJ: Lawrence Earlbaum Associates, Publishers.
- Crawford, V. M., Schlager, M., Toyama, Y., Riel, M., & Vahey, P. (2005). Characterizing adaptive expertise in science teaching. Paper presented at the American Educational Research Association Annual Conference, April 11–15, 2005, Montreal, Canada.
- Chi, T. (2011). Theoretical perspectives, methodological approaches, and trends in the study of expertise. In Y. Li and G. Kaiser (Eds.), *Expertise in mathematics instruction: An international perspective* (pp. 19-39). New York: Springer.
- Davis, B., & Sumara, D. J. (2006). *Complexity and education: Inquiries into learning, teaching, and research*. Psychology Press: Mahwah, NJ
- De Arment, S. T., Reed, E., & Wetzel, A. P. (2013a). ePortfolios: Promoting special education adaptive expertise through reflection in a web-based learning community. *International Journal of ePortfolio* 3(2), 149-160.
- De Arment, S. T., Reed, E., & Wetzel, A. P. (2013b). Promoting adaptive expertise: a conceptual framework for special educator preparation. *Teacher Education and Special Education: The Journal of the Teacher*

- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher* 38(3), 181-199.
- Ericsson, K.A., Charness, N., Feltovich, P.J., Hoffman, R.R. (2006). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.
- Fairbanks, C.M., Duffy, G.G., Faircloth, B.S., He, Y., Levin, B. et al. (2010). Beyond knowledge: Exploring why some teachers are more thoughtfully adaptive than others. *Journal of Teacher Education*, 61(1-2), 161-171.
- Ferrari, M. (2002). *The pursuit of excellence through education*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Corchan-Smith, M., McDonald, M., & Zeichner, K. (2007). How teachers learn and develop. In Darling-Hammond, L., & Bransford, J. (Eds.), *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do* (358-389). Hoboken, NJ: Jossey-Bass
- Hatano, G. (1982). Cognitive consequences of practice in culture specific procedural skills. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, 4(1), 15-18.
- Ho, A. D., Chuang, I., Reich, J., Coleman, C., Whitehill, J., Northcutt, C., Williams, J. J., Hansen, J., Lopez, G., & Petersen, R. (2015). HarvardX and MITx: Two years of open online courses. *HarvardX Working Paper No. 10*.
- IES/NSF (2013). *Common guidelines for education research and development: A report from the Institute of Education Sciences, U.S. Department of Education and the National Science Foundation*. Authors: Washington D.C.
- Janssen, F., Grossman, P., & Westbroek, H. (2015). Facilitating decomposition and recomposition in practice-based teacher education: The power of modularity. *Teaching and Teacher Education*, 51, 137-146.
- Kennedy, M. M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher*, 39(8), 591-598.
- Levy, F., & Murnane, R. (2004). *Expert thinking in the new division of labor: How computers are creating the next job market*. NY: Princeton University Press.
- Lin, X., Schwartz, D.L., & Hatano, G. (2005) Toward teachers' adaptive metacognition, *Educational Psychologist*, 40(4), 245-255.
- Lin, X., Schwartz, D., & Bransford, J. (2007). Intercultural adaptive expertise: Explicit and implicit lessons from Dr. Hatano. *Human Development*, 50, 65-72.
- National Research Council. (2011). *Successful K–12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington, DC: The National Academies Press.
- National Science Board. (2010). *Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital*. Washington, DC: National Science Foundation.
- Nilsson, P. (2014). When teaching makes a difference: Developing science teachers' pedagogical content knowledge through learning study. *International Journal of Science Education*, 36(11), 1794-1814.
- Penuel, W.R., Fishman, B.J., Haugan Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, 40(7), 331-337.
- Reiser, B., J. (2013). What professional development strategies are needed for successful implementation of the next generation science standards? *Invitational Research Symposium on Science Assessment*.
- Rozenszajn, R. & Yarden, A. (2013). Expansion of biology teachers' pedagogical content knowledge (PCK) during a long-term professional development program. *Research in Science Education*, 44, 189-213.
- Soslau, E. (2012). Opportunities to develop adaptive teaching expertise during supervisory conferences. *Teaching and Teacher Education*, 28, 768-779.
- Tsui A.B.M. (2009). Distinctive qualities of expert teachers. *Teachers and Teaching: Theory and Practice*, 15(4), 421-439.
- Van Driel, J.H., & Berry, A. (2012). Teacher professional development focusing on pedagogical content knowledge. *Educational Researcher*, 41(1), 26-28.
- Wilson, S. M. (2013). Professional development for science teachers. *Science*, 340, 310-313.

Acknowledgments

We thank Eric Klopfer, Josh Sheldon, Ilana Schoenfeld, Daniel Wendel, and Hal Scheintaub for their collaboration on the project. This work was funded by the U.S. National Science Foundation Discovery Research K–12 (DRL 1019228).

What Do Learning Scientists Do? A Survey of the ISLS Membership

Susan A. Yoon, University of Pennsylvania, yoonsa@upenn.edu
Cindy Hmelo-Silver, Indiana University, chmelosi@indiana.edu

Abstract: This study responds to an often asked question, what do learning scientists do? Earlier attempts to answer this question came from a need to define a new field of educational research. However, as we have grown into a robust and highly productive society, self-referential methods for describing the membership is prudent to gain a more accurate understanding of where, for whom, and what kinds of learning sciences research and practices take place. Here we report on the responses of 253 ISLS members from a survey conducted in 2014. We discuss implications of the findings in terms of the types of impact learning scientists have, how we might use these results for advising prospective students, and ways that we might use this to create our future.

Keywords: learning sciences, community, members, research fields

Introduction

What do learning scientists do? Earlier attempts to describe the learning sciences (e.g., Kolodner, 2004; Nathan, Rummel, & Hay, forthcoming; Sawyer, 2014) have come from a need to distinguish the field from other areas of educational research. Such characterizations have served important roles for defining a more nascent field. However, as we have matured, it becomes important to define ourselves in terms of what we do rather than in terms of differences with other fields. Evidence of this maturity is found in ISLS having two high impact flagship journals, two robust international conferences, a quarter century of scholarship since the launch of the *Journal of the Learning Sciences* (JLS), and learning scientists connected in large numbers across the globe. It now seems prudent to let the learning science membership define what we do as an educational field. Such self-referential activities are common in social systems. Following Luhmann (1990) we apply the term autopoiesis as a rationale for understanding our field in this way. Autopoietic systems are, “networks of productions of components that recursively, through their interactions, generate and realize the network that produces them and constitute, in the space in which they exist, the boundaries of the network as components that participate in the realization of the network” (p. 3). Taking stock of how organizations and fields evolve and self-organize provide lenses for which to understand centers of gravity and boundaries, while at the same time providing insights into limitations and possible directions for growth. We argue that now is that time to take stock. In this report, we present the results from a membership survey aimed at understanding the breadth of research, practices, methods, and occupations, articulated by members of the *International Society of the Learning Sciences* (ISLS) as of 2014. We believe that the results will have great utility in a number of institutional and professional activities including recruitment of graduate students, degree-bearing program development, and identification of potential areas of research interest, gaps, and growth.

Earlier efforts to define the Learning Sciences

Kolodner (2004) notes that in establishing *JLS*, she wanted to create a venue for cognitive scientists of all kinds to publish research about learning in real world settings and about big ideas in learning. The inception of *JLS* signaled the arrival of a new and growing field of educational research. Over the years, the learning sciences (LS) field has moved from being largely cognitive science focused to one that more fully integrates social and cultural theories. The Learning Sciences, while still drawing on other methodologies, has also developed design-based research as its signature research methodology (e.g., Bielaczyc, 2006; Brown, 1992). The field in general has defined itself with a focus on learning in complex domains with an interest in the dialectic between theory and design—using theory to guide design and enactments of those designs to guide theory (Bielaczyc, 2006).

From the proceedings of a workshop aimed at defining the learning sciences and the implications of such a definition for graduate education, Nathan, Rummel, & Hay (forthcoming) open with four themes they considered to be central contributors to the development of the learning sciences field. These included (1) the design of learning environments and practices, (2) use-inspired basic research, (3) using authentic practices and settings to test hypotheses, and (4) an engineering ethos that envisions new practices and resources to support learning. Delving into the specifics of the themes, among the interesting findings that emerged from workshop discussions were challenges in what was termed “branding” the learning sciences along disciplinary and methodological lines that has implications for the hiring and tenure processes. Another challenge that emerged was the difficulty in

explaining to prospective students what the career paths are for those with LS degrees. We take up some of the challenges raised in that report in the present study.

Recently, in the introduction to the 2nd edition of the *Handbook of the Learning Sciences*, Sawyer (2014) summarizes various research foci that have historically characterized learning sciences research such as the importance of anchoring learning events in prior knowledge, the role of expert knowledge, learning through social interaction, the role of design and application in scaffolding levels of understanding, and technological supports for building knowledge. Teaching and learning in schools is also heavily emphasized. However, quite apart from what is described as “The New Science of Learning,” a number of these strands have evolved beyond their historical roots with nuanced approaches, designs, and environments that blur the boundaries of where, for whom, and what kinds of LS research and practices take place. Here, surveying the membership can provide a more accurate picture of what the field does and how we conceptualize our own LS identities. Next we describe the methods and findings of the ISLS membership survey.

Methods

The survey was commissioned by the president of ISLS in the year 2014 (second author) and charged to the chair of the ISLS Membership Committee (first author) to conduct. The process of constructing the survey was collaborative and iterative. First, the president and chair met to discuss potential questions. Next, a small team of Membership Committee affiliates worked to provide details and refine the survey. Categories of research and scholarship were taken from existing application protocols for ISLS membership as well as conference proposal taxonomies for reviewing. A final draft was sent to the full Membership Committee and the ISLS Board of Directors to vet and revise. In January 2014, the draft survey in Survey Monkey was piloted with approximately 20 learning scientists who worked in various known LS-oriented organizations, e.g., research scientists at SRI, university professors, graduate students, and post-doctoral fellows. Responses in the pilot phase were reviewed by the small team of Membership Committee affiliates (see acknowledgements), revised, and then sent out for a final review with the full Membership Committee. The final survey was sent out through the ISLS listserv on February 5, 2014 with multiple periodic email reminders over the course of 2 months. In total, 253 people responded to the survey out of a possible 450 members, which constitutes 56% of the membership.

Survey questions

The survey was comprised of 32 multiple choice, fill in the blank, short answer, and drop down menu questions. Members responded to only the subsection of questions that pertained to the kind of job they did, e.g., professor, student, research scientist. Table 1 provides a modified version of the survey questions completed by professors.

Table 1: ISLS Membership Survey 2014 questions for professors

1. Do you consider yourself a learning scientist?	10. Do you advise students who are doing LS research?
2. Do you see the work that you do as related to LS?	11. Of the masters students in the last 5 years, list their current occupations.
3. Please select the category that best fits your career status.	12. Of the doctoral students in the last 5 years, list their current occupations.
4. What is your rank or job title?	13. On average, how many students do you advise who are doing LS research each year?
5. What is the name of the department in which you work?	14. Please indicate your areas of [domain] interests (selected from 24 choices, e.g., Computer Science, Learning Technologies, Psychology).
6. What is the name of the department where you obtained your doctoral degree?	15. Please indicate the primary methodological approach (e.g., quantitative, qualitative, mixed methods)
7. Did you receive your degree from an LS program?	16. Please indicate your research focus (selected from 39 choices including “other”, e.g., Assessment, Gender, Scaffolding).
8. What were your research interests while you pursued your degree?	17. Please indicate the contexts of your work (e.g., informal learning settings).
9. What are your current research interests?	18. Please indicate the main population of your research.

Analysis

For multiple choice, fill in the blank, and drop down menu questions, frequencies of responses were tallied and reported. For short answer questions, responses were qualitatively mined to find general population categories

and trends. Category descriptions were negotiated by the small team of Membership Committee affiliates.

Findings

Table 2 summarizes key findings for the global population surveyed in terms of occupation, rank, degree and advising. This is followed by more detailed findings for questions about masters and doctoral student occupations, domain interests, methodologies, research foci, contexts, and populations.

Table 2: Combined population survey responses for occupation, rank, degree, and advising

Question	Response	Frequencies	Percentage
Combined Population		253	
Considers himself/herself learning scientist	Yes	230	90.9
	No	23	9.1
Current job status	Work in academia (e.g., professor, lecturer, post-doc)	160	63.2
	Professional staff working in academia (e.g., researcher, evaluator, project manager)	19	7.5
	Professional staff working outside of academia (e.g., SRI, Google, a museum)	27	10.7
	Doctoral student	47	18.6
	Master's student	0	0
Academics (e.g., professor, lecturer, post-doc)		160	
Current Rank or title	Full Professor	40	25
	Associate Professor	39	24.4
	Assistant Professor	33	10.6
	Non-Tenured Associate Professor	3	1.9
	Lecturer	7	4.4
	Post-Doctoral Researcher	13	8.1
	Other/Unspecified	25	15.6
Received degree from a learning sciences program	Yes	55	34.4
	No	105	65.6
Advise students doing learning sciences research	Yes	109	68.1
	No	51	31.9
Average number of students advised per year	Masters	Average: 3	
	Doctoral	Average: 3	
Professional staff in or outside academia (e.g., researcher, museum)		46	

As expected, a majority of respondents (63.2%) reported working in academia as professors, lecturers, post-docs, or doctoral students. However, we were surprised to see that nearly one-fifth of the members who responded worked as professional staff in and out of the academy. Of those who said that they worked in academia, not surprisingly, only 34.4% of the respondents said that they received their degree from a learning sciences program, however, a majority of them said that they advised students who are doing learning sciences research. This has implications for the specificity and need for learning sciences programs offered.

To obtain an understanding of what learning sciences masters students do, the specific wording of the question was, "Of the MASTERS students whom you directly advised in the last five years, please list their current occupations to the best of your knowledge." Similar wording was used for doctoral student occupations. Figures 1 and 2 shows the distribution of responses. For masters students, occupations included data analyst, game designer, financial analyst, banker, human resources development, and restaurateur. For doctoral students, interesting occupations included running a start-up company in digital media, leader of an educational NGO, instructional tech support in higher education, and assistant secretary of education. Of particular interest is that the largest number of former masters students are K-12 teachers.

Figure 3 shows what the membership revealed about their domain interests as delineated in four categories of *Academics*, *Doctoral Students*, *Professional Staff in Academia*, and *Professional Staff Outside of Academia*. Not surprisingly, the categories of *Learning Environments Design*, and *Learning Technologies* were mentioned the most followed by higher frequencies in the categories of *Curriculum*, *Social Sciences*, and *Professional Development*. However, there were nearly 40 responses in the *Other* category, which included Dentistry, Philosophy, Architecture, Acupuncture, Measurement, and Civic Engagement. Similarly, Figure 4 shows the details about the membership's research foci. Again not surprisingly, we see high frequencies in the categories of *Collaborative knowledge building*, *Inquiry learning*, *Communities of practice*, and *Computer*

supported collaborative learning. The *Other* category included open education, social computing and epistemic feedback.

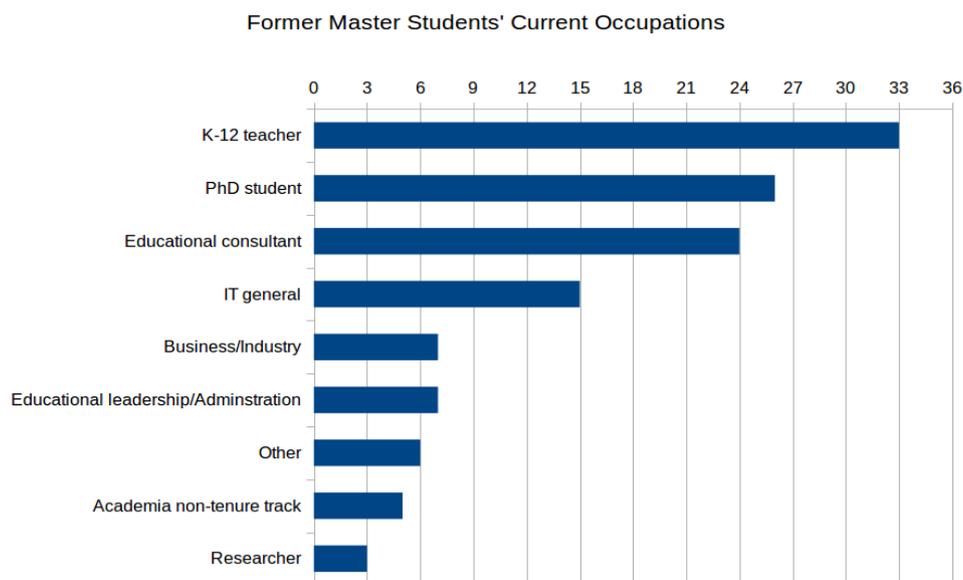


Figure 1. Masters Students' Current Occupations.

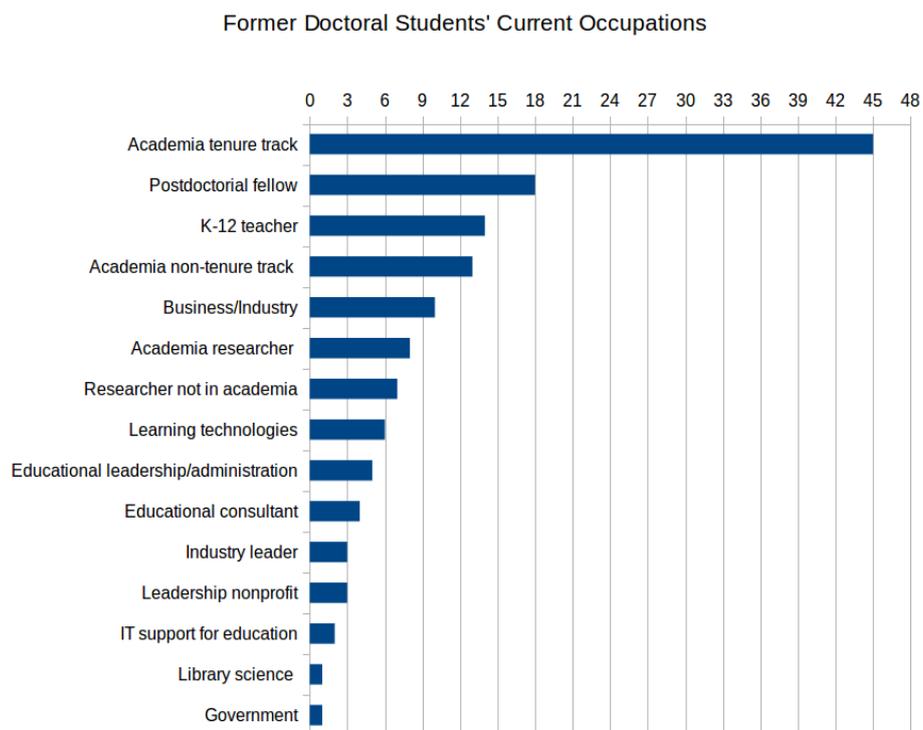


Figure 2. Doctoral Students' Current Occupations.

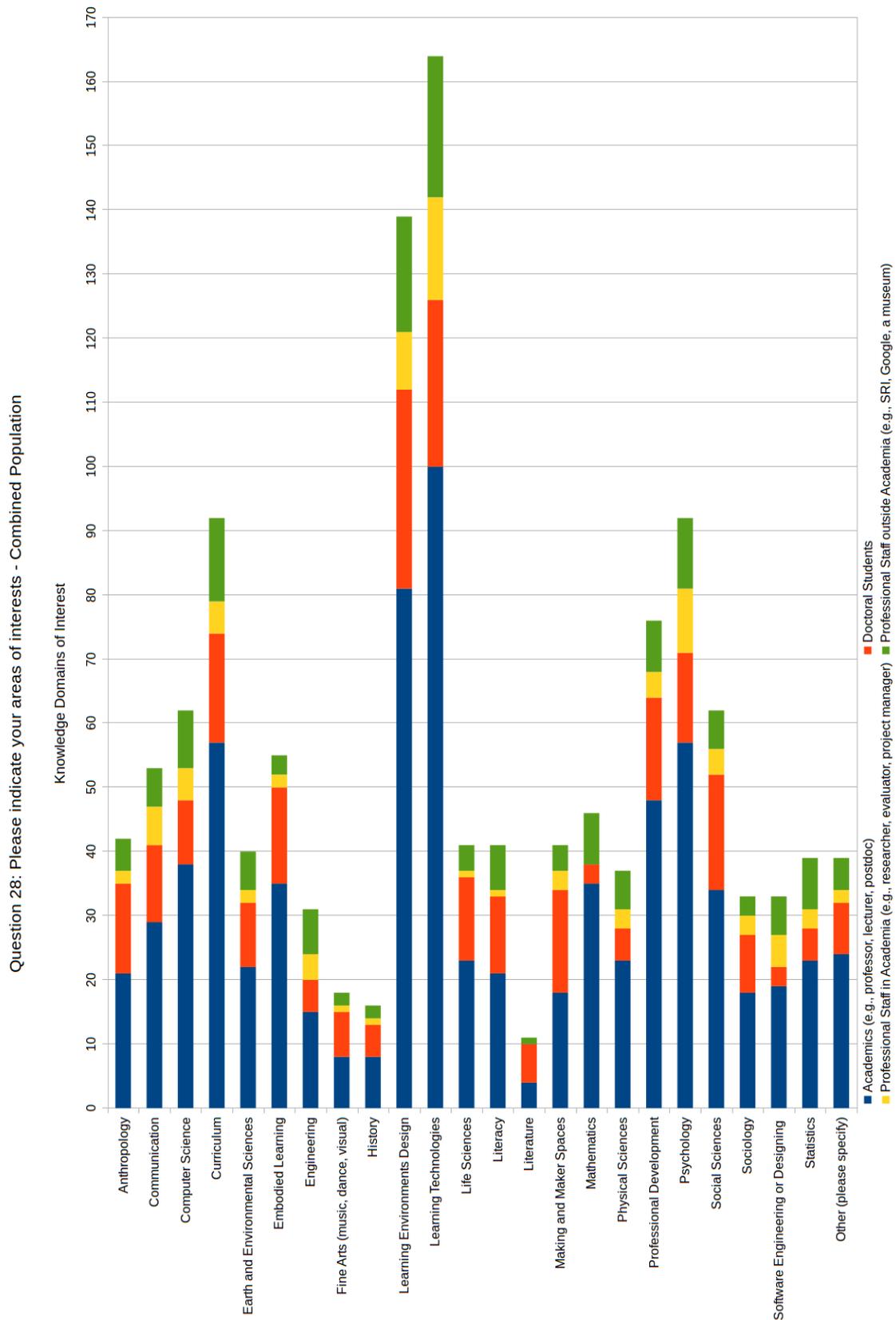


Figure 3. Combined Domain Interests.

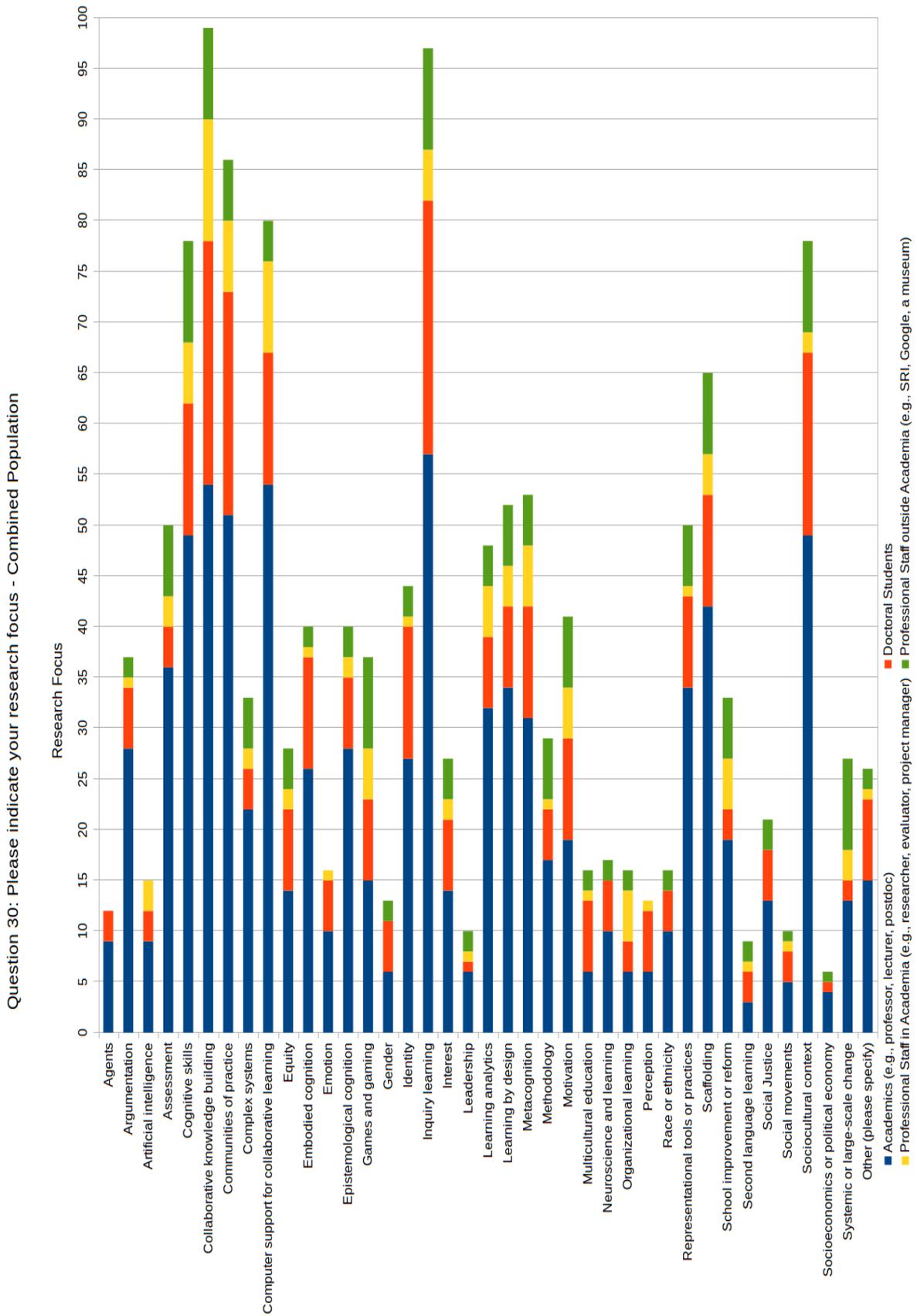


Figure 4. Research Foci.

With respect to the question on the primary methodological approach, about 53.8% of respondents said that they used mixed methods, and 23.3% said that they used qualitative methods. Just 3.6% said that they used quantitative methods only and of those there were no doctoral students represented. Another 1% said that they used other methods and about 10% did not respond.

About half of the membership surveyed (49.4%) said that they work in formal learning settings. Working only in informal settings was indicated by 7.1% of the membership followed by 28.4% working in both formal and informal, and 4.3% working in other learning settings that included medical education and training, life long learning, technical education, work sites, and everyday settings.

Finally, Figure 5 shows the distribution of populations survey respondents said they worked with.

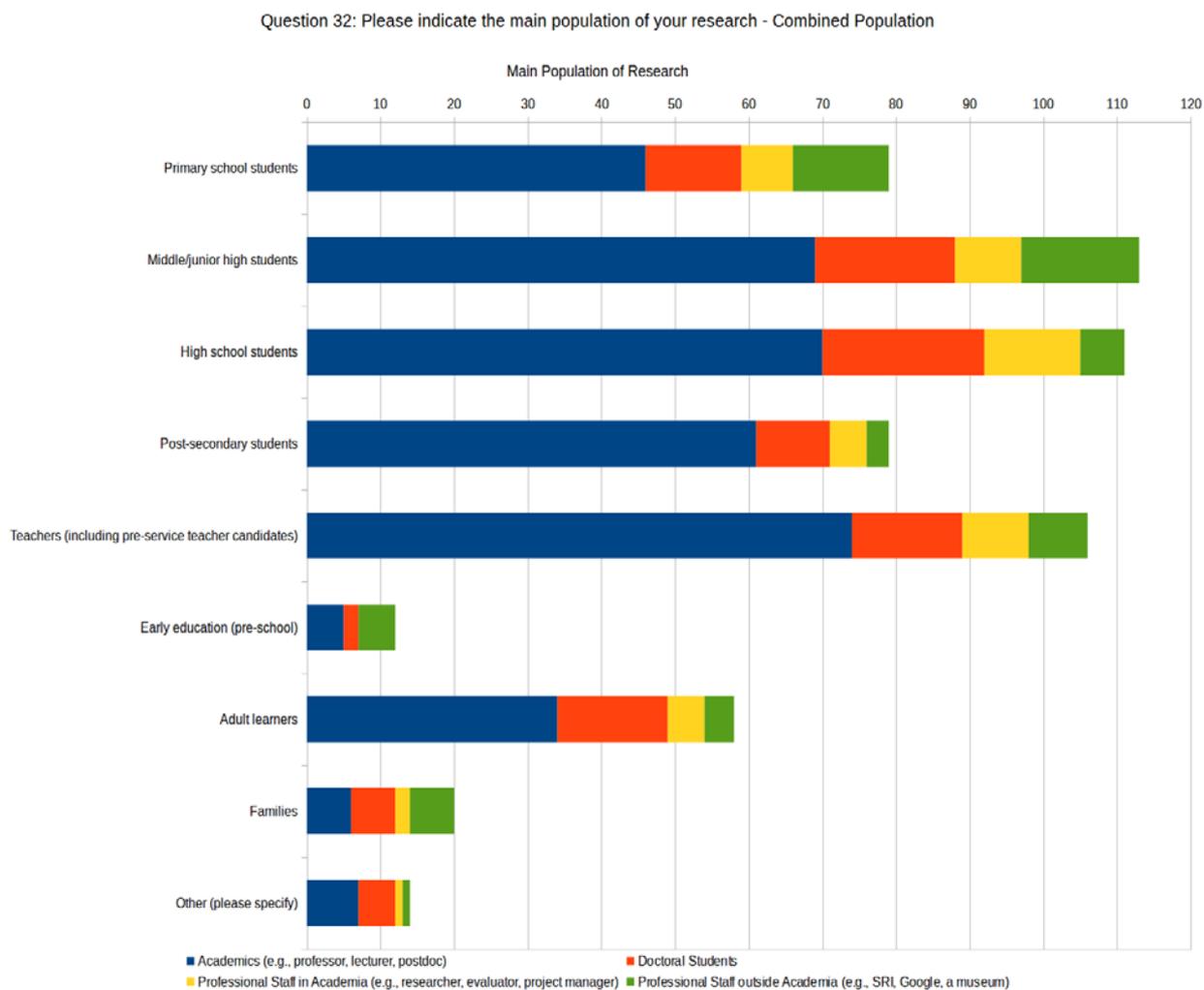


Figure 5. Main Population of Research.

We can see from the graph that a large portion of survey respondents work with students. However, we were surprised that nearly half said that they worked with teachers. We were also surprised at the numbers of people working with adult learners (23 percent) and with families (8 percent). Responses in the *Other* category included physicians, nurses, health professionals, disabled persons, and front-line youth workers. This suggests that other than early education, learning scientists work with learners across much of the life span.

Conclusions and implications

Acknowledging the autopoietic nature of the LS, the survey aimed at investigating how the LS field has self-organized. Our first answer to the question “What do learning scientists do” is that we do quite a bit. But the results of this survey provide suggestions that are more nuanced than that in terms of how LS has grown as a field, the types of impact learning scientists have, how we might use these results for advising prospective students, and ways that we might use this to create our future. The first point of note is that while many of the faculty do not have learning sciences degrees, the students they advise are in learning sciences programs—this is consistent with the growth of learning sciences programs internationally (e.g., see NAPLeS at <http://isls-naples.psy.lmu.de/>).

A second point of note is the diverse settings in which learning scientists work—that provides both evidence of impact and information that we can use to advise prospective masters and doctoral students. For masters graduates, K-12 schools are a frequent place of employment. This is particularly interesting because this is one mechanism for learning sciences inspired theories and designs to have an impact on practice. It is also clear that another place where learning scientists are influential and provide possible career paths are in informal settings, disciplinary education (e.g., medical, dental education), and in the private sector such as game design.

The third point answers directly the question posed in the title of the paper. Many learning scientists take a traditional route through academia, but at both the masters and doctoral levels, the survey results show that there are other pathways as well – as consultants, in information technology fields, in non-profit and research organizations, for example. Most learning scientists, not surprisingly have interests in learning environments design and learning technologies but beyond that, they work in a range of disciplinary fields. Learning scientists are also involved in professional development and curriculum, which goes along with developing learning environments and technologies, which after all, are embedded in curriculum and the need for professional development support. One question that is not addressed here is the degree to which teachers are engaged in participatory design. The survey results also show that a substantial number of learning scientists are working in emerging areas such as embodied learning and makerspaces. In addition, it shows the range of research methods used: most LS researchers are using mixed methods rather than purely quantitative or qualitative methods. Together these findings are particularly important as LS programs market to potential students, make decisions about faculty hires, or advise current graduate students.

Finally, it is important to think about how the results of this survey can help the learning sciences create our future in line with autopoietic systems. In ISLS Board meetings, one question that has arisen is whether we need special interest groups for various sub-communities. While this survey cannot answer this question, it suggests some possible areas such as informal learning and other outside academia contexts. It might also suggest ways that the society can help with brokering, or at least making visible, different forms of expertise so that new and productive collaborations might form. To support continual growth of the society, connections can also be made between organizations that overlap with membership research and interests such as the Association for Educational Communications and Technology (AECT), Learning Analytics & Knowledge (LAK), the National Association for Research in Science Teaching (NARST), the European Association for Learning and Instruction (EARLI), and Computer-Supported Cooperative Work (CSCW). We hope that these results will help current and future learning scientists envision a range of possible futures.

References

- Bielaczyc, K. (2006). Designing social infrastructure: Critical issues in creating learning environments with technology. *Journal of the Learning Sciences*, 15, 301–329.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions. *Journal of the Learning Sciences*, 2, 141–178.
- Kolodner, J. (2004). The learning sciences: Past, present and future. *Educational Technology*, 44(3), 37–42.
- Luhmann, N. (1990). *Essays on self-reference*. New York: Columbia University Press.
- Nathan, M. J., Rummel, N., & Hay, K. E. (forthcoming). Growing the Learning Sciences: Brand or big tent? Implications for graduate education. In M. A. Evans, M. J. Packer, and R. K. Sawyer (Eds.) *Reflections on the Learning Sciences*. (pp. xx-xx). New York: Cambridge University Press.
- Sawyer, K. (2014). Introduction: The new science of learning. In K. Sawyer (Ed.) *Cambridge handbook of the learning sciences*. New York: Cambridge University Press.

Acknowledgments

The authors would like to thank Murat Oztok, Emma Anderson, and Jessica Koehler-Yom for their work on this study.

The Interactional Work of Configuring a Mathematical Object in a Technology-Enabled Embodied Learning Environment

Virginia J. Flood, University of California, Berkeley, flood@berkeley.edu
Benedikt W. Harrer, San José State University, benedikt.harrer@sjsu.edu
Dor Abrahamson, University of California, Berkeley, dor@berkeley.edu

Abstract: We present a detailed account of interactional mechanisms that support participation in STEM disciplinary practices as an adult and a child explore a technology-enabled embodied learning environment for mathematics. Drawing on ethnomethodological studies of technology-rich workplaces, we trace the process of transforming a vague reference into a mutually available mathematical object: a covariant variable. Our analysis reveals that this mathematical object is an interactional achievement, configured via a *reciprocal* process of instructing one another's attention. In particular, we demonstrate how participants' explicit responsiveness to indexical and multimodal resources achieves this object.

Keywords: objectification, collaborative imagining, ethnomethodology–conversation analysis, multimodality, technology-enabled embodied learning environments

Introduction

Discovery-based learning environment designs are increasingly enlisting new user interface technologies (e.g., Microsoft Kinect) to incorporate the body into computer-mediated, collaborative explorations of STEM phenomena. Technology-enabled embodied learning environments (TEELEs) engage groups of learners in a wide range of topics, from models of meteor orbit trajectories (Lindgren & Moshell, 2011) and atomic interactions (Enyedy et al., 2012) to polar bear energy expenditure (Lyons et al., 2012), and even mathematical objects (Nemirovsky et al., 2012). As these technology-rich environments for inquiry gain popularity in classrooms and informal learning centers, there is a growing need to investigate how social interactions in TEELEs can support or constrain children's opportunities for participation in STEM disciplinary practices.

In this paper, we contribute a close examination of a productive sequence of interaction between a child and an adult tutor as they use a TEELE called the Mathematical Imagery Trainer for Proportions (MIT-P). Inspired by ethnomethodological workplace studies of technology-rich environments (e.g., Goodwin, 1996; Hindmarsh & Heath, 2000), we uncover participants' methods for transforming an initially vague reference into a ratified, mutually established mathematical object. Our fine-grained analysis allows us to reconstruct the interactional work that leads to the joint accomplishment of the covariant variable *distance*. We find two key interactional mechanisms involved in this process: (1) responsiveness to indexicality leading to increased specificity; and (2) responsiveness to multimodality in order to demonstrate mutual understandings.

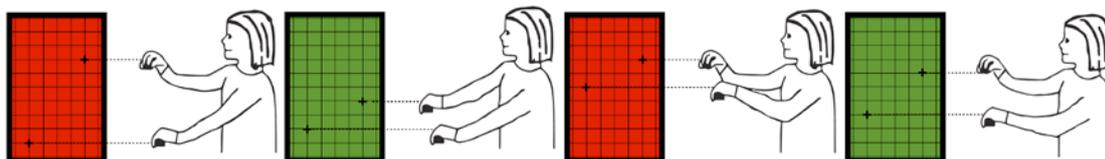


Figure 1. When the Mathematical Imagery Trainer for Proportion (MIT-P) is set to a 1:2 ratio, the screen is green only when the right hand remote is twice as high as the left hand remote; otherwise it is red.

The work of negotiating objects in complex perceptual fields

“The unity of a geometrical configuration or a melody is a general problem of organization.”
(Gurwitsch, 1964, p. 55)

The MIT-P provides an interactive context for users to explore ideas about proportionality through sensorimotor activity. Children operate two independent, hand-held remotes to manipulate the heights of two corresponding cursors on a computer screen (see Figure 1). The system generates green feedback only when the cursor heights correspond to a particular pre-set, concealed ratio (e.g., 1:2 in Figure 1). Whenever the cursor heights do not fulfill the secret ratio, the screen is red. Children engage in guided exploration as they develop and articulate strategies

for turning the computer screen green. Adult tutors guide children's discoveries and facilitate shifts in their strategies for generating green feedback by introducing a series of canonical mathematical artifacts—first a Cartesian grid, then numerals to label that grid. Children enlist these artifacts to generate more robust descriptions of strategy, identify patterns, and develop sophisticated quantitative models that account for the behavior of the device (Abrahamson et al., 2012).

The MIT-P constitutes a complex, dynamic perceptual field of potentially relevant phenomena. Learners discover many perceptual patterns (e.g., haptic, visual) in this phenomenal field of potential relevancies, but a key challenge is getting others—either peers or adult tutors—to perceive these collections of features as significant patterns, as well. In the rich multisensory, multimedia landscapes of TEELEs, this does not merely entail directing one another's attention to visually-present physical features (e.g., in the case of the MIT-P, the cursors on the computer screen). Perceived patterns consist of visually available features in the environment, but they may also incorporate imaginary realms of possibilities that move beyond the immediate spatio-temporal circumstances (Nemirovsky et al., 2012).

To illustrate this, consider, for example, the famous Rubin Vase (see Figure 2). Perceiving the objects *face* or *vase* is not merely determined by the invariant, physical sense data (e.g., the colors, or the boundary between the colors). The same physical features support both imaginary perceptual possibilities. In order to perceive one possibility over the other, the same data has to be *organized* into different configurations (Gurwitsch, 1964). Organizing particular features into certain relationships with one another gives an assemblage of details coherence as an object (Garfinkel, 2008), whether it is real or imaginary.



Figure 2. Rubin Vase.

Configuring objects for one another by instructing each other's attention is a ubiquitous practice in technology-rich technical and scientific workplace settings (Garfinkel et al., 1981; Goodwin, 1996; Hindmarsh & Heath, 2000; Stevens & Hall, 1998). Identifying patterns in vast amounts of information and interpreting their significance is a daily endeavor at these sites. Personnel work together to monitor multiple, heterogeneous streams of information within complex multimedia environments. They must actively negotiate, constitute, and render particular physical and imaginary features of the scene relevant to the projects at hand in order to orient each other to potentially significant patterns (Garfinkel et al., 1981; Hindmarsh & Heath, 2000). Professionals use a variety of situated, embodied strategies as part of this *collaborative imagining* (Murphy, 2009)—including purposefully vague references (Garfinkel et al., 1981; Goodwin, 1996) and gestural mimicry (Becvar et al., 2005)—to configure and mutually establish objects. Finding creative ways to highlight and render phenomena of interest publicly inspectable from an abundance of data is an important STEM disciplinary practice, essential to building scientific and mathematical knowledge together (Latour, 1999).

In the guided-discovery learning that is characteristic to TEELEs like the MIT-P, participants must work to coach each other in how to attend to the rich haptic and visual phenomena they deem relevant and important in a particular moment. Following the lead of Stevens and Hall (1998) and Koschmann and Zemel (2009), we believe that much can be learned about processes of discovery and objectification in pedagogical environments by looking for parallels in the interactional methods employed by professional scientists and technical workers. Inspired by studies of technical workplaces, we attempt to understand the complexity of the interactional work that goes into configuring mutually available objects in the complex field of potential relevance in discovery learning with the MIT-P. We approach learning as a process of producing intersubjectively-achieved ratified understandings (Koschmann & Zemel, 2009; Schegloff, 1991; Stevens & Hall, 1998). Our goal is to reveal the specific *interactional methods* employed by a learner and an adult tutor to accomplish a ratified understanding of a mathematical object.

A methodology for tracing intersubjectivity as an interactional achievement

A powerful architecture for *intersubjectivity* is built into the sequential nature of interaction: Each conversational turn consists of an updated display of the speaker's current understanding of the project underway, which is designed in response to an antecedent display of understanding by the prior speakers (Schegloff, 1991). To trace how an initially vague, proposed perceptual pattern was transformed into a clearly defined intersubjectively-achieved mathematical object, we engage in a fine-grained analysis of the participants' embodied interactional practices. Our praxeologically-based study of interaction is informed by ethnomethodology, an analytical approach that focuses on the *in situ* methods that people use to coordinate mutually intelligible courses of action together (Garfinkel & Sacks, 1970).

Following ethnomethodology, we adopt the view that the meaningfulness of the situations in which participants find themselves is not given a priori and need not be accounted for by partially overlapping or matching psychological states (Schegloff, 1991). Instead, what enables participants to coordinate a given course of action (e.g., having an argument, making a discovery) in any actual situation is a sense of mutual intelligibility

that must be built sequentially, moment-by-moment, with the local resources at hand (e.g., turns at talk, bodies, spatial arrangements, tools) through a relentless process of incremental, public displays of revisions and ratifications of meaning (Garfinkel & Sacks, 1970; Goodwin, 1996; Sacks, 1992; Schegloff, 1991). Our analysis seeks to identify the particulars of the interactional resources that participants mobilize to create these scenic displays of the sense they are making of a situation (Garfinkel and Sacks, 1970).

From evidently-vague reference to reified mathematical object

Our episode is drawn from a corpus of task-based interviews with 4th, 5th, and 6th-grade children using the MIT-P. It occurs about 40 minutes into an interview, in which Boaz, a 5th-grader, has been working with two adult tutors, Dean and Devon. The MIT-P is currently set to the secret ratio of 1:2. In the prior 40 minutes, Boaz has offered several accounts of how to make the screen green (e.g., the right-hand cursor always has to be higher than the left-hand cursor). With the newly introduced Cartesian grid, Boaz proposes he could use “the squares” to see how much higher the right cursor is compared to the left cursor.



Figure 3. Devon (adult tutor) and Boaz (child) each operate a remote to control a cursor on the MIT-P screen, while Dean (a second adult tutor) looks on.

Devon suggests that he and Boaz try this out. Devon moves the left cursor up one grid box at a time, and Boaz adjusts the vertical position of the right cursor to make the screen green (see Figure 3). They advance the cursors up the screen in this fashion, until suddenly Boaz gasps “OHHhhhhh!,” marking what comes next in Excerpt 1 as newly revealed information: a *proposed discovery* (Koschmann & Zemel, 2009).

Excerpt 1¹

- 1 **Boaz:** Each time it's (.) increasing the square (1.0)
 2 so that one was one then we got to two (.)

Boaz moves cursors to...	...then moves them to...	...then he moves to...	...and finally...
<p>3 Boaz</p> <p>and I'm thinking °sort of like° o::ne</p>	<p>4 Boaz</p> <p>and then this one's two</p>	<p>5 Boaz</p> <p>then like here this one is (.) THREE</p>	<p>6 Boaz</p> <p>and then this one is (3.5) it goes upta FOUR</p>
<p>7 SO:o it increases the number of squares each time.</p>			

In Excerpt 1, line 1 (1.1), after both Boaz and Devon have lowered the remotes, Boaz offers this candidate proposal of a discovery: “Each time it’s increasing the square.” With this vague reference, the sense of this discovery is not yet publicly available to Devon and Dean. The statement functions as a *prospective indexical* (Goodwin, 1996) that projects reference into the future toward a forthcoming resolution: its sense will need to be further specified and elaborated as the sequence unfolds. A word, phrase, or expression is *indexical* if its intelligibility is only recoverable by considering the context in which it is embedded (Garfinkel & Sacks, 1970).

Boaz’s turn in 1.1 initiates Devon and Dean into attending to and examining subsequent turns in the sequence for the sense of his vague reference. The prospective indexical sets up an initial framework for interpretation: It renders what follows (in 1.2 and 1.4–1.6) as *instructions for perceiving* the specific phenomenal features that make up the sense of “it’s increasing the square.” By pointing forward toward something whose details are not yet specified or ratified by participants, prospective indexicals give objects-in-progress a quasi-presence in the interaction, allowing for their ongoing negotiation (Goodwin, 1996). Garfinkel et al. (1981) first revealed the function of such *evidently-vague reference* in the context of professional scientific work, tracing its

key role in the embodied, social negotiation of a certifiable scientific object of knowledge (the astronomical discovery of a pulsar). Koschmann and Zemel (2009) showed how, in a similar fashion, prospective indexicals allow students to negotiate and establish ratified kinematics discoveries while exploring a pedagogically-designed physics microworld.

In 1.2, Boaz offers a first, verbal elaboration. Then, he picks up both remotes and in 1.3 through 1.6, he narrates a sequence of activity with the MIT-P. He pauses rhythmically after positioning both cursors and matches each position to a verbalized numerical value. Starting at 1.4, he also verbally demarcates each configuration with “this one.” Through this multimodal performance, Boaz offers an ordered series of parallel cases of *it’s increasing the square* for inspection. Boaz has narrowed the complex perceptual field of the MIT-P and configured an ephemeral, animated realm of possibilities for Devon and Dean to search and examine for potentials of the proposed discovery. In 1.7, Boaz uses a similar statement to 1.1 to now *retrospectively* index his just-prior performance (he creates a *retrospective indexical*).

Despite his embodied demonstration, Boaz’s instructions for how to appreciate his proposal are not unequivocal for his audience. The numbers Boaz calls out are consistent with at least two possibilities during each event: (1) the positions of the left cursor in each case; and (2) the total number of squares between the vertical positions of the left and the right cursors. There are also several “increases” one might perceive and attend to: (1) the height of the left cursor; (2) the height of the right cursor; (3) the heights of both cursors; and (4) the distance between the cursors. Notably, all these potential patterns contain imaginary features: The properties height and distance have no material instantiation (as opposed to, for example, the cursors or the lines of the Cartesian grid on the screen) in the phenomenal field. A sufficient amount of interpretive work is still necessary to negotiate what features constitute the pattern Boaz is proposing. In Devon’s adjacent, subsequent turn, he shows responsiveness to the ambiguity of Boaz’s explanation (Excerpt 2).

Excerpt 2

- 1 **Devon:** And (1.2) so when you say it increases what- if you were speaking
- 2 to someone on the phone and they can't- they can't see what's
- 3 going on here (.) and they say it inc- WAIT WOAHH woah woah, what
- 4 are you- what are you talking about, what is increasing?

In Excerpt 2, Devon laminates the immediate spatio-temporal present with a hypothetical situation by voicing an imaginary interlocutor who would have limited perceptual access to the current scene (no visual or haptic access) at the other end of a telephone line. We can appreciate the function of this turn as a “repair initiator” (Schegloff, 1991). It marks a source of trouble in understanding Boaz’s just-prior formulation. Participants in interactions have a variety of coordinated repair practices at their disposal to mark and resolve troubles in speaking, hearing, and understanding. These practices provide mechanisms for incrementally establishing intersubjectivity turn-by-turn (Schegloff, 1991).

In this case, the use of “what” has multiple functions. (1) It pinpoints the exact location of the trouble: Devon has swapped Boaz’s “it” with “what” to build the hearably parallel construction, “*what* is increasing” (2.4), which targets the *it* as what needs repair. (2) The use of “what” also displays a claim of understanding: It operates as a candidate understanding—an ontological interpretation of what Boaz originally meant. The understanding Devon displays here is that Boaz’s *it* is some kind of *thing* (as opposed to a person—who; a time—when; or a place—where). Finally, (3) Devon’s “what” projects a format for how Boaz should re-design his initial description (how he should do the third-turn repair): He should further specify the nature of this *thing*.

Devon’s turn explicitly attends to the prospective indexical nature of Boaz’s prior turns and *renews* Boaz’s antecedent coda (from 1.7) as prospectively indexical: The lamination with the hypothetical telephone situation positions 1.7 as *continuing to be* prospectively indexical. It renders Boaz’s explanation as still unresolved, unfinished, and in need of elaboration. Devon’s turn encourages Boaz to elaborate the instructions for perceiving the phenomenon in a way that would be sensible for someone not immediately part of the current scene (the person at the other end of the phone line; 2.2). A hallmark of the work of scientists involves the design of representations to capture and archive features of phenomena, and to make them available in places and times beyond the immediate material circumstances in which they were first discovered (Latour, 1999). Thus, this turn creates an opportunity for Boaz to participate in the disciplinary practice of generating a context-independent representation of a phenomenon. In what follows, Boaz does re-design his description, but this does not immediately result in consensus about the object-in-progress.

In Excerpt 3.1, Devon displays his orientation to Boaz’s initial description of the pattern (“by one”) as a potential object (“that thing”). He makes a strong claim that his understanding is the same as Boaz’s and that all that remains for them to figure out is what to call *it* (3.2). At first, Boaz does not contest Devon’s development of

his idea. He suggests naming *it* “plus” and goes on to fit “plus” into Devon’s construction from 3.1: “plus is increasing” (3.3). Devon’s repetition in 3.4 is another third-turn repair initiator, now locating trouble in “the plus” and providing an opportunity for Boaz to rephrase or elaborate what the plus is in the subsequent third turn. However, Boaz does not elaborate in 3.5 and instead simply says “yeah.” In 3.6, Devon reveals his understanding that “plus” refers to only one cursor at a time. Boaz responds that “both” should be called plus (3.7). Again, there are multiple possible interpretations of Boaz’s assertion: It could refer to *the pair* of the left and right cursors, or something else entirely, such as the increasing space the cursors bound when they are moved together in the configurations from Boaz’s demonstration (cf. Excerpt 1).

Overall, in turns 3.6–3.9, Boaz and Devon display significant difficulties in understanding each other. Unable to project the end of each other’s turns, both Boaz (3.7) and Devon (3.8) interrupt each other mid-sentence. This trouble projecting the boundaries of turns signals a failure to appreciate their ideational content.

Excerpt 3

- 1 **Devon:** How should we call that thing that is increasing by one?
- 2 I can see it also but I'm not sure what to call it.
- 3 **Boaz:** Hmmmm call it plus, plus is increasing=
- 4 **Devon:** =The plus?=
5 **Boaz:** =°Yeah°=
- 6 **Devon:** =Wa-wh- ss the one on the right? Or the-
- 7 **Boaz:** -All right, both of them should be called plus cause that's-
- 8 **Devon:** -Yeah, yeah, we call these things crosshairs, (°but whatever°), but
- 9 oka::y so they're increasing by one
- 10 **Boaz:** °Right°. NO not they increase- the first one increases by one and it
- 11 gets higher and then it goes two, three, four, five, si- it goes
- 12 um um odd even odd even each time

The breakdown in Excerpt 3 powerfully demonstrates that mutual understandings are incrementally renewed and re-achieved on a turn-by-turn basis and can be quite suddenly lost. Participants *ongoingly* generate resources for their partner(s) to determine whether they are still oriented to the same phenomenon in the same way *and* their interlocutor(s) must continually monitor these resources for evidence of misalignments. Each subsequent response contains further evidence for the original speaker that their instructed perceiver is or is not orienting appropriately to the features, displaying successful or unsuccessful apprehension of the intended phenomenon (Hindmarsh & Heath, 2000). In our case, Boaz is instructing Devon’s perception. He must carefully attend to Devon’s responses for evidence that Devon is apprehending the same phenomenon, and move to remedy detected displays of misunderstandings.

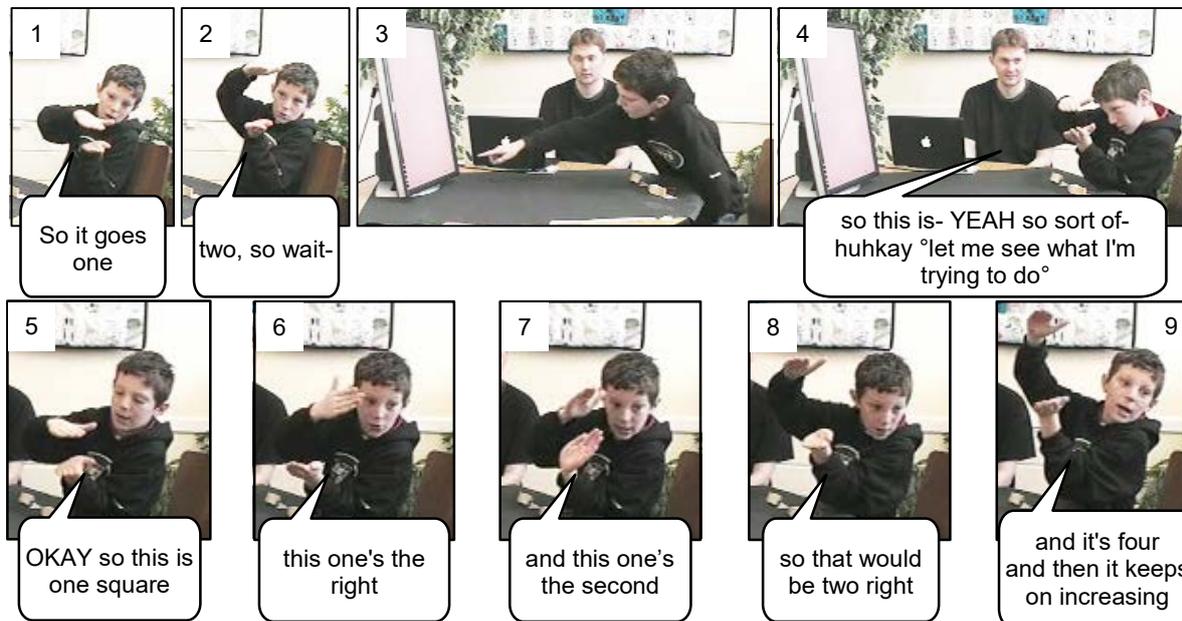
In Excerpt 2, Devon had both *claimed* and *demonstrated* an understanding of Boaz’s initial proposal (Excerpt 1) that Boaz could monitor (Sacks, 1992; p. 252). Boaz had not challenged this demonstration and seemed to treat it as evidence that Devon had attended to the same features Boaz was instructing him to perceive. Now, in 3.8–3.9, Devon’s turn provides another key demonstration of his understanding for Boaz to assess. Devon demonstrates that he is orienting to Boaz’s “plus” (from 3.7) as a reference to each individual cursor by proposing the alternate, pluralized label, “crosshairs” (3.8). Then, he carries the plural into the indexical description “they’re increasing” (3.9). With this turn, Devon creates a public display of what features he is attending to: the two cursors. However, now, Boaz’s challenge (3.10–3.12) suggests that he treats Devon’s latest display as evidence that Devon *has not* appropriately oriented to what Boaz wishes to highlight for him.

The trouble in Excerpt 3 leads to productive work for the negotiation of the object-in-progress. Devon’s display in 3.8–3.9 creates a new priority for Boaz to re-design his approach. In 3.10–3.12, Boaz repairs his description by elaborating it significantly in response to Devon’s displayed misapprehension. In Excerpt 4, Boaz further instructs Devon (and Dean who has remained silent but still participates as an on-looking audience) in how to see the phenomenon: He uses a new approach for assembling its features, re-designing his description through a rich performance of gestures and speech to orient his audience.

Boaz starts with his right hand (palm-down) above his left hand (palm-up) in 4.1. Then, he moves both hands farther apart as he lifts both of them higher in the air (4.2). In 4.3, he leans forward to point at the screen. However, unable to reach, he moves back into his seat and creates a space between his hands, sighting through it at the monitor (4.4). In 4.5, with this bounded space assembled, he turns his body toward Devon. He instructs Devon how to see each of his hands: The right hand stands in as the right cursor (4.6) and the left hand stands in

as the second, left cursor (4.7). Boaz then increases the spacing between his hands and moves them both upward (4.8). Then he raises both hands even further up and makes the space between each hand even larger (4.9). Each time he lifts his hands and increases the spacing between them, Boaz calls out increasing numbers. He concludes with the retrospective indexical, “it keeps increasing,” pointing back at his performance.

Excerpt 4



Through this multimodal performance, the *material nothingness* surrounding Boaz’s two hands is transformed into something significant: Boaz generates a perceptually available *empty, vertical space* (Nemirovsky et al., 2012) for Devon and Dean to attend to. To create this empty space from nothingness, Boaz configures the outline of the space with his hands, lifting it as a figure from the ground, much like bounding the contour of the vase in Figure 2 lifts it out of a formless background. The empty space is given quasi-presence through bodily activity; it is an imaginary feature that Boaz must instruct Devon and Dean to experience as part of the unfolding object. This now quasi-present imaginary feature allows Boaz to juxtapose several cases of the increasing, covariant quantity: The empty space is getting bigger *as* it moves upward.

Excerpt 5

1 **Devon:** Should we come up with some kind of word=



In Excerpt 5, Devon displays his understanding by recreating Boaz’s gestures. He elaborates the original gestures by explicitly tracing the space between his two hands (5.2–5.3). As he says “thing” (5.4) he collapses and re-opens his hands to emphasize his attention to the space between his own hands. Co-timed with the word “growing,” Devon increases the space between his hands as he moves them upward (5.5–5.6). This multimodal revoicing of Boaz’s idea, with its explicit effort to highlight the empty space, publicly demonstrates that Devon has organized the same features in the same way as Boaz.

Boaz does not challenge any part of Devon's interpretation, suggesting that he now finds Devon's orientation appropriate: Devon's multimodally displayed understanding of the indicated empty space as a *thing that grows as it is elevated* is visibly in alignment with Boaz's displayed understanding. Boaz demonstrates his satisfaction with the currently intersubjectively-achieved configuration of real and imaginary features—i.e., the presently accomplished organization and mutual availability of a perceptual pattern—by now going along with the proposed project of naming it. He asks if Devon has any suggestions: “Do *you* have a name for it?” Devon offers the label “distance” that Boaz promptly affiliates with: “Oh, ye:::ah, the distance.” When Devon further elaborates, “the distance between the hands” while repeating his gestures from 5.2–5.4, Boaz smiles, laughs, and accepts the name: “yeah.” Devon finally proposes a candidate consensus conclusion, “so the distance is growing?” and Boaz confirms it: “The distance is growing.”

This name—“distance”—now comes to stand in for a reified assemblage of an ordered collection of real and imaginary features. With this, a *mathematical object* has been jointly established: It is a covariant variable that increases while the hands—and therefore the cursors on the screen—move upward. This mathematical object, now called *distance*, itself is an imaginary object that has no material presence. Rather, it is a *property* of the configuration of the cursors that was made perceivable through Boaz's and Devon's figuring of the height-dependent empty space between their hands.

Conclusions and implications

In this article, we demonstrated how a young learner and an adult tutor working with a technology-enabled embodied learning device for mathematics jointly establish a mathematical object—a covariant variable. Both active participants engage in nuanced forms of multimodal interactional work to achieve the mutual availability of the object as a coherent assembly of physical and imaginary features. The process of reifying this mathematical object is distributed temporally across the interaction, across embodied activity (gestures), across material artifacts (the MIT-P), and notably, across individuals. In this case, the objectification that occurred is irreducible to any individual mind. In particular, we identified two key interactional mechanisms occurring between adult and learner as vital to the intersubjective achievement of the ratified object: (1) Participants' attention and propagation of prospective indexicals to occasion their elaboration; and (2) participants' attention to and mimicking of gestures (multimodal revoicing) to demonstrate mutual understanding. Our findings elaborate current frameworks that trace objectification as a form of ontogenesis (Radford, 2003).

While specificity and preciseness are considered hallmarks of sophisticated scientific and mathematical practice, our findings suggest that indexicality and vagueness play a beneficial role in configuring and establishing mutually available mathematical objects. In our case, Boaz's initially vague, prospective indexical description provides the key resources for interpreting his subsequent performance as the proposal of a significant pattern to be attended to. Devon's responsiveness to this prospective indexical generates an environment that allows for the continuing negotiation and specification of the object-in-progress. It also affords an opportunity to engage in the STEM disciplinary practice of developing context-independent representations of phenomena. We therefore add the use of prospective indexicals as a productive form of “systematic vagueness” that enables discourse and negotiations between teachers and learners about disciplinary content (Newman et al., 1989; p. 62). An implication of this study is that explicit responsiveness to indexicality and multimodality in learners' explanations can foster productive engagement in STEM disciplinary practices within TEELEs.

Our episode also challenges interpretations of zones of proximal development where learning is seen as the result of agentive adults or experts structuring a situation for a child or novice, and regulating their participation. What unfolds in our episode is not a simple case of an adult knowing the cultural future and bringing it to bear on a child's proposed interpretation of events (cf. prolepsis, Stone, 1993). Conceptualizations of scaffolding as an adult-initiated, asymmetrical process cannot capture the reciprocal instruction necessary for jointly establishing the covariant variable in our episode. Boaz must instruct his adult tutors in how to experience what he is experiencing in order to make the subsequent negotiation of meaning possible. He must also carefully monitor his tutor's response for evidence as to whether they have attended to and interpreted what he has highlighted, and he must be responsive to their interpretations. Our case supports recent efforts to re-capture and highlight the symmetry of participation in interactional zones of proximal development by analyzing instruction as a *bi-directional* process between experts and newcomers (Roth & Thom, 2009).

With respect to this year's conference theme, we suggest that empowering learners involves appreciating their symmetrical role in the negotiations of meaning in pedagogical situations. We emphasize the importance of finding ways to support learners' use of indexicality and multimodality. As demonstrated in our episode, both are productive strategies for building mutual understandings, and they should be considered valuable and relevant ways of practicing and knowing in scientific and technical disciplines.

Endnotes

- (1) Jeffersonian Transcript Conventions (Jefferson, 2004): (.) for a micro-pause <1 sec; (2.5) for timed pause in seconds; CAPITALS for loud speech; = for latching; :: colons for elongated speech; underline for emphasis; ° for quiet speech.

References

- Abrahamson, D., Gutiérrez, J., Charoenying, T., Negrete, A., & Bumbacher, E. (2012). Fostering hooks and shifts: Tutorial tactics for guided mathematical discovery. *Tech, Knowledge and Learning*, 17(1-2), 61–86.
- Becvar, L. A., Hollan, J., & Hutchins, E. (2005). Hands as molecules: Representational gestures used for developing theory in a scientific laboratory. *Semiotica*, 4(156), 89–112.
- Enyedy, N., Danish, J., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *Intl Journal of Computer-Supported Collaborative Learning*, 7(3), 347–378.
- Garfinkel, H. (2008). *Toward a sociological theory of information*. (A. Rawls, Ed.). Boulder: Paradigm.
- Garfinkel, H., Lynch, M., & Livingston, E. (1981). The work of a discovering science construed with materials from the optically discovered pulsar. *Philosophy of the Social Sciences*, 11, 131–158.
- Garfinkel, H., & Sacks, H. (1970). On formal structures of practical actions. In J. C. McKinney & E. Tiryakian (Eds.), *Theoretical sociology: Perspectives and developments* (pp. 337-366). NY: A-P-C.
- Goodwin, C. (1996). Transparent vision. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 370–404). Cambridge: Cambridge University Press.
- Gurwitsch, A. (1964). *The field of consciousness*. In R. M. Zaner & L. Embree, (Eds.), New York: Springer.
- Hindmarsh, J., & Heath, C. (2000). Sharing the tools of the trade: The interactional constitution of workplace objects. *Journal of Contemporary Ethnography*, 29(5), 523–562.
- Jefferson, G. Glossary of transcript symbols with an introduction. In G.H. Lerner, (Ed.) *Conversation analysis: Studies from the first generation*. John Benjamins, Amsterdam, 2004, 13–31.
- Koschmann, T., & Zemel, A. (2009). Optical pulsars and black arrows: Discoveries as occasioned productions. *Journal of the Learning Sciences*, 18(2), 200–246.
- Latour, B. (1999). *Pandora's hope: Essays on the reality of science studies*. Cambridge: Harvard Univ. Press.
- Lindgren, R., & Moshell, J. M. (2011). Supporting children's learning with body-based metaphors in a mixed reality environment. In *Proc. of the IDC Conf.* (pp. 177–180). NYC: ACM.
- Lyons, L., Slattery, B., Jimenez, P., Lopez, B., & Moher, T. (2012). Don't forget about the sweat: Effortful embodied interaction in support of learning. In *Proc. of TEI Conf.* (pp. 77–84). New York: ACM.
- Murphy, K. M. (2009). Imagination as joint activity: The case of architectural interaction. *MCA*, 11(4), 37–41.
- Nemirovsky, R., Kelton, M. L., & Rhodehamel, B. (2012). Gesture and imagination: On the constitution and uses of phantasms. *Gesture*, 12(2), 130–165.
- Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive change in school* NY: Cambridge University Press.
- Radford, L. (2003). Gestures, speech, and the sprouting of signs: A semiotic-cultural approach to students' types of generalization. *Mathematical Thinking and Learning*, 5(1), 37–70.
- Roth, W.-M., & Thom, J. S. (2009). The emergence of 3D geometry from children's (teacher-guided) classification tasks. *Journal of the Learning Sciences*, 18(1), 37–41.
- Sacks, H. (1992). *Lectures on conversation: Volume II*. (G. Jefferson, Ed.). Cambridge: Blackwell Publishing.
- Schegloff, E. (1991). Conversation analysis and socially shared cognition. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 150–170). D.C.: APA.
- Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. In *Talking mathematics in school: Studies of teaching and learning* (pp. 107–149). NY: Cambridge University Press.
- Stone, C. A. (1993). What is missing in the metaphor of scaffolding? In E. A. Forman, N. Minnick, & C. A. Stone (Eds.), *Contexts for learning: Sociocultural dynamics in children's learning*. NY: Oxford University Press.

Acknowledgments

This material is based on work supported in part by the National Science Foundation under Award No. 1321042 and in part by the Institute of Education Sciences under Award No. R305B090026.

Developing Pre-Service Teachers' Professional Vision Through Collaborative Multimedia Artifacts

Andi M. Rehak, Andrea S. Gomoll, Cindy E. Hmelo-Silver, and Joshua A. Danish
amstrack@indiana.edu, agomoll@indiana.edu, chmelosi@indiana.edu, jdanish@indiana.edu
Indiana University

Abstract: Effective teachers are constantly adapting their classroom practice to best meet students' needs. However, it is difficult to adapt one's teaching practices without first noticing the key features of classroom activity. This kind of noticing—professional vision—is a cultivated skill that requires a combination of understanding theory and experience with practice. When theory-practice connections are not strong, teachers tend to fall back on “what works.” Demands to support pre-service teachers in developing stronger theory-practice connections are increasing, yet this remains a challenging goal. The qualitative analysis presented in this paper offers insight into how creation of multimedia artifacts can support pre-service teachers in connecting theory to practice while developing their professional vision. Findings suggest the process of creating and collaboratively refining multimedia artifacts in small groups improves pre-service teachers' ability to make theory-practice connections and provides opportunity for them to exercise their emerging professional vision.

Keywords: pre-service teachers, collaborative learning, professional vision, activity theory, video case studies

Introduction

Effective teachers need to understand student learning and adapt classroom practice to meet learners' needs on a daily basis. In the demanding classroom environment, teachers tend to fall back on “what works” approaches if strong connections between theory and practice are not made during pre-service experiences (Ball & Forzani, 2009). Often, teacher education relies on practicum experiences to establish these connections, yet there is little evidence to support that pre-service teachers make theory-practice connections during practicums (Allsopp, DeMarie, Alvarez-McHatton, & Doone, 2006). Furthermore, teacher training programs and assessments are beginning to ask pre-service teachers to observe, describe, and evaluate others' teaching practices and their own in order to obtain licensure (e.g., edTPA, 2015). This study explores pre-service teachers' development of this ability to see the classroom in this expert manner as *professional vision*—the ability to use a theoretical lens to support noticing key features in a learning environment, interpreting them, and then responding (Sherin, 2001). We explored the development of this skill by looking at the ways in which students annotated classroom video and gave one another feedback. Viewing classroom video provides a unique opportunity for students to explore the complexities of practice and to see how learning theories apply to the classroom (Van Es & Sherin, 2002).

To explore the potential for using shared video resources to support the development of professional vision, we made use of the RUanalytic tool (see Maher, Palius, Maher, Hmelo-Silver, & Sigley, 2014). The RUanalytic tool is a video annotation tool that was designed to work with a large video repository, the Video Mosaic Collaborative (VMC). This video repository houses an extensive video collection from a longitudinal study of mathematics learning in K-12 classrooms (<http://videomosaic.org/>). Users of the tool are able to create a multimedia artifact consisting of “events” clipped from VMC videos and written annotations of these events. The final product is a VMCAlytic, a multimedia artifact that plays the clipped events in sequence, accompanied with the author's text descriptions. This video annotation tool has been developed and used in previous research to support pre-service and in-service teachers' development of pedagogical skills, including teacher professional development with a focus on mathematics education (Maher et al., 2014).

Informed by the parameters of the edTPA assessment (edTPA, 2015) and university guidelines (Kunzman, 2014), two class sessions of activities involving the video annotation tool were collaboratively designed by a team of researchers of which the course instructor was a member. Our conjecture was that creating multimedia artifacts would provide a unique opportunity for pre-service teachers to observe moments in a classroom and then connect learning theories to these moments as a means of developing professional vision. This study was driven by the following research questions: (1) How do these pre-service teachers use learning theory when describing episodes of teaching and learning? (2) How do these pre-service teachers' multimedia artifacts change during class activities with peers? (3) What is the nature of these pre-service teachers' professional vision as displayed in their multimedia artifacts?

Theoretical perspectives

Our design was guided by activity theory (Engeström, 2014), which is rooted in Vygotsky's (1978) sociocultural theory of learning. Activity theory asserts that collective human activity and thus learning, is a process mediated by both individual and sociocultural factors (Engeström, 2014). Mediators "stand between" a subject and object of activity; subjects view their objective through the lens of the mediators they are using, and the mediator thus transforms the entire activity. The subject is the individual or group of individuals (e.g., students) engaging in activity. Object refers to the overarching motivations for the activity as realized by the subject (Cole, 1996). Because the object of activity is constructed by the subject, it may or may not be aligned to what the designer of activity intended the object to be (Rajala & Sannino, 2015). From an activity theory perspective, mediating artifacts are the products and practices of the culture. Although distinct, these mediators cannot, and do not, function independent of one another and thus need to be analyzed as interrelated (Roth, 2007). Here, professional vision acted as a mediator, shaping student activity and learning. We were interested in the groups' objects of activity, as well as the ways in which the RUanalytic tool mediated students' articulation of professional vision.

A crucial element of teaching is being able to see key features of a learning environment and respond to them. The development of professional vision involves honing the practice of teaching through attention to noticing these salient features in the classroom (Borko, 2004). Professional vision involves "socially organized ways of seeing and understanding events that are answerable to the distinctive interests of a particular social group" (Goodwin, 1994, p.606). Goodwin's definition highlights how cognition and perception are intertwined with the social norms of a discipline. Thus, developing expertise within a given discipline requires the ability to "see" the world through the lens of that social group (Stevens & Hall, 1998). In a classroom setting, teachers' professional vision mediates active noticing and the interpretation of key elements of the environment, including student reasoning and features of peer interaction (Maher et al., 2014; Sherin, 2001; 2007). This skill includes identifying student strategies, inferring student levels of understanding, and developing a response based on these levels of understanding. Noticing in this capacity is complex and variable (Jacobs, Lamb, & Phillip, 2010). The reflective use of classroom video over time mediates changes in noticing and interpretations of what was noticed (Sherin & van Es, 2005). As pre-service and in-service teachers develop expertise in the skill of interpreting classroom situations, they use their understanding of learning theory to isolate the important features of what is happening. In this process of noticing, they then connect the specific evidence and learning theories to support conclusions about how to respond. Professional vision mediates their teaching activity and is continually under construction. Professional vision develops as students become enculturated into the teaching community.

Additionally, we were interested in how the opportunity to engage in activity with a particular division of labor informed by *collaborative learning* (CL; O'Donnell & Hmelo-Silver, 2013) contributed to pre-service teachers' multimedia artifact construction. In CL, group members mutually influence one another and participate equally in knowledge building (O'Donnell & Hmelo-Silver, 2013), making CL valuable for helping students to see themselves as members of a community (Hakkarainen et al., 2013). This sociocultural perspective of CL frames our work to integrate learners into not only a classroom community, but also into a larger community of educators. To explore the impact of peer interactions, we examined pre and post-collaboration versions of multimedia artifacts, and analyzed video of two groups of pre-service teachers creating their multimedia artifacts through the lens of activity theory to better understand student interactions.

Methods

Participants

The participants in this study included 27 pre-service teachers enrolled in "Educational Psychology for Elementary Teachers," at a large Midwestern University. The course covers seven core units: behaviorism, social cognitive theory, information processing theory, constructivism, sociocultural theories, motivation, and assessment. These students intended to teach primary grade students (5-12 year olds). The students in this course ranged from first to third-year undergraduate students; 22 of the 27 were female. They worked in groups of 3-5 students that were previously organized by the instructor. In our activity theory driven analysis, the interactions and multimedia artifacts of two randomly selected groups of three students are examined with the aim of understanding the role of collaboration in the articulation of pre-service teachers' professional vision.

Activity design

Our goal was to provide an opportunity for the students to see theory in action (see Gomoll et al, 2016), while beginning to recognize how learning, and principles of learning theories, might be experienced in real classrooms. The research team selected short video clips (7-12 minutes long) from the repository that showed both students

and teachers in action. These were selected with attention to length and connection to several theoretical frameworks covered in the course (i.e. behaviorism, constructivism, social cognitive theory, information processing, and sociocultural theories).

Prior to the first day of activities, students were asked to independently view an assigned video clip and to write down three “interesting” moments. Students in each small group worked from the same video clip, while each group had a different video. The open-ended prompt was intended to help the students see the video as something relevant to them, and to help prepare them for the class session’s activities. In class, students were asked to examine the same videos again, but to explain their highlighted moments using learning theory. To do this, students independently created a multimedia artifact where they connected their chosen events and provided accompanying descriptions, which articulated how learning theories covered within the course would explain the success or failure of selected events. Students were provided with the opportunity to identify new events if they felt their initial choices did not support this analysis, but few did. In the following class session, students of the same group provided feedback on one another’s multimedia artifacts—allowing them to explore the differences in how they had connected theory to the same videos. Following this feedback, students had the opportunity to edit their work to reflect the feedback they had received. The final product created by students was a multimedia artifact with three “events,” or short clips, selected from the longer video. Each of these events, about thirty seconds to one minute in duration, was annotated with a description pointing the reader to important episodes of teaching and learning.

Data sources

All version histories of multimedia artifacts were archived to track changes in each student’s artifact over time. During the two class periods in which students worked on their multimedia artifacts, three video cameras captured whole class activity and two small groups’ interactions (Group A and D). Only two students in group A were present on the second day of the intervention, Natalie and Isabella (pseudonyms). Addison, Riley, and Isaac comprised group D. Students also took notes about their peer feedback during the second day via Etherpad—an online platform that allows multiple participants to type in a text document at once. Students participated in an in-person whole-class debriefing session at the end of the experience. While prior analysis has addressed the multimedia artifacts created by all students during this intervention (see Gomoll et al., 2016), here we focus on the multimedia artifacts of two groups, A and D that were randomly chosen to be video recorded. This additional data source allowed a more fine-grained examination of group collaboration in the face-to-face classroom environment as well as students’ final multimedia artifacts. The last version of the VMCanalytic from day 1 is considered the pre-collaboration analytic, and the final version of the VMCanalytic, the one students submitted for the course assignment was considered the post-collaboration version.

Analysis

Viewing each small group as an activity system helped us to better understand the nature of students’ professional vision articulated in their multimedia artifacts. We analyzed student artifacts and video of two groups of pre-service teachers’ collaborative work constructing and refining multimedia artifacts—comparing the activity systems of these two groups to each other and to the activity system envisioned by the instructor. The two groups show the curriculum being reinterpreted by the students as they constructed unique group-specific activity systems, allowing us to construct contrasting cases. Neither of the activity systems was identical to the activity system intended by the designers, yet both led to productive outcomes in the form of satisfactory multimedia artifacts. We began with a representation of the *intended* activity system for small group work as envisioned by the curriculum designers (see Figure 1). Underlying the object of students articulating their professional vision and revising their multimedia artifacts is the assumption that the assignment requirements students mediate activity in such a ways that students would utilize learning theory in their description, and therefore find theory useful for making sense of classroom interactions. Utilizing theory as a mediator in examining classroom interactions is a precondition for students to enact professional vision. Students were placed into small groups where the physical arrangement and shared video were intended to foster collaboration. The first day was largely individual work as students constructed the initial version of their artifacts; we focus here on the second day in which there was intentional peer collaboration. In the second session, the video annotation tool, instructor prompt, and nature of collaboration mediated students’ activity and revisions in all three activity systems— researchers’ intended, group A’s enacted, and group D’s enacted. These mediators did however function in different ways in the three activity systems. Three key differences stand out across these three activity systems: (a) the constructed object, (b) the role of the Etherpad document, and (c) the resultant multimedia artifacts.

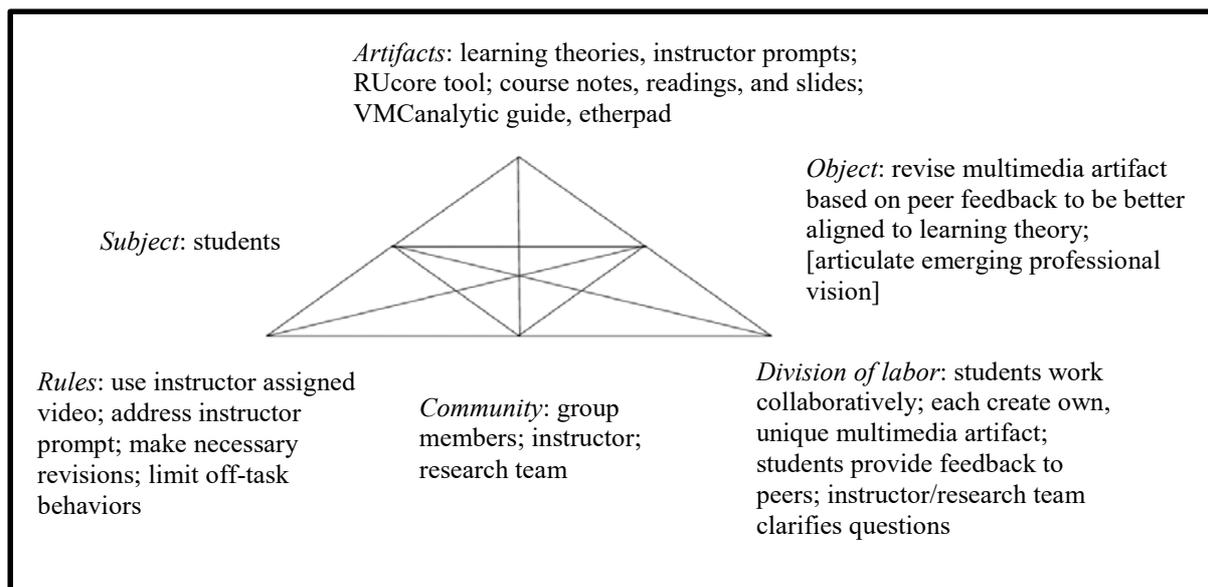


Figure 1. Intended activity system for class session two.

In group A, the students' object seemed to be to finish their multimedia artifacts and to document review of each other's artifacts in the group Etherpad. The nature of Group A's collaboration was not as rich and interactive as the designers had intended, but their activity was neither disruptive nor inappropriate according to norms of this particular classroom. With the object of completing the task, the group's key mediators were the instructor prompt and course slides, and less so learning theory. In group D, the students' object seemed to be to finish their multimedia artifact, but also to provide peers with meaningful feedback, and address this feedback in revisions to their multimedia artifacts. The students in group D were not as attentive to taking notes in their Etherpad, and instead engaged in conversation debating specifics of theories and concepts. Learning theory was a key mediator in their activity. Group discussions led two of the students, Isaac and Addison, to refer to their course notes and the instructor's course slides. While the richness of their collaboration was not entirely captured in their Etherpad, all three students referred back to the notes as they revised their multimedia artifacts. While we saw that students were not directly oriented to articulating their emerging professional vision, our conjecture that due to the assignment requirements, students would apply learning theory and therefore make claims rooted in theory concerning teaching practice providing evidence of their emerging professional vision was supported. Further, we believed that the differing nature of collaboration would have a direct impact on students' professional vision in their revised artifacts.

In order to gain insight into how the differing activity systems affected students' revisions to, and final multimedia artifacts, a qualitative coding scheme was developed to trace professional vision. The multimedia artifact version histories for the five students were examined to get a sense of the levels at which students were able to connect theory to practice before and after collaboration. The first and final versions of each event were coded using a coding scheme adapted from Sherin & van Es's (2002; 2009) teacher professional vision scoring rubric. Event versions were coded as (a) atheoretical noticing (b) theory-based noticing without interpretation (c) theory-based noticing and interpretation, or (d) theory-based noticing, interpretation, and response. Table 1 explains the code criterion and how the codes build on one another. In some events students merely mentioned a learning theory without defining it in any way or articulating a connection between it and the rest of their annotation. Even though these annotations technically named a theory, such events were classified as atheoretical noticing. Examining the multimedia artifacts through the professional vision codes allowed us to better understand how students' professional vision and noticing changed (if at all) after collaboration with their peers. By allowing us to trace the impact of collaboration on students' multimedia artifacts, these professional vision codes supported our activity theory analysis by providing a means of comparing professional vision across students, events, and revisions. Together, our activity theory analysis and professional vision coding provide a picture of emergent professional vision for two student groups—highlighting improvements in theory-practice connections and differences in the groups' objects of activity and use of mediators.

Table 1: Professional vision codes

Code	Description	Example excerpts
Atheoretical noticing	Annotations only explain what happened in the event, but do not connect the actions of the students and/or teachers to theory.	“In this clip, two students work to solve a fraction problem. The instructor comes to the table and give (sic) them feedback. The feedback the teacher gives is bad because the students do not seem to understand it. The teacher walks away from the table.”
Theory-based (T-b) noticing only	Annotations include references to learned theories and concepts, and provides explicit support from the video for these references. Implications for teaching and learning are addressed at the level of description and some evaluation, but do not reach interpretation of teaching practice.	“Meredith is brought in the center of the screen during this clip. Interestingly enough, though, she finds herself not raising her hand when the teacher asks if it’s (sic) okay for two/tenths to be called one/fifth...It relates to Information Processing because she is modeling the information that she believes is true. Although she was the first student to prove that two/tenths is equal to one/fifth, when she sees another example she challenges her understanding because the example given by the two boys before her is not the same. This can also be Information Processing Theory because Meredith may have a schema that there may only ever be one correct answer, or way, to show that something is the same.”
Theory-based (T-b) noticing and interpretation	Annotations provide a description of the important features of the interaction, address a learning theory or concept, ties this theory to explicit video evidence, and evaluate this evidence as it relates to teaching practice (i.e. this was “good” because...)	“Here, the two girls seem to divide themselves into separate roles with relatively little communication. [Student name] takes each of the stacks and determines the duplicates which she then gives to [Student name] to deconstruct. One of the elements within Sociocultural theory is the Activity Triangle diagram. One of the triangles in this diagram is focused on division of labor; who is responsible for what. Due to the clarity of each students (sic) distinct role, this theory would confirm that this is a beneficial strategy.”
Theory-based (T-b) noticing, interpretation, and response	Annotations include description of classroom actions, evaluation, and interpretation. Authors introduce a concept or theory, tie it to explicit video evidence, and interpret this evidence to make clear suggestions about how it influences their work as future teachers.	“...The girl in this video has a misconception as to what 14-5 is because she believes it is 8, but the middle boy believes it is 9 (he is right). Although he is right, he seems to be second guessing himself and modeling after the girl. They then work together at the end of this clip, all doing the same thing, using the cubes. This relates to the cognitive theory, and the teacher must address the problem so that Jamie builds correct schema for future math.”

As the coding examples show, not all students applied learning theory accurately, despite making theory based observations and interpretations. Because of this, an additional set of codes was then developed and used to address how appropriately the learning theories were used in the multimedia artifacts. The first and final versions of each event were then coded for application of learning theory using four codes: (a) no/unclear use of theory (b) inaccurate or inappropriate use of theory, (c) approaching appropriate use of theory, and (d) appropriate use of theory. Finally, the authors noted instances when post-collaboration event versions included new theoretical concepts, extended interpretations, or included evaluative statements (i.e. “I think this is bad for learning ...”). Together these two coding schemes and our observations of post-collaboration additions allowed us to explore changes in students’ use of learning theories and articulation of professional vision.

Results

While the pre-service teachers did not take up the designed activities exactly as the research team had intended, all students submitted satisfactory multimedia artifacts. The instructor and research group’s intended object of activity was to help students further connect learning theories to actual teaching practice, but both groups analyzed were more oriented towards directly fulfilling assignment requirements than understanding and applying learning theories. However, only one post-collaboration event was categorized as atheoretical noticing; all other events reached at least t-b noticing. Four students’ (Isaac, Riley, Natalie, and Isabella) post-collaboration modifications resulted in a change in their articulation of professional vision. Isaac and Riley both moved from atheoretical noticing to t-b noticing in two of their event descriptions. Because professional vision is develops over time and through practice, it is not surprising that few students’ articulation of their professional vision changed throughout

this one-week intervention. Yet, evidence of post-collaboration changes in students' multimedia artifact descriptions suggests that some students benefited more than others from discussing their own and each other's multimedia artifacts.

Four of the six students achieved t-b noticing + interpretation in at least one of their final events. This improvement is meaningful given the brief nature of the intervention. As students prepare for student teaching and the challenges of everyday classroom practice, they must notice, interpret, and respond. In these students' multimedia artifact construction, we see progress towards this long-term objective on a short time-scale. Moving from t-b noticing + interpretation to t-b noticing + interpretation + response requires pre-service teachers to articulate how they would handle this particular episode in their own classrooms or how the teacher in the video could have improved his/her response. Overall, students' multimedia artifact descriptions improved post-collaboration for students in both groups. Students increased the number of theories referenced in their final versions of events and articulated more appropriate uses of learning theory suggesting the RUanalytic tool and the activity design constructed a useful space for students to practice making theory-practice connections.

In both groups, two of the three students improved their articulation of professional vision after collaboration in at least one event. These shifts were supported by the detailed feedback provided in the group Etherpad documents, and suggest that discussing multimedia artifacts with peers typically had a positive impact on articulation of professional vision and use of learning theory. Isaac's first version of event 3 was categorized as no/unclear use of theory; he writes, "...Although [Meredith] was one of the students to prove that two-tenths is equal to one/fifth, when she sees another examples she challenges her understanding because the example given by the two boys before her is not the same. It's almost like she is changing her mind, not because she wrong, but because she trying to grasp a different understanding." In reviewing this event description Riley and Addison write, "Clip 3: I didn't notice this at first but LOVE that you caught it and explained it. Add modeling to the end. Great use of imformation [*sic*] processing! Misconceptions are also what she is struggling with." While Isaac did not state that he was using information processing his two group members assume he is drawing upon something having to do with this theory. Isaac took up this feedback and revised his event description to explicitly define and use information processing to articulate Meredith's struggle as a misconception that there is only one way to show two fractions are equivalent. Isaac's final event 3 was classified as appropriate use of learning theory and t-b noticing.

In collaboration, small groups were able to draw connections from interactions in their video clips to course theories, and some students were able to connect these noticings to the kinds of responses they would take as teachers. While reviewing each other's multimedia artifacts, groups re-watched certain events repeatedly to reach a consensus on what was occurring in the interaction. Furthermore, they encouraged each other to look beyond their first impressions of which theories applied to an event—seeing that multiple theoretical concepts could be applied to just one interaction. This kind of connection from theory to practice was the primary learning outcome intended by the research team, and seems to have been reached by all participants in these two groups, though there was room for improvement in how theory was applied by most students. In each of Isabella and Addison's final event descriptions their use of theory was categorized as either appropriate use or approaching appropriate use of learning theory. This means both pre-service teachers had few, if any, misconceptions concerning the learning theories each chose to support her observations and interpretations. While Riley, Isaac, and Natalie made many modifications to their use of learning theory post-collaboration, they each had one final event that was categorized as inappropriate or inaccurate use of learning theory.

Presenting and discussing their multimedia artifacts in their groups did improve all students' definitions of concepts and overall application of learning theory, but some students still struggled to define or apply theories accurately despite discussing them during collaboration. For the students with events classified as inaccurate use of theory post-collaboration, either the Etherpad was not employed as an artifact to guide their revisions, or the nuances of the theory discussed were not taken up by the student or effectively captured in the Etherpad. This is visible in the changes Isaac made to his event 1 description. An excerpt from Isaac's first version of this event read as follows, "...This thought process could be relevant when speaking about the Zone of Proximal Development in Social Cultural Theory. // Meredith is trying to understand the concept of fractions before sharing it with her classmates. This enhances her ZPD because her thoughts go from a personal learning experience to a whole class learning experience." Upon review of Isaac's work, Riley and Addison provided him with the following feedback regarding the first event in their Etherpad, "Clip 1: Add modeling to the begining [*sic*] of the clip where Meredith is modeling her way of figuring out the problem. Add what ZPD means. Did you include the concepts of building new schemas in clip 1?" Here, Riley and Addison suggested adding additional theoretical connections to an event and modifying use of theory to be more articulate. Review of the video recordings of this group's interactions revealed that Isaac and Addison returned to their class notes to clarify their understanding of information processing. Ideally, Isaac would have also reviewed his notes from the unit on sociocultural theories

as well and recognized that his use of ZPD was not aligned with how Vygotsky defined it. Examination of Isaac's final event descriptions suggested that he took up the suggestion to define ZPD and ignored the recommendation of adding concepts from other theories. Isaac's final version had an additional sentence (where the double slashes are in the event description above), which read, "The Zone of Proximal Development is the area and environment that a student feels comfortable with." But because Isaac's definition of ZPD was not accurate, his application of this theoretical concept remained an inappropriate use of theory. While not accurate, Isaac utilizes his understanding of ZPD as some sort of movement from a personal to a social learning experience to understand why Meredith first works independently before offering her solution to the rest of the class. Isaac's event one description is an instance of t-b noticing that is based on an inaccurate understanding of learning theory. The struggle to appropriately employ learning theory is not surprising considering this is a well-documented challenge in teacher education. Further, students' use of the RUanalytic tool was limited to one week of activities. In future iterations we intend to integrate the tool throughout the semester to provide greater opportunities for connecting theory to practice.

In analyzing the interactions and artifacts created by the pre-service teachers in this educational psychology course, we found that multimedia artifact creation allowed students to (a) parse student-teacher and peer interactions that are typically overlooked when just watching a video of classroom activity or observing a classroom in person, (b) relate aspects of different learning theories to concrete moments of teaching and learning, (c) solidify or reify their beliefs on how teaching and learning should unfold, and (d) review course materials and notes productively. Many students recommended greater integration of the tool into course activities across the semester because they found it useful in making sense of the learning theories and connecting theory to practice. The course instructor also believed this was a useful tool for similar reasons.

Significance

Through analysis of two groups of pre-service teachers' collaborative activity systems and their descriptions of classroom videos, we gain insight into how to support future teachers to think differently about theory and the ways it can play out in the classroom. Our activity theory analysis reveals that although students did not take up prompts in the way the instructor/research team intended; students constructed activity systems that were productive for their own objects of completing the assignment. These results suggest that students may not view theory as necessary to interpreting classroom interactions, and therefore did not appropriate the intended object due to a lack of relevance. Activity theory was a productive lens for understanding how students take up curriculum. Although group A's collaboration was more mechanical than desired, both groups were productive in connecting theory to practice, highlighting the value of integrating video of classrooms into course activities.

Teaching is a practice-centered career. As pre-service teachers gain experience refining their own teaching practice (through student teaching and extensive classroom observation), they fine-tune important skills. When pre-service teachers engage in teaching practice early on, the application of theory is often lost. As first through third year pre-service teachers, it is not surprising that these students' professional vision did not improve dramatically across one week. What is impressive is that all students made theory-based noticings. The activity of annotating real classroom video allowed students to demonstrate their ability to articulate theory-practice connections. Our results show that constructing a multimedia artifact provides the opportunity for pre-service teachers to see how theory can be part of their professional vision, but that they may also need to engage in collaborative discussions to make the most out of these opportunities, at least initially.

Together, the findings suggest that integrating more multimedia activities in pre-service teaching coursework has the potential to greatly improve students' ability to contextualize course concepts and more directly support their emerging professional vision. Interacting with the VMC repository and RUanalytic tool throughout an entire course, rather than in a week-long intervention, may allow students to more closely appropriate the intended activity system of using a video annotation tool to mediate evidence-based use of theory and development of professional vision. In future work, we hope to examine how video annotation practices might be integrated throughout pre-service teachers' undergraduate careers—moving from the annotation of VMC video to the annotation of peer teaching and eventually of their own teaching practice. As pre-service teachers move through course and practicum experiences, they must learn to meaningfully observe students and appropriately respond (Ball & Forzani, 2009; Sherin, 2001). Observation in this capacity is a complex skill that involves fine-tuned attention, disciplinary knowledge, and an applied understanding of theory (Eberbach & Crowley, 2009). Video annotation in general, and the RUanalytic tool in particular, might be leveraged as a form of scaffolded training for pre-service teachers as they develop professional vision.

References

- Agnew, G., Mills, C. M., & Maher, C. A. (2010, January). VMCAnalytic: Developing a collaborative video analysis tool for education faculty and practicing educators. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (pp. 1-10). IEEE.
- Allsopp, D. H., DeMarie, D., Alvarez-McHatton, P., & Doone, E. (2006). Bridging the gap between theory and practice: Connecting courses with field experiences. *Teacher Education Quarterly*, 19–35.
- Ball, D.L., & Forzani, F.M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, 60(5), 497-511.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, 33(8), 3-15.
- Cole, M. (1996). Putting culture in the middle. In *Cultural Psychology: A once and future discipline* (pp. 116-145). Cambridge, MA: Blknpa Press of Harvard University Press.
- Eberbach, C., & Crowley, K. (2009). From everyday to scientific: How children learn to observe the biological world. *Review of Educational Research*, 79, 39-68.
- edTPA. (2015). *About edTPA: Overview*. Retrieved from <http://edtpa.aacte.org/about-edtpa>
- Engeström, Y. (1987/2014) *Learning by expanding: An activity-theoretical approach to developmental research* (2nd ed.). New York, NY: Cambridge University Press.
- Goodwin, C. (1994). Professional vision. *American anthropologist*, 96(3), 606-633.
- Gomoll, A., Rehak, A. M., Novak, W., Andrade, A., Saleh, A., Hmelo-Silver, C., & Danish, J. A. (2016). Paper to be presented at the American Education Research Association National Conference in Washington DC.
- Haikkarainen, K., Paavola, S., Kangas, K., Seitamaa-Hakkarainen, P. (2013). Sociocultural perspectives on collaborative learning: Toward collaborative knowledge creation. In C.E. Hmelo-Silver, C.A. Chinn, C.K. Chan, & A. O'Donnell (Eds.), *International handbook of collaborative learning* (pp. 57-73). New York: Routledge.
- Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 169-202.
- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2005). *Student success in college: Creating conditions that matter*. San Francisco, CA: Jossey-Bass.
- Kunzman, R. (2014). "It could be so many learning theories": Multimedia artifacts and professional vision in pre-service teachers. IUB teacher education: In pursuit of coherence, consistency, and quality. Bloomington, IN: School of Education, Indiana University.
- Maher, C. A., Palius, M. F., Maher, J. A., Hmelo-Silver, C. E., & Sigley, R. (2014). Teachers can learn to attend to students' reasoning using videos as a tool. *Issues in Teacher Education*, 23(1), 31-47.
- O'Donnell & Hmelo-Silver, (2013). Introduction what is collaborative learning?:An overview. In C.E. Hmelo-Silver, C.A. Chinn, C.K. Chan, & A. O'Donnell (Eds.), *International handbook of collaborative learning* (pp. 1-16). New York: Routledge.
- Rajala, A., & Sannino, A. (2015). Students' deviations from a learning task: An activity-theoretical analysis. *International Journal of Educational Research*, 70, 31–46.
- Roth, W. M. (2007). On mediation: Toward a cultural-historical understanding. *Theory Psychology*, 20, 641-665.
- Sherin, M. G. (2001). Developing a professional vision of classroom events. In T.Wood, B. S. Nelson, & J. Warfield (Eds.) *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 75-93). Hillsdale, NJ: Erlbaum.
- Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In Goldman, R., Pea, R., Denny, S. (Eds.) *Video research in the learning sciences* (pp. 383-395). New York: Routledge.
- Sherin, M., & van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of technology and teacher education*, 13(3), 475-491.
- Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. In Lampert, M., Blunk, M. (Eds.) *Talking mathematics in school: Studies of teaching and learning* (pp. 107-149). Cambridge, UK: Cambridge University Press.
- van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571-596.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Talking Back to the Future: Anatomy of Reflection as Collective Practice

Alfredo Jornet, University of Oslo, a.g.jornet@iped.uio.no

Abstract: Reflection is recognized as crucial for deep learning to occur; yet how it is involved in learning is less understood and often taken for granted as an individual cognitive process. Building on Vygotsky's methodological insights on *unit analysis*, I elaborate on an approach that studies reflection as a collective, social and material phenomenon. I define the *joint production of accounts of prior experiences* as the unit of social interaction that contains all the essential features of anything we may ascribe to reflection as a cognitive process. Drawing from empirical (interaction) analyses conducted as part of a larger multi-data project on debriefing and retrospective situations, I describe *constitutive* features of reflection as collective practice. I show how this kind of analysis contributes to an explanation of reflection as learning process in its own right, and not just as a factor that "affects" learning. I conclude discussing theoretical and practical implications.

Keywords: accountability, conversation, reflection, unit analysis, Vygotsky

Introduction

Research in the learning sciences has consistently shown that reflecting upon prior experiences is a crucial feature of how people learn and a requisite for deep learning to occur (Sawyer, 2006). Students who engage in reflective tasks significantly outperform students who do not (Etkina et al., 2010) and specific forms of prompting learners to reflect upon their performance have been found to positively impact learning outcomes (Davis, 2003). Less is known, however, about *the mechanisms* involved in these outcomes. Most often, reflection is treated as an individual mental process without further critical analysis. Yet, if reflection is a private process that goes on within the individual mind, how can researchers and educators unproblematically identify instances of reflection? More importantly, if reflection cannot be directly observed, how can learners ever learn it in the first place?

The purpose of this paper is to examine reflection as a process that is publicly available for those already competent to recognize it, and for those not yet competent to engage in reflective practice even if they had never done so before. This study thus aligns with recent calls in the learning sciences to consider the larger communicational space constituted by several participants as the minimal unit of analysis (Greeno & Engeström, 2014). Although assumptions that learning is social have become commonplace, we do not yet have a conceptualization of reflection as collective practice and its implications for learning research and practice. Taking a Vygotskian perspective, and drawing from experiences gained during ongoing research on debriefing and retrospective situations, in this paper I contribute to a methodologically and empirically sound conceptualization of reflection as collective practice.

Reflection as collective practice

In much research about learning, the term *reflection* is used unproblematically to designate particular designed or observed activities, without necessarily engaging in a critical discussion of the concept or the processes it denotes. Thus, it is not difficult to find instances where the term reflection is used to define the very process that it is supposed to describe. This happens when researchers define "scaffolding ... reflection" as "helping learners ... *reflect* on their solutions in ways productive for learning" (Reiser & Tabak, 2014, p. 48, emphasis added), or when metacognition is defined as "the ability to ... *reflect* on what and how one learns" (Nathan & Sawyer, 2014, p. 32, emphasis added), only to subsequently find that reflection and metacognition come to be treated interchangeably in several other parts of the same volume. When conceptualizations are provided, these tend to be articulated in terms of individual cognition. Thus, Piagetian notions of *reflective abstraction*, which describe reflection as the learners' formation of internal (mental) representations of their own activity, are present in the literature along with emphases on the need to consider social and situated aspects (Lobato et al., 2012). A recent review shows that reflection has mostly been studied with regard to different types of knowledge and measures of individual variables such as motivation, beliefs, expectations, or self-regulation (Winne & Azevedo, 2014).

Though often conceptualized as individual competence, the fact that learning scientists are able to use *reflection* unproblematically already suggests that, whatever is designated by the term, it is objectively present as a witnessable feature of the situations and designs being so named. In fact, and perhaps not surprisingly, reflection is most often characterized in terms of tangible public actions. It is, for example, described as the learners' ability "to pause and evaluate their progress" (Nathan & Sawyer, 2014, p. 32), and as enabling learners "to compare their own problem-solving processes with those of an expert, another student, [or] an internal cognitive model of expertise" (Collins & Kapur, 2014, p. 114). A similar description was also given by American pragmatist John Dewey, who defined reflection as "the postponement of immediate action," a "stoppage of the immediate manifestation of impulse until that impulse has been brought into connection with other possible tendencies to action" (Dewey, 1938/1997, p. 64). Unlike the cases cited above, however, the philosopher is clear that these actions are not individual or mental. Instead, they comprise "a co-operative enterprise" that constitutes a form of "social intelligence" (p. 72). Rather than internal mental process, reflection consists of observable, public social actions that lead to the possibility of turning (prior, current) experience into the object of individuals' experiences. A methodological implication, therefore, is that reflection does not have to be assumed but can be directly studied in and as conversational process.

Conceptualizations of reflection as involving conversation exist in the literature. Donald Schön's (1983) work *The Reflective Practitioner* emphasizes that professionals shape work situations "*in conversation*" (p. 73) with them, a process he coins as *back-talk*. More recently, a view has been advanced that discourses' *recursivity*, language's property of *being about* or *referring to itself*, is foundational to individuals' competence to think and reflect. Accordingly, reflection could be thought of as a form of communication, a *commognitive act* (Sfard, 2008). However, the few studies that focus on reflection as social phenomenon still maintain the duality between the social and the individual, where, even if the role of community practices is acknowledged, "it is the individuals who reflexively monitor and discipline themselves" (Ovens & Tinning, 2009, p. 1126).

Reflection as collective practice is consistent with Vygotsky's thesis that "every function appears on the scene twice in the child's cultural development, i.e., on two levels, first the social, and then the psychological, first between people as an interpsychological category, and then within the child" (Vygotsky, 1989, p. 58). Vygotsky further cites Piaget as one of the first scholars "to confirm the thesis that thinking in preschool children does not appear before the debate appears in their collective" (Vygotsky, vol. 3, p. 95). The thesis that cognitive functions first exist as social relations has been taken up in the literature, a growing body of studies having empirically substantiated and validated it (Mercer, 2008). Accordingly, researchers have begun to investigate reasoning as a feature of the collective that exists in conversational forms (Michaels, O'Connor, & Resnick, 2008). Yet, here too the focus tends to be on finding *what works*, with less attention given to conceptual considerations of reflection as learning *practice* in its own right, and the implications to research and practice. If all possibility for individual reflection exists first *as* a social relation, a study of the collective achievement of reflecting bears not only promise to improve our understanding of how to help learners become more reflective, but also makes empirically available the mechanisms by means of which we come to learn from prior experiences—a foundational task in the learning sciences.

Methodological approach

For the purpose of developing a functional and empirically grounded conceptualization of reflection as collective practice, in the remainder of this paper I present findings derived from an ongoing multi-data project on debriefing and retrospective practices. The research, itself a collective endeavor, involves a variety of data sets and the development of a particular methodology inspired by ethnomethodology and cultural-historical psychology. Central to this methodology is the premise that cognitive functions exist first in the form of social relations, that is, as *talk-in-interaction*. However, not any stretch of talk-in-interaction constitutes an instance of reflection. It is necessary to first establish a *unit* that captures the basic features of the cultural phenomenon as a whole.

Joint production of accounts of prior experience as the unit of analysis

The research reported here follows a methodology based on Vygotsky (1987), who distinguishes between *analysis by elements* and *unit analysis*. Analysis by elements proceeds "with the decomposition of the complex mental whole into its elements" (p. 35). This treatment is exemplified with the chemical analysis of water, which takes the elements hydrogen and oxygen as the minimal units. The elements, however, "lack the characteristics inherent in the whole" (p. 35). For this reason, "the man who resorts to the decomposition of water into hydrogen and oxygen in his search for a scientific explanation of the characteristics of water, its capacity to extinguish fire ... will discover that hydrogen burns and oxygen sustains combustion" (p. 35). This form of analysis, which tends to rely on inductive approaches, reduces complex dynamic phenomena into

abstract categories or variables (e.g. motivational, social), thereby missing the integrity of psychological phenomena as living wholes.

By contrast, unit analysis aims to preserve the details and richness of the empirical whole under analysis, while still being able to derive general laws. The term unit “designates a product of analysis *that possesses all the basic characteristics of the whole*” (Vygotsky, 1987, p. 46). Studying the relation between thinking and speech, Vygotsky found such a unit in *word meaning*. In word meaning, both thinking and speech are united objectively and intrinsically: “Word meaning is a phenomenon of thinking only to the extent that thought is connected with the word and embodied in it” (p. 244). It is in the *concrete act* of speaking that thinking and word are united. Accordingly, unit analysis aims to find analytical objects that are, at once, both an immediate situation and the whole basis of the cultural phenomenon. In this regard, as Vygotsky (1971) notes paraphrasing K. Marx, unit analysis constitutes a form of *microscopic anatomy*, where general constitutive features are empirically found in their real and dynamic connection.

To achieve a functional anatomy of reflection as collective practice, the question needs to be asked, *what is the social unit that retains the whole of what reflection can be as a cognitive learning process in its own right?* Dewey defines reflection as stoppage of action to connect whatever has been done with other possible tendencies to action. If any higher psychological function *was* first a social relation (Vygotsky, 1989), then the social situation where an account of a prior event or experience is demanded and produced in conversation for the purpose of re-orienting action is that unit that exhibits all the properties of reflection as a whole. In such situations, reflection is not achieved individually, but is jointly and publicly accomplished. Individuals hold and are held accountable to what has happened, and it is then when one’s own experience can become the object of one’s own experience. Rational accountability has in fact been identified as a crucial aspect for learning (Suchman, 2003) and discussed as a desirable goal for learning (Greeno & Engeström, 2014). The pertinence of examining cognitive phenomena as conversational accomplishments is also well established, some researchers admitting that, to study cognitive (reasoning) processes, “it [is] necessary to take into account structures of conversation and then attempt to detect the logic within them” (Michaels et al., 2008, p. 287). These authors have identified *accountable talk* as crucial for productive learning, and have described three dimensions as normative models to develop in instruction. The research presented here expands on these findings empirically and methodologically by focusing on and describing reflection as a social, public process, and revealing not just its social nature, but also its nature as teaching/learning mechanism.

Data and participants

The research reported in this paper is part of a larger project investigating the social, material, and developmental aspects of debriefing practices designed to learn from prior experience in school and workplace settings. For the purpose of presentation, here I use excerpts from a database on science education. The database consists of 220 hours of video and other ethnographic materials recorded during two consecutive iterations of a design-based technology-rich research project aimed at connecting inquiry activities in the school lab, a science museum, and the classroom. As part of the intervention, inquiry activities alternate with sum-up sessions in which recordings and other products of inquiry are debriefed in plenum. A total of 64 students from an upper-secondary school and their science teacher participated in the studies, which were run across 2 academic years.

Analyses

Analyses are performed in individual and collective data sessions employing *interaction analysis* (Jordan and Henderson 1995). Collective data sessions include scholars and practitioners with experience in the learning sciences, computer sciences, engineering, and other domains of relevance depending on the nature of the data. The analyses aim to capture and reproduce the integral dynamics of social events, which involves investigating the turn-taking and tool-facilitated conversational practices of the participants. The analysts do not attempt to interpret what the participants’ utterances mean or what they individually think, but follow the means and consequences of social actions as these become relevant to the participants. Observations should lead to what has been termed *ethnographically adequate descriptions*, where “the ethnographer’s adequate account of what natives do together must follow from the way in which the natives structure a situation to allow their participation” (McDermott, Gospodinoff, & Aron, 1978, p. 246). In this way, and taking the joint production of accounts of prior experiences as unit, the form and function of reflection is examined as it is endogenously produced in practice. The resulting descriptions are therefore rigorous with respect to the empirical details but, in that they describe cultural possibilities of social interaction, they also have general character. Transcripts follow Jeffersonian transcription conventions, the level of detail being adapted to the corresponding analyses.

Findings: Anatomy of reflection

In this section, I report on three general moments of reflection as collective practice that have emerged in the course of investigations on debriefing and retrospective situations. These moments are not detachable from each other and describe general features that exist united not abstractly but contingently, in actual social situations. Because the methods do not fall on either side of the inductive/deductive dichotomy, the presentation retains both conceptual and empirical aspects. The goal is to summarize, in the space available, phenomenal properties of reflection as collective practice that might be useful for guiding future research and practice.

Jointly stopping/troubling action

A first important moment in reflection concerns, as thematized in Dewey's definition, stoppage of what otherwise may be described as *absorbed coping* (Dreyfus 2002). In absorbed coping, the person is so familiar with the means-consequence relations that her actions go on seamlessly, without having to *think about* them. To reflect involves stepping out of that absorbed coping and make salient otherwise undifferentiated aspects of what one is doing. That *doing* may be one that is now happening and is interrupted, or may be a happening that took place in the past. In either case, there must be work to make the situation such that whatever has been done needs to be looked upon again.

Consider the following excerpt, where a group of upper-secondary students are experimenting with a spray can of compressed air as part of their inquiry-based curriculum on energy, which includes computer prompts to help them reflect.

Excerpt 1

- 01 S1: ARG:: ((shakes hand)) look at my skin, ((laughing)) it gets cold.
it gets freaking cold.
02 S2: let me feel it ((touches base of the spray can))
03 S1: it becomes freaking cold. just feel it (.) feel up there, ((touches
upper part)) it becomes cold on the top, kind of.

In the excerpt, a student (S1) produces interjections after having pressed the valve of the spray can, releasing the air and drastically decreasing the can's temperature due to (a) a drop in the pressure, and (b) the associated phase transition from liquid to gas, which draws from the heat around. This scientific account however is not available to the students; such account is precisely what the students are supposed to learn. A second student (S2) then requests to "feel it" and S1 instructs him where to touch to feel what he felt. The students' talk is mostly about their sensuous experience and is closely connected to the material situation by means of indexical gestures and touching. At stake is the making present of a presence, which can be equally accessed by the students given they provide each other tutorials on how to touch so as to feel it in specific ways. Although the students already *talk about* the event, the nature of their actions is such that actions go on seamlessly. This, however, is about to change as the students read the computer-provided prompt, "feel, observe and discuss," and wonder "what happened?"

Excerpt 2:

- 19 S1: I don't know what happened. I think it happened down there and
suddenly it happened something cold up there
20 S3: no, there is much pressured air inside, which makes it to come out
at once
21 S2: pressure makes it to become cold, perhaps, I don't know

An event that first is treated as unproblematically present is now accounted with doubt. Their descriptions now include qualifiers such as "I think" and "I don't know." Although the event is the same, its mode of orientation has completely changed: Something that was immediately apparent to the students now becomes uncertain. A description that still preserves sensuous features of the event no longer is accepted and leads to descriptions that resemble scientific talk (turns 20–21). This shift does not come immediately but requires (practical, social) work. Thus, the prompt is attended for the first time several turns prior to Excerpt 2, when S2 visibly orients gaze and body towards the screen where the prompt is displayed, reads the task "feel, observe, and discuss," and asks, "what happened?" as if anything the students had said or done so far were not adequate accounts. S1 then responds by inviting S2 to "try and press," which he does, leading to emphatic exchanges that S3 formulates as "funny" (turn 14), laughing as the students spray over each others' hands. There is a new visible orientation

towards the task's prompt, however, which re-states the question again, "what happened?" (turn 15). S2's repeated orientation to the task's prompt becomes visible bodily and verbally, standing out against the background of laughing as an invitation to return to the seriousness typically associated with science and science classrooms.

Excerpt 3:

- 12 S1: just try and press yourself.
(3.8) ((S2 sprays, all touch the cold can and laugh))
- 13 S1: LOOK AT MY SKIN- ((laughing)) LOOK At my hand, IT feels like-
(2.0) ((all laughing))
- 14 S3: SO funny,
- 15 S1: yeah=you are smart. well. ok. ((gazes towards prompt in screen))
what happened?

Jointly orienting prior events

Excerpts 1–3 exhibit how an initially unproblematic situation comes to stand as problematic and as requiring revision. But mere stoppage of action is not yet reflection. The latter requires that the prior event becomes not just salient, but does so *in a new light*. Reflection as collective practice, thus, is not mere *retrieval* but requires *work to orient towards* and make a prior experience present again in a way in which it has not been experienced before. We already see part of that work in the excerpts above, where making salient certain perceptual features already brings a new orientation and form of accounting for the event. The first and second moments of reflection therefore are not sharply divided phases. Rather, they are interconnected *constitutive* moments that, together, begin to distinguish reflection from other possible functions, such as recollection.

The reader may better appreciate the work involved in making a prior event present again when considering what it takes to get a child to tell what they did earlier during the day. Parents with school-age children know the time and energy (work) that it takes to get them *orient to* and *talk about* what has happened at school in a way that the parents can accept as appropriate responses. In such situation, rather than some given or "natural" form of retrieval there is *instruction* on how a prior experience must be recollected. The work is joint (social, irreducible to the individual) precisely because requests need to be posed in such a way so as to provide with the necessary hints that will elicit adequate answers. The resulting recollection—the child's telling—then is not product of the child's mind but rather results from the social relation that leads to its production. We observe orienting work in the excerpt below, which takes place as part of the same curricular unit but during a next iteration two academic years later. Prior to this situation, the students have already conducted their hands-on tasks during a previous lesson, when they had produced digital videos of their observations and explanations of "what had happened," among other things, with the spray can of compressed air. In the current situation, the students' video-recorded accounts are being presented in plenum, where the teacher leads a reflective discussion concerning the students' accounts. There has been one video clip projected and the teacher picks up on the students' "last sentence" in the video, namely that "the temperature is higher in the room than in the can."

Excerpt 4

- 01 T: and THEN my QUESTION to you Is; (1.1) I::::S It, SO:. (1.5) that;
BASically; (1.4) the temperature is higher in the room than in the
can?
- 02 (3.8) ((complete silence first, then murmuring))

The teacher poses the question whether the student's observation corresponds to what actually had happened during the event. He produces the question in such a way that culturally competent listeners can hear it not just as a genuine request of information, where the inquirer asks in order to gain some information she currently does not have, but rather as questioning the adequacy of the students' account from the perspective of someone who already knows the answer. This hearing is made possible not just because of the teacher's institutional role but also in and through the teacher's emphasis in accenting, intonating, and lengthening parts of the question, as shown in the transcription conventions. The question is not responded to, and the teacher repeats the question, this time addressing a particular student in the classroom, namely the student that had produced the assertion in the displayed clip. Here we observe a work similar to that reported in excerpts 1–3, where S2 reformulates the question "what happened?" which marks whatever has been produced as inadequate response. In the current situation, the student being addressed responds in the negative, and the teacher asks,

“why not?” There is again a long silence, and the teacher reformulates once more, but this time changes also the lexical content of the question.

Excerpt 5

- 15 T: if the temperature would have been ... higher in the room than in the can, what should one have as well?
16 S4: nitrogen?
17 T: ye::a::?
18 S4: no=no=no, I mean, you know, liquid nitrogen-
19 T: yes; it is in a way almost the same as- almost the same as what we have here. it is liquid air. it's the same principle as fluid nitrogen.
20 S5: are you thinking of a gas now?

The teacher reformulates his question now so that what is requested concerns what should be observed if the account given had been so, namely that the temperature outside of the can was higher. One student responds “nitrogen,” which the teacher takes up in a way that invites repair (turns 17–18). The student then clarifies, he means “liquid nitrogen” and the teacher acknowledges that that is almost the same as what they “have here” (turn 19). Another student then asks what the teacher is thinking. In this sequence thus we observe that the work to achieve the production of the account of what happened does not fall to one single participant. The teacher is performing work to formulate his questions in ways that elicit adequate answers. But we also observe the students performing work to understand how they have to hear the question so that they can provide a successful answer, which here involves asking whether the adequate account has to do with “a gas”. The teacher then performs further work, which anchors the account to the event being reflected upon.

Excerpt 6

- 21 T: yea: is it so that-it- (.) when one has the gas container standing on the table; (.) ((*gestures as in figure*)) is the temperature higher inside the gas-can, than in the room then; ((*points from side to side*))
22 S: no;
23 T: no::hh;



The teacher acknowledges the student’s query, but offers a new reformulation. This time he makes present, by means of talk and bodily performing, the original situation in which the spray can lies on a table. He draws on the gestural set up to make salient the verbal distinction of “inside” and outside the can, which this time elicits an appropriate response, which he evaluates as such by emphatically repeating it (turn 23). Few lines later, the classroom will achieve a new account of the event, where it becomes articulated and visible to all that there was no temperature difference at the starting point. If there had been, the difference would have balanced somehow, allowing for further discussion on thermodynamics. Important to our analyses is the work that is performed so that a prior event is oriented to in particular ways, which here is not reduced to verbally eliciting one or other aspect of the event. If that had been enough, we would not have observed the repeated reformulations, which shift not only lexically but also prosodically and gesturally. Most centrally, the orienting work involves a form of *framing*, where what changes is not content but *context*. The event is not oriented to in and of itself precisely because it was not available in the form that would allow for learning to happen. To make the situation present again in a way so that learning something from it is possible, the event needs to be brought back so that it can be experienced again differently, where the crucial features that were missed become salient now.

Jointly orienting to the future

In the prior fragments, I show how reflecting on a (distantly or immediately) prior event involves first a stoppage of action, which requires work so as to make something salient and so no longer seamlessly tied to one’s own action; and also an orientation so that that which becomes salient is approached as part of some whole, as text within a context, that gives it a new sense that was not available before. This context is not achieved in the abstract workings of individual minds, but requires joint work. Thus, reflection, as practical

achievement, always involves a relation between what has been done and the what-for of ongoing action. This third moment is necessary in the constitution of reflection as a function that is different from recollection (in its direction to the past) or prediction (in its direction to the future). It is not the later because its function is to find out things that had not been found in the original event. But this orientation is not initially a property of the individual mind, but of collective activity, which always implies a motive (Leont'ev, 1978). It is this collective motive what, in and through participation, comes to shape whatever is refracted in the individual's experience [*perezhivanie*] (Vygotsky, 1994).

In the sequences described above, part of the work of making the prior event present and intelligible in a new light concerns looking for that frame of action within which a (a) different account can be produced and (b) be accepted as adequate to that frame of action. We have seen how that work is performed in and through conversation and the back and forth of requests and provision of accounts, a *field of expectations* emerges as a visible and accountable feature of the setting that now allows the students new forms of reflective participation. Thus, several turns into the lesson described in excerpts 4–6, the teacher will request one student to summarize the account of the spray can event that they have arrived at. As part of this account, the student states that, when the liquid air is released “[the can] takes in energy from the environment with high temperature.” The teacher then asks what heat is, and the students respond, “energy that moves,” which the teacher completes, “from higher temperature to lower temperature.” At that point, a student raises his hand and asks, “but is not the can at room temperature?” The former question is taken up by the teacher, who rather than responding, evaluates it positively, “good, good, good”, and throws the question back to the classroom. In this situation, there is not just a new opportunity for the whole classroom to further develop a scientific account of the event, but we observe how one student now marks as problematic and demands an account of what has just been said and done, precisely in the same manner that the teacher had done before for the students. A reflective function that was organized collectively is now observed in the individual, who exhibits new orientations and expectations towards the prior event and how it is to be accounted for in the present, and in the future.

Conclusions and implications

In this paper I discuss and examine reflection as collective practice. Although social aspects are often acknowledged in the literature, *what* is social about reflection has not been elaborated. Building on Vygotsky's unit analysis, I identify the joint production of accounts of prior experience as the (social) unit that exhibits all the properties of reflection as cognitive process. One influential constructivist analysis of reflection begins with a formal description of the operations that are involved in reflective situations such as when one, “having just eaten an apple, takes a bite out of a second one, and is asked which of the two tasted sweeter” (von Glasersfeld, 1991, p. 24). In his formal description, E. von Glasersfeld describes the following sequence: “the sensations that accompanied the eating of the first apple would have to be remembered... then they would have to be re-presented and compared ... with the sensations accompanying the later bite from the second apple” (p. 24). Glasersfeld, taking the individual mind as the point of departure, admits that his is only a hypothesis, as “we cannot observe how such a judgment is made” (p. 24). A similar approach is observed in contemporary literature, where analyses of reflection build on formal descriptions of the tasks, different typologies of being derived, e.g., with respect to the types of knowledge involved (Winne & Azevedo, 2014).

The formal analysis presented here is not based on formal abstractions or elements, but describes actual, irreducible social phenomena. Although formalisms, the three moments presented here describe concrete work that the participants make visible to each other and to the researchers alike. Unlike other (important and needed) studies that look at reflection in terms of “what works”, the present study focuses on features that are *constitutive* of reflection as a unitary and recognizable phenomenon, which distinguish it from other cognitive functions such as memory. The connection between the three moments is not externally imposed, but exist in as a real connection. Furthermore, whereas constructivist analyses assume that whatever is to be reflected upon (sweetness in the apples' example) must appear within the intellectual horizon of the individual, in the cases analyzed here I show how students achieve accounts of their prior experience that are beyond their current intellectual competence. Reflection thus is here observed not as an external factor, but as *a learning process in its own right*. Learners reflect not because they find in their mind what they need, but because the making present of prior experience develops in and through public, visible conversational relations. In *addressing* and making talk about prior experiences intelligible to each other, participants develop new expectations and ways of relating to their own actions. It is the public work of posing and answering questions what allows learners to participate in reflective practices about aspects of events that otherwise are not within their reach. This work involves not just talk but also the material transformation of the situation so as to become an intelligible *context*. Reflection thus emerges here not as an internal cognitive competence that somehow matures with age, but as a artifactual, *cultural* process through and through. The analyses presented here thus make salient aspects of

bodily and affective engagement that remain uncovered in current conceptualizations, and which may potentially expand our understanding of how to better design for reflective practices.

References

- Collins, A., & Kapur, M. (2014). Cognitive Apprenticeship. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences (2nd Edition)* (pp. 109–126). New York, NY: Cambridge University Press.
- Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences, 12*, 91–142.
- Dewey, J. (1997). *Experience and education*. New York: Touchstone. (Originally published in 1938)
- Dreyfus, H. L. (2002). Refocusing the question: Can there be skillful coping without propositional representations or brain representations? *Phenomenology and the Cognitive Sciences, 1*, 413–425.
- Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R., & Hmelo-Silver, C. E. (2010). Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories. *Journal of the Learning Sciences, 19*, 54–98.
- Greeno, J. G., & Engeström, Y. (2014). Learning in activity. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences (2nd Edition)* (pp. 128–147). New York, NY: Cambridge University Press.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences, 4*, 39–103.
- Leont'ev, A. N. (1978). *Activity, consciousness and personality*. Englewood Cliffs, NJ: Prentice Hall.
- Lobato, J., Rhodehamel, B., & Hohensee, C. (2012). 'Noticing' as an alternative transfer of learning process. *Journal of the Learning Sciences, 21*, 433–482.
- McDermott, R.P., Gospodinoff, K., & Aron, J. (1978). Criteria for an ethnographically adequate description of concerted activities and their contexts. *Semiotica, 24*, 245–275.
- Mercer, N. (2008). Talk and the development of reasoning and understanding. *Human Development, 51*, 90–100
- Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education, 27*, 283–297.
- Nathan, M. J., & Sawyer, R. K. (2014). Foundations of the learning sciences. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences (2nd Edition)* (pp. 21–42). New York, NY: Cambridge University Press.
- Ovens, A., & Tinning, R. (2009). Reflection as situated practice: A memory-work study of lived experience in teacher education. *Teaching and Teacher Education, 25*, 1125–1131.
- Reiser, B. J., & Tabak, I. (2014). Scaffolding. In R. K. Sawyer (ed.), *The Cambridge Handbook of the Learning Sciences (2nd Edition)* (pp. 44–62). New York: Cambridge University Press.
- Sawyer, R. K. (2006). Introduction: The new science of learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp.1–18). New York: Cambridge University Press.
- Schön, D. A. (1983). *The reflective practitioner*. New York: Basic Books.
- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. New York: Cambridge University Press.
- Suchman, L. (2003). Writing and reading: A response to comments on plans and situated actions. *Journal of the Learning Sciences, 12*, 299–306.
- von Glasersfeld, E. (1991). Abstraction, re-presentation, and reflection: An interpretation of experience and Piaget's approach. In L. P. Steffe (Ed.), *Epistemological foundations of mathematical experience* (pp. 45–67). New York: Springer.
- Vygotsky, L. S. (1971). *The psychology of art*. Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky, vol. 1: Problems of general psychology*. New York: Springer.
- Vygotsky, L. S. (1989). Concrete human psychology. *Soviet psychology, 27*(2), 53-77.
- Vygotsky, L. S. (1994). The problem of the environment. In R. van der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 338–354). Oxford, UK: Blackwell.
- Winne, P. H., & Azevedo, R. (2014). Metacognition. In R. K. Sawyer (ed.), *The Cambridge Handbook of the Learning Sciences (2nd Edition)* (pp. 63–87). New York: Cambridge University Press.

Acknowledgments

The research reported here is jointly funded by the Research Council of Norway and the Co-funding of Regional, National, and International Programs (COFUND) – Marie Skłodowska-Curie Actions under the EU Seventh Framework Programme (FP7), grant number 240246/F10. I thank Wolff-Michael Roth, Chris Campbell, and the anonymous reviewers for their critical comments and suggestions.

Becoming an “Expert”: Gendered Positioning, Praise, and Participation in an Activist Community

Joe Curnow, University of Toronto, joe.curnow@mail.utoronto.ca
Jody R. Chan, Teachers College, Columbia University, jrc2222@tc.columbia.edu

Abstract: This paper argues that gendered participation in a community of practice can create an illusion of expertise and learning based on masculine performances of expertise. Looking at a local campaign of the student climate change movement over the course of one year, we analyze participants’ interactions and examine how gendered positioning produces “experts”. We argue that these experts achieve their expertise not through learning or achieving mastery of the core competencies of the community of practice, but rather by performing dominant forms of masculinity, which are affirmed by other members of the community and recognized as authoritative. We argue that theorizations of situated learning must centre critiques of sociohistorical power relations in order to accommodate for such tendencies.

Introduction

Climate change is a pressing issue, disproportionately impacting low-income communities, communities of colour, and Indigenous peoples (Taylor, 1997). Given the gravity of the situation, we need critical and competent people who actively combat the root causes of climate change. The youth climate movement is actively engaged in this work and in developing the leadership and expertise of the next generation of activists. The experts the climate movement is cultivating believe they know the facts, are willing to act decisively, and speak assertively to stop climate change. The problem is, this group is almost exclusively made up of privileged White men (Goldenberg, 2014). Where are the rest of the experts?

This research asks how gendered forms of participation enable and constrain students’ learning and engagement in a climate activist campaign, such that men identify and become identified as experts while women do not (1). In our analysis we overwhelmingly found men were recognized as experts while women’s contributions were likely to be dismissed or fail to gain traction. We argue that in this community becoming an expert was deeply tied to behaving in typically-masculine ways. When certain members positioned themselves as authorities through engaging in exclusive talk and taking up space, others came to recognize them as experts. We focus on three forms of interaction: ideas being accepted without discussion, exclusive talk, and affirmation. Through these practices, we see expertise being conflated with dominant masculine forms of participation, making those modes of participation difficult, and sometimes impossible, for women to perform.

Our analysis shows that this process reinforced itself in an ongoing way. As dominant men presented themselves as experts, the rest of the group tended to recognize that “expertise”, which reinforced men’s views that they were, in fact, more experienced and more qualified. That sense of being more qualified lead them to engage other male participants in exclusive talk, garnering more recognition and blocking other people from participating. They tended to receive praise and encouragement, which we argue further entrenched men’s sense of expertise. At the same time, women’s contributions received far fewer instances of praise from fewer people, and few women were included in exclusive talk. Women’s attempts to contribute were less successful, and the dismissal of women’s contributions lead them to be less assertive and withdraw vocally and physically.

We argue that becoming an expert in a community of practice may have less to do with adopting shared practices or acquiring new skills, and more to do with performing masculinity, even in spaces that value social justice and anti-oppression. In the past, communities of practice and situated learning scholars have taken for granted that the learning process involved in becoming a full participant necessitates developing understandings of what the community does, why, and how. We argue that this may not always be the case, as participants may become full members based on their socialization and identity performances, regardless of their learning processes, just as other members can be left at the periphery. This raises an important question for situated learning theorists—how do we accommodate for social relations of inequity in communities of practice and how does that impact access to learning and full participation?

Situated learning and communities of practice in a gendered world

Situated learning theory understands learning as a social process of becoming part of a community through adopting its practices (Lave and Wenger 1991, Lave 1996, Contu and Willmott 2003). Learning is an active, ongoing process in which people move into different forms of participation by engaging with the activities, theories, and nuances of the group context. Communities of practice (Lave and Wenger 1991, Wenger 1998) are understood as groups of people engaged in collective meaning-making through joint work and mutual engagement with shared repertoires of action. Legitimate peripheral participation describes the process by which new members become proficient at skills and activities within a community of practice and become part of the community, moving from marginality towards mastery of the practices (Lave and Wenger 1991). Within situated learning theories, learning is considered becoming – becoming a type of person, a member of a community, a recognized expert (Lave 1996). This is about identity development at the individual level, but also about the social production of identity, where other participants in a community of practice support (or do not) and acknowledge one's process of development, and through that process, community identity is developed

Just as communities produce identities and are in turn shaped by the identities of their members, gender is socially produced in context. West and Zimmerman (1987) employ the notion of “doing gender” to explain the ways that gender is a socially produced performance rather than a fixed category. Gender is an historical social relation with consistent meaning under patriarchy, yet gender is produced in an ongoing way and is specific to different contexts and groupings – people perform gender in nuanced ways depending on the norms of the space they are in. Crenshaw (1991) argues that gender is always complicated by intersecting identities and social relations. People perform their genders differently depending not only on the space, but also their social location at the intersections of race, colonialism, dis/ability, class, and other relations. This is not to say that people are free to perform their gender(s) in any way they please. Gender is a collective performance and social relation, and gender policing and recognition can constrain people's modes of doing gender.

Communities of practice tend to reproduce dominant social relations (Curnow 2013). Although they can be subverted, gendered practices are often unintentionally part of the modes of participation, and create inequitable spaces within groups. Gender is performed in ways specific to the community, but embedded in the broader patriarchal context. Within a community of practice, if practices reflect the dominant gender relations and performances, some practices will be out of reach for people from different social locations (Hodges, 1998), and for women and trans people in particular. Salminen-Karlsson (2006), Curnow (2013), and Paechter (2003) examine gendered performances and the ways women access full participation in male-dominated communities of practice, arguing that if women are unable to perform or access certain gendered activities, their mastery of the performances and learning are complicated and often limited.

One of the ways people do gender is through positioning themselves and being positioned by others. Positioning refers to the ways people present themselves through their actions, uses of space, ways of speaking, and physical presentation, among other things, and how other people understand those actions and situate people within the community (Holland, Lachiotte, Skinner, and Cain, 2001). Gendered participation in a community of practice is one aspect of how one positions oneself and is positioned by others, and gendered practices are enacted through different modes of positioning.

We are interested in the confluence of these theoretical frameworks – how men doing gender in dominant patriarchal ways of performing masculinity position themselves, and in turn are positioned by others, as experts within a community of practice. We trace how men in an activist community of practice become “experts” through their positioning and examine what they learn, as well as what women in the group learn, as men's performances of expertise are situated as the full practices of the community, making them unattainable or at least uncomfortable for other-gendered people. Bringing these frameworks together shows how gendered performance is disparately valued. Masculine modes of participation are recognized as expert, rendering women's forms of participation less valuable and limiting their access to the full practices of the community.

This analysis troubles the view in situated learning theory that becoming recognized as a full participant in a position of mastery necessarily denotes learning, or vice versa. Situated learning emerged, in part, from studies of apprenticeship, where skill-based competencies were a core part of the practices necessary, in addition to the social performances of the community, in order to become a full participant. However, what is required to become recognized as a full participant may not reflect performances of particular skills or even adopting a shared frame of the philosophies or logic of the community, as has been suggested (Lave and Wenger, 1991). While we agree that learning occurs in communities and is part of the co-navigation of communities and activities, we offer a caveat that learning skills or shifting consciousness may not be required to become a full participant. As we

demonstrate below, participants learned to work together, and men learned, at some level, that it was not only possible, but also advantageous for them to perform dominant forms of masculinity. However, situated learning assumes people's learning will include competencies and performances tied to the relevant subject – in the case of an environmental activist group, the assumption would be that people would become masters through adopting practices of environmentalism, like strategizing, canvassing, or other core techniques of environmental activist communities. We argue that expertise was decoupled from these types of learning. Dominant male speakers moved into full participation not through their mastery of the full practices of the community related to their work, but through their ability to position themselves and perform dominant masculinity in the group. Learning did occur, but expert status and centrality was not available to every participant – women were prevented from becoming recognized as experts or full participants in the community. In order to address this, we must attend to power and sociocultural relations of inequity in order to mitigate the assumptions of expertise and centre instead on learning, rather than enactments of privilege.

Methods

This work is based on an ongoing participatory action research project that examines how student activists learn about race, colonialism, and patriarchy through their involvement in environmental activist campaigns. Data was collected with a Toronto-based campus climate action group. The group of young people came together to fight climate change, primarily through a local campaign that is internationally coordinated. They meet weekly and plan different strategies to engage students. Students are regularly involved in direct actions, social media campaigns, and social events to move forward an environmental activist agenda. Meetings had from 11-27 participants, representing all years in school, (1,2,3,4,5, Law, MA, PhD). Men and women students attended in roughly even numbers and there were no openly trans or other gendered students. Racial and ethnic make-up shifted over the course of the year, but the group remained predominantly White, even as Indigenous, Black, South-Asian, and East Asian students became increasingly involved. The group had three elected leadership positions, one filled by a White woman, the other two by White men. Decision-making was ostensibly done by what they termed “working consensus”, but in practice, decisions were not made in a consistent way.

These young people are all committed to the joint work of environmentalism and social justice, though they do not always share understandings of what social justice encompasses. The contradiction between the social justice goals, the rhetoric of the group, and the way their processes play out offers a window into how difficult it is for people to identify the dominant practices of their genders, but also the subsuming nature of the dominant practices of North American society vis a vis gender and patriarchy.

Data collection and analysis

Video was collected at fourteen meetings over the course of the academic year at weekly meetings, resulting in 2,009 minutes of video. Videos lasted from 60-180 minutes, depending on the meeting length. For this analysis, we include only the whole group meetings, and do not include the breakout sessions and report backs, limiting our data to 578 minutes of video. Videos are captured from one to four angles and stacked so that all four streams are visible and can be watched and coded simultaneously.

After video was collected, it was content-logged and pre-coded using preliminary codes based on the research question (including rough tags like race, gender, and colonialism). The first substantial analytic pass of coding was conducted by five women participants from the group, – two White, one South Asian, one East Asian, one Indigenous – and one Black man. We watched segments of videos from across the year together and coded “interesting” segments, asking the broad question of how gender is made salient in our group, discussing every instance someone raised and making extensive notes. After conducting “interesting” coding on three segments from the beginning, middle, and end of the year, we reviewed and consolidated them into codes that were most present in the video we reviewed and in our experiences of the group more broadly. The codes we developed included overlapping talk, exclusive talk, affirmation (vocal and gesture), uptake of ideas, positioning, and recognition. The codes were developed into sub-codes, which are described in the findings. All codes were refined by test coding 10 minutes of two videos and iteratively clarifying the codes.

After qualitatively coding, we embarked on a basic quantitative analysis as a way of demonstrating the patterned participation that participants identified. While there are significant limitations to quantifying interaction and speech acts (Schegloff, 1993), we attempted it as a way of convincing participants that gendered patterns existed and were important to acknowledge. Participants requested this data because they felt it could prove or disprove the claims women in the group were making. We asked participants how much of our total

meeting time we would need to analyze to create a compelling case for the gender-dynamics of the group. Those polled preferred randomly selected segments (to avoid “cherry picking” data) across all of the meetings, rather than extended segments from fewer meetings. To meet their requirements, we used a random number generator to select two segments of five minutes each from 14 videos from all regular weekly meetings and did detailed analysis of each. The samples represent 25% of the meeting video.

Results were compiled to create a basic statistical view of the trends. We compiled instances based on the segments, and ran two tailed t-tests to evaluate confidence in the comparisons we examined. Paired t-tests were selected as the most appropriate because of our small sample size, and their ability to evaluate the statistical significance of a difference between two groups or results. All t-tests, except for one, noted below, demonstrated statistical significance in the patterned difference between men’s and women’s participation. Paired t-tests demonstrate the existence of a pattern or difference but not its magnitude, so we provide sample means to express the average number of instances of a code, separated out by gender, within each 5 minute segment. Statistical analysis was complemented with ethnographic vignettes drawn from our coding to illustrate the trends. The initial analysis was circulated among the group who developed the codes for feedback, which was incorporated.

Findings

We look at gendered participation and positioning in three ways: the uptake of ideas, exclusive talk, and affirmation. These three aspects of group process show us different yet complementary things about how members of the community positioned themselves as experts and were positioned as experts by other members. Taken together, these show how members positioned themselves in the way they contributed their ideas and had those ideas adopted, sometimes without question. They also show how members positioned themselves as relevant contributors with expert information or opinions, which were jointly discussed by other experts and in the process reinforced the performance of expertise. Finally, they show how other members of the community of practice affirmed contributions, demonstrating that certain members were deserving of praise and encouragement. These three forms of participation demonstrate the processes of members positioning themselves, positioning others as co-experts, and being positioned by other members. The data highlights how these practices were highly gendered and reinforced the idea that men were experts deserving of the space they occupied in the group.

Adopting ideas

The first area we examined was whose ideas were discussed and adopted in the community. We believe this serves as an indicator of whose ideas were valued, whose contributions were centred in the group, and ultimately, who had the authority to set the agenda for the group. Through our coding process, it became clear we needed to attend to a specific form of ideas being taken up, namely the practice of ideas being adopted immediately and without discussion. This was defined as an instance of someone making a proposal for action, either around the internal process of the group or the external strategy, and having it be taken up by the next speaker as a given without further discussion. We argue that this is a distinct indicator of positioning oneself as an expert and of being positioned by the rest of the group as an expert.

Statistical analysis of our coded sample showed that ideas being adopted without discussion was, in fact, a highly gendered practice. There were 58 instances of this occurring in the sample from men participants, and 11 from women, with men’s ideas being five times more likely to be adopted without discussion than women’s. This is a meaningful act of positioning; it demonstrated how men were understood to be authorities, how men’s ideas drove the group’s process and strategy, and how men effectively decided what everyone would do. Most significantly, this shows how men had their ideas taken as given, as participants either agreed or acquiesced to the direction of the group that dominant male speakers laid out for them.

One example of how ideas might be adopted without discussion and how that had major repercussions for the campaign happened at a meeting in November. One of the men in leadership had talked to another organizer in Ontario about a strategy they were adopting to stop a major pipeline from being built by filing deputations as part of a consultation process. After explaining the context, he paused and asked, “So is there a general consensus that it’s a good way?” He waited 1.76 seconds, received a few nods, and continued, “Ok! Good” and continued to plan the details of the strategy. Though this action had no direct relationship with the goals of the campus campaign, participants spent significant amounts of time in mobilizing around the strategy. This idea, adopted without discussion in a meeting, became the major activity for November and December. While other participants perhaps demonstrated their tacit consent by participating, and shaped the implementation through their engagement, the decision-making process raises questions about what it means for a group when men’s ideas

are disproportionately adopted. This was not an isolated incident, but a consistent pattern across the data, suggesting that men's ideas persistently drove the actions of the group.

Exclusive talk

The second area we coded for was exclusive talk. Exclusive talk is defined as an exchange between two or more people but that does not include the majority of meeting participants, despite being part of a larger conversation. Exclusive talk was established through gaze and body positioning directed to the included speakers, explicitly naming the speakers who should participate, or discussing things only certain people had information about. Exclusive talk often happened at higher speeds of exchange, with no or small gaps between turns, and was often marked by a decision made by the participating speakers without opening the question out to the broader audience. We argue that exclusive talk works to establish who is in charge and whose voice is necessary in discussions and decisions, and is one of the key indicators of full participation in the community of practice.

Our coding revealed 46 instances of exclusive talk across our sample, and it was a common practice in the group. We measured instances of exclusive talk that were mixed gender and compared them against instances of exclusive talk that were men to men. While mixed gender was more common at a sample mean of 1.07 instances of exclusive talk per segment, men to men exclusive talk was also prevalent throughout the sample at a sample mean of 0.62 instances per segment. This demonstrates that men are highly likely to be engaged in exclusive talk that does not include women. Men to men and women to women exclusive talk also showed significant difference, with a sample mean for women to women talk of 0.077 instances per segment. In almost every case, men outnumbered women in instances of mixed-gender exclusive talk. Instances of male-dominated exclusive talk, with two or more men to one woman participating, were more than two and a half times more likely to occur than gender balanced (one man to one woman, or two men to two women) exclusive talk (with sample means of 1.11 and 0.42 instances per segment respectively).

Socially speaking, this is highly significant, in that even when women were included in the exclusive talk, it was at much lower levels of participation. It is both a sign of who was entitled to and given license to take up space and also served as a powerful indicator to other members of the group around who is a competent, central, full member. When a conversation could happen on behalf of the whole group but include only a small segment of the group, it signaled to participants of the exclusive talk and those on the periphery whose voices mattered, whose voices were required, and whose were not.

We argue that exclusive talk was one of the major ways full participation and expertise were established. They demonstrate for new members who must be included in discussions and who is unnecessary to include, for whatever reason. The fact that participation in these exclusive talk instances was so skewed based on gender-identity demonstrates how men's participation was affirmed as expert and necessary, while women's was not. During these exchanges, exclusive talk speakers often shared inside information that not everyone had, further demonstrating that their position was one of access and authority, and through their quick exchanges with other knowledgeable colleagues, they affirmed those involved in exclusive talk as co-authorities, while others did not need to be brought up to speed or included in decision making. The population of those not included in exclusive talk included all people of colour, men and women, (2) and most White women, which had the effect of establishing White men as the full participants in the community who were able and entitled to participate in discussions without including women and people of colour in the process.

Affirmations

Finally, we examined the rate at which speakers received affirmations from other members of the group for their speech contributions. We understand affirmation as acts that express agreement, acknowledgment, encouragement, or enthusiasm for what another speaker is contributing. Affirmation was vocal or gesture, and they were often paired. Any time a speaker received a comment during their speech turn that included "yes", "yeah", "yep", "uh huh", "mmhmm", or explicit praise (ie: "that's a great idea") it was coded as vocal affirmation. For gestures, the vast majority of coded instances were nods, from slight nods timed to affirm a speaker's contribution, to large nods with repeated head bobs and often involved the shoulders. We also see hand gestures like pointing, thumbs ups, and applause.

There were 602 instances of affirmation across our sample, averaging 26 instances per segment. We coded these as man to man, man to woman, woman to man, and woman to woman. First we found men received affirmations at almost three times the rate of women, with sample means of 17.7 times per sample segment for men compared to 7.25 times per segment for women. Looking more specifically, we found men gave affirmation

to other men at almost twice the rate of men's affirmation to women, at 9.6 instances per segment compared to 5.3 instances per segment. Women gave praise to men at four times the rate they affirmed women, at 8.2 instances versus 1.9 instances per segment. It is notable that the difference between women's praise of men and men's praise of men is not significant, at $t=0.4290$, thus we would not assume the rates of affirmation to men by women versus by men are statistically different. On the other hand, men's and women's praise of women is significant; women appear to affirm other women at a rate far less than men do.

After analyzing the data, it became clear that though those are strong trends in the sample, they do not fully account for the positioning dynamics of the group. We identified a pattern in the ways that men received affirmations that seemed distinct from how women received praise. A specific contribution by a man would often receive praise from multiple people, while women received praise from only one, or perhaps two people at a time. We believe receiving widespread praise is a strong indicator of community response and how people in the community position a speaker, so we analyzed affirmations based on the number of affirmation acts per contribution for men and women. The results showed men received affirmation from one speaker almost twice as often as women from one speaker, at sample means of 10.04 compared to 5.25 affirmations per segment. Men received affirmation from two speakers at two and a half times the frequency of women receiving affirmation from two speakers, at sample means of 2 compared to 0.79 affirmations per segment. Men received affirmation from three or more speakers more than six times as often as women receiving affirmation from three or more speakers, at sample means of 1.04 compared to 0.167 affirmations per segment. This shows a clear patterned behavior of men not only receiving more affirmations overall, but also their individual contributions being encouraged and affirmed by a larger segment of people per contribution.

Since men spoke far more frequently and for longer duration, one could argue we should not expect that women would receive validation at comparable rates. However, one factor serves as a control for that skew-- the rate at which men give affirmation to women, a sample mean of 5.3 times per segment. This asserts that there are opportunities for women to receive praise and makes it clear that gendered practices do influence the outcomes, since women only give praise to other women an average of 1.9 times per segment, meaning men are 2.8 times more likely to praise women's contributions than other women are. This raises questions of why women affirm other women differently than men affirm women, while both affirm men's contributions at statistically significant rates.

We found women did not receive praise at the same rates, and most significantly, we found that women hardly ever gave affirmation to other women. While coding the video and recognizing this trend, some of the women participants noted places where they trailed off or abandoned their contribution, and when asked they said they felt like they were "taking up too much space". In the video they were looking for encouragement to keep going or acknowledgement of their ideas (which they said they often gave to other speakers), and when they did not receive it, they assumed other participants disliked their ideas and stopped. For some of them, that experience chilled their participation for the rest of the meeting, resulting in reduced participation.

Affirmations were distinct from the other positioning categories above, in that it is a place for peripheral members to actively position others; while exclusive talk is largely determined by people in positions of mastery, affirmations are given by the larger group. It reinforces and legitimates the participation of those being affirmed by the membership, not just those in positions of mastery. These instances of affirmation were far more often directed at men and included extended series of gestures.

Discussion

These tendencies of small-scale gendered interactions gelled to create an environment that White men did not necessarily identify, but implicitly asserted their expertise as a group. In one exchange at a meeting halfway through the academic year this crystallized in a particularly telling way. A large group of new participants attended the meeting, surprising everyone, and in the attempt to be welcoming to new people, members of the group jockeyed to explain the context of the conversations we were having. Over the course of the first ten minutes in the first meeting of the semester, with 16 women and 9 men in attendance, 4 men spoke a total of 41 times, at a total duration of almost 9 minutes. Two women spoke a total of 20 times, with a total duration of just under one minute. But what we argue is the most revealing was a set of comments made by White men, eight minutes apart, during the beginning of the meeting.

Student 1: Ask us questions. And also, if you have a question you don't think you want to ask in public, just write it down, ask one of the people that talks a lot.

Student 2: If you would really like to come to the retreat, I guess just talk to...any of the people who you see talking a lot.

These statements, combined with a group dynamic in which the men were the people who participated most often and for the longest duration, implicitly advised new members that the men in the room were experts able to help explain things, and guide new members participation. What was also implied was that the women in the room were not expert, not able to answer questions, and not in positions of authority. Though this is just one vignette, it is a pervasive sentiment. Participants identified the most vocal and assertive members of the community of practice as the most central actors, capable of performing the full practices of the community.

This then created a core problem for women to surmount. If the full practices of the community included taking up space the way the White men did and receiving affirmation, other participants are in a bind. Even when women in the group had achieved some level of expertise, it was not recognized – and did not receive affirmation in the same ways men’s contributions did. Additionally, the women were unlikely to position themselves as authorities in the same ways, which led to men asserting their expertise over the expertise of those women. Many instances of this exist in our sample, but one specific example is how a White woman reported on her project and then ended with a question. We interpret this question as an invitation to the group to collaborate, rather than seeking specific information or a definitive response. However, one of the men in positions of mastery who frequently positions himself as an authority immediately responded (leaving no gap between the speaker’s question and his own answer), answering the question as if it were directed to him and a clarification question, rather than a question to the group about strategic direction. His definitive answer stopped further collaboration and ended the sequence, and served to undermine the woman’s credibility as a full member because both the direction she sought for the conversation was squashed and she was assumed to be incapable of answering the question that her male colleague was able to answer immediately and without equivocation.

Examples like these abound in our data and demonstrate that participants’ recognition as experts had less to do with their learning and capacity relative to the tasks of the community of activists, and more to do with their gendered modes of participation. In one example of this, we traced the planning process of an important presentation to the university administration. This was considered a high-stakes activity where people needed to be thoroughly prepared and ready to answer questions. In the planning discussions it was taken for granted that three of the White men were qualified and prepared, but when it came to finding additional participants, people were not sure who else was qualified. The planning group talked about needing someone “quick on their feet” who could answer question swiftly, with full confidence. They dismissed the possibility of several women as participants because they were seen as hesitant and the planning team feared the women would give too many qualifying hedges on their statements. The concerns of the planning team in including women in the presentation reflect assumptions that the administration would perceive only masculine performances of leadership as indicative of competence and expertise.

This presentation was the truest test of who had come into the full practices of the group, and the bar that people had to reach was a highly gendered and White mode of engagement. Women’s expertise in helping develop or edit the document they were presenting was not valued equally, their ability to organize students was not valued, their reproductive labour in sustaining the group was not valued — the test for expertise was the ability to perform masculinity under pressure when dealing with a powerful group of mostly White men.

Implications

This case provides an example of how White masculine performances can be conflated with expertise in a community of practice and how this process can systematically block women from becoming recognized as central members who are capable of participating fully in the core processes of the community. Becoming recognized as an expert can be as much about how one positions oneself as an authority through demonstrations of masculinity as it is about building competency in relevant skills and modes of participation, and people often give and receive affirmation based on gendered performances. Thus researchers need to be careful about assigning expertise and learning to these processes, which entrench and re-entrench performances of White masculinity rather than modes of participation that are more broadly available to members of a community. In order to accommodate the material and symbolic privilege that some participants bring to communities of practice, situated learning theorists must attend to and theorize the ways that gender, as well as race, colonialism, and other sociohistorical relations of power and inequity, influence interactions and access to full participation. Until we

do, our theories will not have the full analytic power they need to articulate the dynamics of learning and positioning within communities.

This study has implications in the political realm, where these participants worked daily. If men's voices are over-represented and disproportionately affirmed and adopted, what does this mean for the political actions that social movements like the climate movement take up? Our case suggests a few outcomes: the first is that men's leadership will be continually affirmed to the point that they understand themselves as experts. While positioning is usually an implicit process, occasionally they explicitly posit that women in the group are not experts and leaders, or are at least not as qualified as they themselves. This creates a feedback loop for participants of all genders, where men seem to become more confident in their contributions, while women become increasingly withdrawn and less willing to contribute. This also has the impact of shaping the political agenda; where there was contestation over group direction in our data, men's visions prevailed. Social movements need diverse strategies that reflect the experiences of all participants, and the more leadership and expertise is confined to White men, the less able we are to build effective strategy and movements.

Endnotes

(1) In our sample, no participants openly identified as other gendered, and so our discussion centres on cis gendered participants' experiences. We would posit that trans and other gendered people would experience specific but similar forms of gendered discrimination in participation, but it is outside the scope of our data.

(2) The racialized nature of exclusive talk, and many of the other modes of participation is significant and linked to the gendered patterns, and is the subject of a separate, more extensive analysis.

References

- Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford law review*, 1241-1299.
- Contu, A. & H. Willmott. (2003). Re-embedding Situatedness: The Importance of Power Relations in Learning Theory. *Organization Science*, 14(3), 283-296.
- Curnow, J. (2013). Fight the power: situated learning and conscientisation in a gendered community of practice. *Gender and Education*, 25(7), 834-850.
- Goldenberg, S. (2014, May 8). Why are so many White men trying to save the planet without the rest of us? *The Guardian*. Retrieved from <http://www.theguardian.com/commentisfree/2014/may/08/White-men-environmental-movement-leadership>.
- Hodges, D.C. (1998). Participation as Dis-Identification with/in a Community of Practice. *Mind, Culture, and Activity*, 5(4), 272-290.
- Holland, D., W. Lachicotte, D. Skinner & C. Cain. (1998). *Identity and agency in cultural worlds*. Cambridge, Mass.: Harvard University Press.
- Lave, J. (1996). *Understanding practice: Perspectives on activity and context*. Cambridge University Press.
- Lave, J. & E. Wenger. (1991). *Situated Learning Legitimate Peripheral Participation*. New York, NY: Cambridge Press.
- Paechter, C. (2003). Learning masculinities and femininities: Power/knowledge and legitimate peripheral participation. *Women's Studies International Forum*, 26, 541– 552.
- Salminen-Karlsson, M. (2006). Situating gender in situated learning. *Scandinavian Journal of Management*, 22.
- Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on language and social interaction*, 26(1), 99-128.
- Taylor, D. E. (1997). American environmentalism: The role of race, class and gender in shaping activism 1820-1995. *Race, Gender & Class*, 16-62.
- Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge: Cambridge Press.
- West, C., & D. H. Zimmerman. (1987). Doing gender. *Gender & Society*, 1(2), 125–151.

Acknowledgments

We thank participating activists, especially Amil Davis, Tresanne Fernandes, Keara Lightning, and Ariel Martz-Oberlander for their collaboration. We also thank Andrew Kohan, Omar Sirri, Victor Veitch and the GLITTER lab for their feedback on drafts. This work was generously funded by the Vanier Canada Graduate Scholarship.

Prior Knowledge and Mathematics Different Order Thinking Skills in Multimedia Learning

Thomas K.F. Chiu, The University of Hong Kong, tchiu@hku.hk

Abstract: This experimental study investigated prior knowledge and instructional visual aid on different mathematical order thinking skills of remembering, understanding and analyzing in multimedia learning. One hundred and twenty-three secondary school students was randomly assigned to a condition, in a 2 (prior knowledge: weaker vs stronger) x 2 (aid: with vs without) between subjects factorial design. In the experiment, the aid was evolved from variation theory and multiple representations and the materials were designed using multimedia learning principles. The results showed that the stronger group learned better on remembering, only the weaker group benefited from the aid on understanding, and the aid was more beneficial on analyzing. We suggest that multimedia presentation designs should take into account prior knowledge and order thinking. Limitations and future research are discussed.

Keywords: prior knowledge, multimedia learning, mathematics, order thinking skill, cognitive processing

Introduction

Multimedia presentations should display images and words simultaneously for learning (Mayer, 2009). The use of the presentations are considered effective in learning for mathematics, as it results in better remembering (retention) and understanding (transfer) skills (Chiu & Churchill, 2015a, 2015b; Mayer, 2009). Learner prior knowledge can influence the effectiveness of learning with multimedia presentations (Kalyuga, 2014; Mayer, 2009; Schnotz & Lowe, 2003). Many experimental studies showed that learners of different prior knowledge level responded differently to a multimedia design (Kalyuga, 2014; Leslie, Low, Jin, & Sweller, 2012, Liu, Lin, & Paas, 2014; Potelle & Rouet, 2003; Spanjers, Wouters, Van Gog, & Van Merriënboer, 2011). In their experiments, the tests used assessed remembering and understanding, but not higher order thinking skills. Their additional designs, treatments, were not customized for a specific domain (e.g. Spanjers et al. 2011; Leslie et al., 2012). Higher order thinking skills that require dynamic and relational internal (mental) representations would receive different effects from the additional designs. Therefore, this study aims to investigate the effects of prior knowledge and visual aid on three different order thinking skills of remembering, understanding and analyzing for algebra. Since many studies were conducted for remember and understanding, this study focuses on the higher thinking skill of analyzing. It is expected that the aid would differently affect the development of the three thinking skills. The aid was designed specifically for algebraic learning.

Prior knowledge and multimedia design

The cognitive theory of multimedia learning suggests that learners select relevant multimedia messages from the presentations (extraneous processing), organize them into a mental structure (essential), and finally integrate it with relevant prior knowledge (generative) retrieved from long-term memory (Mayer, 2009). Therefore, prior knowledge level and recalling process directly influence the effectiveness of the integration process – acquiring new knowledge – in multimedia learning (Chiu, 2015; Kalyuga, 2014; Mayer, 2009).

Many experimental studies support prior knowledge has impact on remembering and understanding in multimedia learning (Chiu, 2015; Kalyuga, 2007, 2014; Kalyuga et al., 2000; Leslie et al., 2012; Rey & Fischer, 2013; Spanjers et al., 2011). The experimental materials included additional designs, such as presenting aids audibly and/or visually, and controlling the pace of learning (see segmenting principle). For example, the design that showed steps to learn with images presented on screen worked best for weaker learners, but not for stronger learners (Kalyuga et al., 2000); visual representations helped younger children (less prior knowledge) learn science, but not older children (Leslie et al., 2012); segmented animations were more effective than continuous animations for less knowledgeable learners (Spanjers et al., 2011); and adding expository examples and illustrations was more beneficial for weaker undergraduate students rather than stronger students when developing the statistical skill (Rey & Fischer, 2013). The studies suggested that the designs – treatments – helped weaker learners understand the images and words presented. The designs provided information or environments to guide weaker learners to connect images and words presented thereby easing the cognitive processing for searching or recalling. In contrast, stronger learners may found the information is duplicated or the environments are discouraging. The designs became a burden, which required additional cognitive processing (Kalyuga, 2014).

This processing was unnecessary and resulted in less cognitive capacity for other kinds of processing, which is more important for stronger learners (Kalyuga, 2014). Moreover, most of the studies indicated incorporating additional design in multimedia representations was effective for both learners on remembering, but was effective for weaker learners only on understanding when learned from the presentations (Leslie et al., 2012; Rey & Fischer, 2013; Spanjers et al., 2011).

Instructional design for presentation – algebra

An important factor in experimental studies is the additional design. Instructionally providing appropriate and relevant learning messages for a specific-domain is beneficial for learners (Brophy, 2001; Marton et al., 2004; NCTM, 2000). In algebra teaching, numerous studies on presenting various forms of learning information for students have been conducted. Rittle-Johnson and Star (2007, 2009) endorse comparing and contrasting solution methods – students learn better by comparing an equation and its different solution methods, or by comparing different forms of an equation and their solution method. By applying variation to algebra teaching, students understand concepts better by seeing and experiencing different algebraic forms and solving methods simultaneously (Marton et al., 2004; Mok, 2009; Mok et al., 2002). Other than teaching strategies evolved from the variation, NCTM (2000) suggests that mathematics concepts should be presented in four forms that are numerical, graphical, algebraic and descriptive simultaneously to ensure effective algebra learning.

Mathematics orders of thinking skills and cognitive processing

The revised Bloom's Taxonomy categorized skills into six cognitive process dimensions (Anderson et al., 2001). The taxonomy suggests six orders of thinking skills. They are remembering, understanding, applying, analyzing, evaluating and creating. Remembering requires learners to retrieve, recognize and recall relevant knowledge from long-term memory; understanding requires them to construct their knowledge by way of classifying, summarizing and comparing; applying requires learners to implement procedure; analyzing requires learners to determine how parts relate to each another and to an overall idea; evaluating requires students to make judgments and explain their decisions; and creating requires students to reorganize what they have understood into a new pattern.

In mathematics, a higher order thinking skill requires more complete understanding (Berger & Torner, 2002; Derry, 1990; Rabinowitz, 1988) – a large number of knowledge and concepts needed for its development (Derry, 1990; Sweller & Cooper, 1985). Instructional methods for the higher order thinking skill will be more effective if they helped learners acquire prerequisite knowledge beforehand (Derry, 1990; Sweller & Cooper, 1985), for example, acquiring procedures of reassigning variables in an algebraic equation before solving conventional problems (Sweller & Cooper, 1985). Learning methods of lower order thinking skills can contribute to development of higher order thinking skill (Derry, 1990; Silver & Marshall, 1990). Developing a higher order thinking skill can involve different types of thinking/cognitive processes of its own and/or other lower order thinking skills (Derry, 1990; Silver & Marshall, 1990). In other words, more different thinking processes are likely to be involved when developing higher order thinking skills, which may lead to heavier learner essential and generative processing.

According to cognitive theory of multimedia learning, essential processing comprises selecting multimedia messages from a presentation (Mayer, 2009). An instructional design better instructs learners how to learn with images and words, which requires less time in selecting process, suggesting a more-structured design has less essential processing than a less-structured design. In other words, in less-structured environment, learners will have heavier essential processing and need more help. Less-structured environments are more effective to develop higher order thinking because the environments do not impede conceptually oriented interactions (Cohen, 1994). An instructional design, such as visual aid, that complicates development of lower order thinking skills may become more helpful for stronger learners when developing higher order thinking skills in less-structured environments.

Methods

The present study

The present study aims to investigate prior knowledge and visual aid on remembering, understanding and analyzing for secondary level algebra in multimedia learning. The materials used in the experiment were designed using Mayer's multimedia learning principles; the aid was designed for algebraic learning. We hypothesized that (1) the stronger students would performed better on remembering, (2) only the weaker students benefited from the aid on understanding (Leisle et al., 2012; Spanjers et al., 2011), and (3) with the aid was a better design than without on analyzing.

Participants and design

We invited 140 senior secondary level students aged from 16 to 18 years from a Hong Kong school to participate this study. Only 123 (around 60% boys) completed the experiments. We also invited two teachers in the school to conduct the experiments. A 2 x 2 between subjects factorial design with the factors prior knowledge (weaker vs stronger) and visual aid (with vs without) was used. This resulted in the four experimental conditions – 30 weaker students learning with the aid, 32 weaker students learning without the aid, 31 stronger students learning with the aid, and 30 stronger students learning without the aid.

Materials

This experiment included learning material and a posttest. We used the learning materials of Chiu and Churchill (2015b), see their online video in the experiment and Figure 1. The materials were developed using Mayer multimedia learning design principles under a design-based approach. The design aims to maximize learner cognitive capacity. Moreover, the visual aid design, evolved from variation theory, presented different forms of a quadratic equation and its different solving methods, and the four-section presentation – graph, equation, solving method and description. The description and solving method sections demonstrated the relationships between the graph and equation sections.

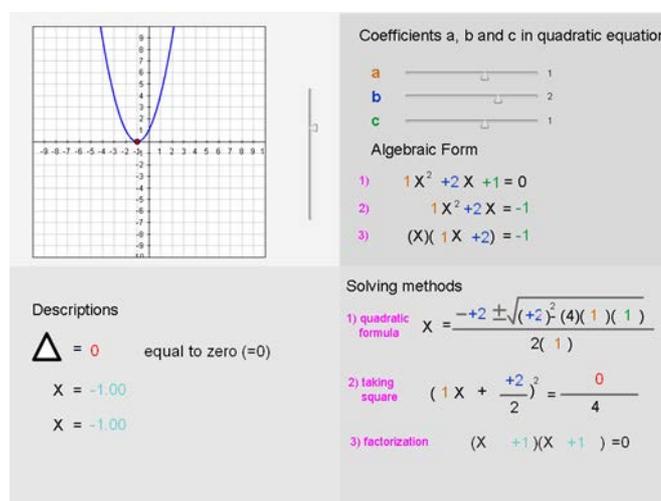
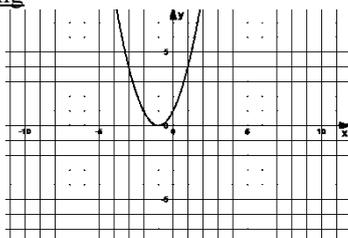


Figure 1. The material in the treatment group.

In the posttest, the questions were tested in the study of Chiu and Churchill (2015a). The questions assessed remembering (retention, level 1 in Anderson, Krathwohl, & Bloom, 2001), understanding (transfer, level 2) and analyzing (transfer, level 4). In the questions of remembering, students were required to write down the value(s) of roots and discriminant of a graph, see Figure 2; in that of understanding, the students were asked to identify graph(s) of a quadratic equation or a condition; and in the analyzing questions, students were required to consider two pairs of statements or expressions and decide whether they were related (or true) or not related (or false). Each of the questions was scored out of 1; and each skill was scored out of 12.

Remembering



What are the values of roots of an equation $f(x)=0$?
 What is the sign of the discriminant of an equation $f(x)=0$?

Understanding

Which of the following graph is $y=ax^2+bx+c$ if when $a<0$?

Analyzing

Statement or expression 1	Statement or expression 2
$\Delta=b^2-4ac$	Determines number of x-intercepts
value of a	Determines shape of the function $y=ax^2+bx+c$.

Figure 2. Questions used in the post-test.

Procedure

We first got the consent of the principal, teachers and parents. A pilot study was used to determine the time allowed for learning tasks. Before the experiment, the students completed an online 10-minute multiple-choice question quiz. The students were divided into the four experimental conditions based on a median split using the quiz scores obtained in the pretest (Kalyuga, 2007). We conducted the experiments in a computer room. In the experiments, the students were assigned to an individual seat in front of a personal computer. We first briefed them on the procedure of the experiment, and explained how to control the materials and what they would learn from the learning activities. The students had 40 minutes to manipulate the multimedia materials assigned, to understand the relationships between the graph and equation for learning. After the experiment, the students completed the posttests in 30 minutes.

Results

A t-test analysis showed that there was a significant difference between the weaker and stronger learner groups in the quiz, all p values <0.001 . This showed that stronger group had significant better skills to remember, apply and analyze graphical representations of algebraic equations.

We conducted univariate ANOVAs with remembering, understanding and analyzing as dependent variables. The results of Levene's tests, all p values >0.05 , indicated that all dependent variables met the assumption of homogeneity of variance. Means and standard deviations for scores of the three skills for both weaker and stronger students were shown in Table 1.

Table 1. Means and standard deviations of remembering, understanding and analyzing in the posttest.

Condition	Remembering		Understanding		Analyzing	
	M	SD	M	SD	M	SD
Weaker group with the aid (n=30)	7.47	1.00	7.40	1.30	6.87	1.20
Weaker group without the aid (n=32)	6.88	1.13	6.34	1.31	5.69	1.18
Stronger group with the aid (n=31)	9.42	1.29	7.16	1.50	7.35	1.40
Stronger group without the aid (n=30)	9.13	1.50	8.10	1.71	5.97	1.03

The results of ANOVA on dependent variable remembering showed the main effect of prior knowledge was found, $F(1, 119) = 83.66$, $p < 0.001$, partial $\eta^2 = 0.41$, indicating that there was a significant difference for the stronger ($M = 9.38$, $SD = 1.50$) over weaker groups ($M = 7.16$, $SD = 1.10$). There was no main effect of multimedia design, $F(1, 119) = 3.722$, $p = 0.056$, partial $\eta^2 = 0.03$, nor was there a multimedia design by prior knowledge, $F(1, 119) = 0.339$, $p = 0.561$, partial $\eta^2 < 0.01$.

With regard to the dependent variable understanding, univariate ANOVAs showed that there was no significant effect of multimedia design, $F(1,119) = 2.60$, $p = 0.110$, partial $\eta^2 = 0.021$. A significant main effect was found for prior knowledge, $F(1, 119) = 17.96$, $p < 0.001$, partial $\eta^2 = 0.13$, such that the stronger group ($M = 8.00$, $SD = 1.62$) outperformed the weaker group ($M = 6.85$, $SD = 1.40$). A significant interaction effect was found, $F(1, 119) = 5.52$, $p = 0.020$, partial $\eta^2 = 0.04$, see Figure 3. A significant simple effect was found for the weaker group who learned better with the aid, $F(1, 60) = 10.12$, $p = 0.002$, partial $\eta^2 = 0.144$; a significant simple effect was for the stronger group who learned better without the aid, $F(1, 59) = 0.22$, $p = 0.32$ (one-tailed), partial $\eta^2 = 0.004$. A significant simple effect was found for the multimedia design without the aid, such that the stronger

group learned better than the weaker group, $F(1, 60) = 20.78, p < 0.001, \text{partial } \eta^2 = 0.257$; no simple effect was found for the material design with the aid, $F(1, 59) = 1.87, p = 0.177, \text{partial } \eta^2 = 0.031$.

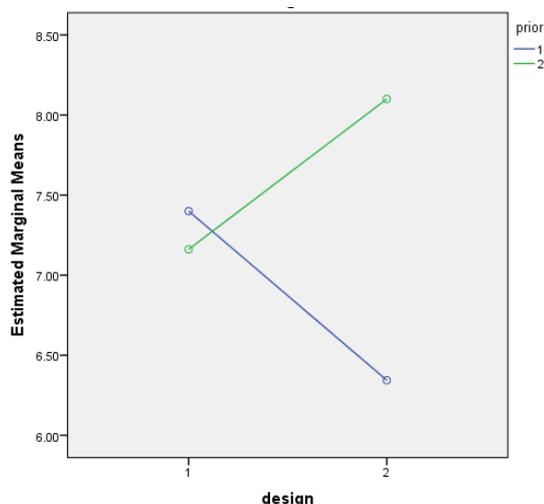


Figure 3. The interaction between prior knowledge and visual aid on remembering.
 Note: Design 1 – with aid; 2 – without aid. Prior 1 – novice group; 2 – advanced group.

For dependent variable analyzing, a two-way ANOVA indicated there was a significant main effect of multimedia design, $F(1, 119) = 34.55, p < 0.001, \text{partial } \eta^2 = 0.225$. The groups learned better with the aid ($M=7.11, SD=1.32$) rather than without the aid ($M=5.83, SD=1.11$). No significant effects were found for either prior knowledge, $F(1, 119) = 3.09, p = 0.082, \text{partial } \eta^2 = 0.025$, or interaction effect, $F(1, 119) = 0.23, p = 0.633, \text{partial } \eta^2 = 0.002$.

Overall, these results show that the aid have negative effects on the stronger group when developing understanding skills in multimedia environments, but not for remembering and analyzing.

Discussion and conclusion

The experiment reported in this paper was designed to investigate the effect of the instructional visual aid that is designed using variation theory and four-section representation in digital multimedia learning environments for mathematics students with different levels of prior knowledge on different order thinking skills – remembering, understanding and analyzing skills. The goal of this study was to investigate prior knowledge and visual aid on remembering, understanding and analyzing for secondary level algebra in multimedia learning.

The results of remembering and understanding supported the studies of Leslie and colleagues (2012), and Rey and Fischer (2013). Remembering was influenced by prior knowledge. The stronger group remembered better than the weaker group, suggesting the visual aid is not necessary for facilitating remembering. On understanding, the weaker group who received the aid outperformed those weaker students did not receive the aid. In contrast, the stronger group who received the aid performed less well than the strong group who did not receive the aid. These results suggest that the aid did help the weaker group see the relationships between the equation and the graph to understand the properties of the graph better. The description section in the aid, which may be seen as an explanation, directly explained the relationship between the graph and equation. The section appeared to be redundant for the stronger group who may have stronger graphical property skills. Processing the aid was extraneous processing in working memory and thereby reduced cognitive capacity available for other processing. Therefore, this demonstrates that for weaker group, graphs and equations might be facilitated by the inclusion of a visual instructional aid that facilitates essential processing. Also demonstrated was that the same visual aid had negative consequences for stronger group.

The results further showed that the with the aid group outperformed the without group in developing analyzing. A plausible explanation is that the learning task for analyzing involved heavier cognitive processing. The analysis questions required the students to justify if there were any relationships between pairs of statements. The students were required to see the connections among most multimedia messages when the learned with the material, which became less structured environment. The cognitive process in this environment is heavier

(Nievelein et al., 2013) for both weaker and stronger students. The students would need more help to connect multimedia messages for constructing a more complete understanding. Processing the aid that may provide essential information could be necessary for all the students, therefore, the aid were not be redundant for the stronger students, but facilitated essential processing.

Implications and suggestions

The findings also show multimedia materials were more effective when designed for learners of different levels of prior knowledge (Kaluga, 2014; Mayer, 2009) and order thinking skills. The studies have three implications. First, the findings confirmed that the visual aid format – variations and multiple representations – were more effective for weaker students in understanding. The aid explained the relationships between graphs and equations, and thereby helped weaker students have better understanding algebra (see Leslie et al., 2012; Rey & Fischer, 2013) in multimedia learning. Second, if not carefully orchestrated in order thinking skill, students of different prior knowledge may not receive the best design. In our experiment, for weaker students, the aid was more effective in understanding compared to analyzing and remembering. There may have better designs for remembering and analyzing skill development. As discussed before, numerous experimental studies support the effects of prior knowledge on multimedia learning, but most of them did not consider higher orders of thinking skills. Our findings suggested that the order of thinking skills could influence the effects of prior knowledge on the instructional design in multimedia learning. Third, processing the aid for higher order thinking skills can facilitate essential processing for better learning outcomes. In other words, designs that are ineffective for stronger learners in structured tasks may become effective for them in less structure tasks.

The results also afford two suggestions. First, multimedia presentation designs should consider order thinking skill and prior knowledge. Using order thinking skills to identify instructional formats or procedures can balance the degree of guidance offered to learners. Second, in algebra learning, providing graphs and equations can be enough for students when developing remembering, but when developing understanding, an aid should be added to weaker students only. An aid should be included for all the students when developing the higher order thinking skill of analyzing.

In conclusion, the findings can contribute to the design principal for learner prior knowledge in multimedia learning, such as Kalyuga (2014) and Mayer (2009). One multimedia learning design cannot fit all learners of different prior knowledge levels (Kalyuga, 2007, 2009, 2014, Mayer, 2009). This study further suggests that one multimedia learning design also cannot fit all different order thinking skills. Instructional designs for the weaker learners may become useful for the stronger learners when developing higher order of thinking skills that require higher cognitive processing.

Limitations and future directions

The present findings are also relevant to adaptive digital multimedia learning environments. Multimedia learning will be used in many adaptive learning environments (Van Merriënboer & Sweller, 2005) in the future. Most studies suggest using learner behavior, characteristics and prior knowledge (Astleitner & Wiesner, 2004; Chiu, 2016; Chiu & Churchill, 2015c; Kalyuga, 2006, 2008) to modify the environment or to give personalized feedback to learners. This present study suggests that the adaptive environment should include order thinking skills and learner prior knowledge to identify multimedia presentations or tasks for delivery to promote individual learning.

It is important to better evaluate the effects of instructional designs and learner prior knowledge level on different order thinking skills. The results of the present experiment could also be extended by additional studies on other higher order thinking or in other subject domains.

Overall, future research on adaptive learning environments should focus on cognitive processing, and interactions among learner prerequisites, multimedia presentations and learning outcomes.

References

- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Boston, Massachusetts: Allyn & Bacon.
- Astleitner, H., & Wiesner, C. (2004). An integrated model of multimedia learning and motivation. *Journal of Educational Multimedia and Hypermedia*, 13(1), 3-21.
- Berger, P. & Törner, G. (2002). Some Characteristics of Mental Representations of Exponential Functions: A Case Study with Prospective Teachers. In D. Mewborn, P. Sztajn, D.Y. White, H.G. Wiegel, R.L. Bryant, &k. Nooney (Eds), *Proceedings of the 24th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol.1, pp. 223-231), Columbus, OH: The Eric Clearinghouse for Science, Mathematics, and Environment Educaiton.

- Brophy, J. E. (2001). *Subject Specific Instructional Methods And Activities*, 8. West Yorkshire, UK: Emerald Group Publishing.
- Chiu, T.K.F. (2015). Effects of prior knowledge on mathematics different order thinking skills in mobile multimedia environments. In D. Churchill, J. Lu, T.K.F. Chiu, & F. Bob (Eds), *Mobile Learning Design: Theories and Applications* (pp.373-386). Springer: New York.
- Chiu, T.K.F. (2016). Introducing electronic textbooks as daily-use technology in schools: A top-down adoption process. *British Journal of Educational Technology*.
- Chiu, T.K.F., & Churchill, D. (2015a). Design of Learning Objects for Concept Learning: Effects of Multimedia Learning Principles and an Instructional Approach. *Interactive Learning Environment*, 1,1-16. DOI: 10.1080/10494820.2015.1006237
- Chiu, T.K.F., & Churchill, D. (2015b). Exploring the Characteristics of an Optimal Design of Digital Materials for Concept Learning in Mathematics: Multimedia Learning and Variation Theory. *Computers & Education*, 82, 280-29. DOI:10.1016/j.compedu.2014.12.
- Chiu, T.K.F., & Churchill, D. (2015c). Design of learning objects for concept learning: effects of multimedia learning principles and an instructional approach. *Interactive Learning Environment*, DOI: 10.1080/10494820.2015.1006237
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1-35.
- Derry, S. J. (1990). Learning strategies for acquiring useful knowledge. In B.F. Jones, & L. Idol (Eds), *Dimensions of thinking and cognitive instruction* (pp. 347-379). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dori, Y. J., Tal, R. T., & Tsaushu, M. (2003). Teaching biotechnology through case studies—can we improve higher order thinking skills of nonscience majors?. *Science Education*, 87(6), 767-793.
- Kalyuga, S. (2006). Assessment of learners' organised knowledge structures in adaptive learning environments. *Applied Cognitive Psychology*, 20(3), 333-342. DOI: 10.1002/acp.1249
- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19(4), 509-539.
- Kalyuga, S. (2008). Managing cognitive load in adaptive multimedia learning. Hershey, PA: IGI Global.
- Kalyuga, S. (2014). The Expertise Reversal Principle in Multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*, (pp. 576-597). New York, NY: Cambridge University Press.
- Leslie, K. C., Low, R., Jin, P., & Sweller, J. (2012). Redundancy and expertise reversal effects when using educational technology to learn primary school science. *Educational Technology Research and Development*, 60(1), 1-13.
- Liu, T. C., Lin, Y. C., & Paas, F. (2014). Effects of prior knowledge on learning from different compositions of representations in a mobile learning environment. *Computers & Education*, 72, 328-338.
- Marton, F., Tsui, A.B.M., Chik,P., Ko, P.Y.,& Lo, M.L. & Mok, I.A.C. (2004). *Classroom discourse and the space of learning*. New Jersey, NY: Lawrence Erlbaum
- Mayer, R. E. (2009). *Multimedia learning*. New York, NY: Cambridge Press.
- Mok, I. A. C. (2009). In search of an exemplary mathematics lesson in Hong Kong: An algebra lesson on factorization of polynomials. *ZDM*, 41(3), 319-332.
- Mok, I.A.C., Runesson, U., Tsui, A., B.M., Wong, S.Y., Chik, P., and Pow, S. (2002). Questions and Variation. In F. Marton & P. Morris (Eds.) *What matters? Discovering critical conditions of classroom learning*, (pp.75-92). Göteborg: Acta Universitatis Gothoburgensis.
- National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- Nivelstein, F., Van Gog, T., Van Dijck, G., & Boshuizen, H. P. (2013). The worked example and expertise reversal effect in less structured tasks: Learning to reason about legal cases. *Contemporary Educational Psychology*, 38(2), 118-125.
- Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58, 327-345.
- Rabinowitz, M. (1988). On teaching cognitive strategies: The influence of accessibility of conceptual knowledge. *Contemporary Educational Psychology*, 13(3), 229-235.
- Rey, G. D., & Fischer, A. (2013). The expertise reversal effect concerning instructional explanations. *Instructional Science*, 41(2), 407-429.
- Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, 99(3), 561–574.

- Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101*(3), 529–544.
- Schnotz, W., & Lowe, R. (2003). External and internal representations in multimedia learning. *Learning and Instruction, 13*(2), 117-123.
- Silver, E. A., & Marshall, S. P. (1990). Mathematical and scientific problem solving: Findings, issues, and instructional implications. In B.F. Jones, & L. Idol (Eds), *Dimensions of thinking and cognitive instruction* (pp. 265-390). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Spanjers, I. A., Wouters, P., Van Gog, T., & Van Merriënboer, J. J. (2011). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior, 27*(1), 46-52.
- Sweller, J., & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59–89.
- Van Merriënboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147-177. DOI: 10.1007/s10648-005-3951-0

Prior Knowledge for the Construction of a Scientific Model of Equilibration

Hillary Swanson, Northwestern University, hillary.swanson@northwestern.edu

Abstract: The Next Generation Science Standards charge U.S. teachers with the task of including *patterns*, as a crosscutting concept, in their science curriculum. This study explores prior knowledge that is relevant to students' construction of a scientific model of an equilibration pattern. A *Knowledge in Pieces* lens is applied to video transcript of a class discussion on thermal equilibration in order to identify elements of students' prior knowledge that might be productive for their construction of a scientific model of equilibration. The discussion occurred near the beginning of a unit on equilibration, in the middle of a yearlong pattern-based curriculum. Twenty-one 8th grade students participated in the class. Six previously undocumented knowledge elements are identified, characterized, and considered in terms of their potential productivity for helping students construct a *difference drives rate* model of equilibration. These findings contribute to literature concerned with the productivity of prior knowledge in conceptual change.

Keywords: Conceptual change, Knowledge in Pieces, prior knowledge, Knowledge Analysis

Introduction

The Next Generation Science Standards promote 7 crosscutting concepts as recurring themes throughout the K-12 science curriculum. The first of these is *patterns*. The Framework for K-12 Science Education argues that patterns are important for science and science learning because they are readily observed in nature across dimensions of structure and process; noticing patterns is a first step to asking deeper questions about the mechanisms that drive their emergence; and patterns can be used as rules for categorizing similar phenomena within a unified explanatory framework (Schweingruber, Keller & Quinn, 2012).

This study focuses on one specific process pattern: *equilibration*. Equilibration can be found in a range of contexts across the physical sciences. It can be seen in both the warming to room temperature of a cold liquid and the diffusion of gas across a semipermeable boundary. Both phenomena follow a pattern in which their *rate* of equilibration is directly proportional to the *difference* between their current state and their equilibrium state. For example, a cold liquid's *rate* of temperature change is directly proportional to the *difference* between its temperature and the temperature of the environment with which it is equilibrating. When the temperature *difference* is large, the *rate* of temperature change is large. As the *difference* decreases, the *rate* of change decreases. A *difference drives rate* model of equilibration can be a powerful tool for scientists as it predicts and explains a range of observable phenomena. Despite its power, the *difference drives rate* pattern is usually introduced later in the science curriculum when it is taught in the form of a differential equation, for example as Newton's law of heating.

It is the goal of this study to identify elements of prior knowledge that might serve as a foundation for younger students' construction of a conceptual understanding of the *difference drives rate* model of equilibration. This understanding might in turn help them make sense of phenomena like thermal equilibration before they have the mathematical machinery to interpret Newton's law of heating; it may also serve as a foundation for their construction of related mathematical concepts later on. The research reported here looks at a single class discussion during which students generated explanations for an instance of thermal equilibration and asks the question: "What prior knowledge emerged as a resource for students' construction of a *difference drives rate* model of equilibration?"

By identifying elements of prior knowledge that are potentially fruitful for science learning, this study makes theoretical and empirical contributions to the conceptual change literature. The larger work within which the study is situated makes a practical contribution to classroom science through insights into how instruction can be designed to leverage students' prior knowledge. Such instruction not only supports learning, but positions learners as agents of knowledge construction. This, in turn, may have the power to enhance students' picture of the scientific enterprise, making science a discourse where their voices matter, and to which they might one day make meaningful contributions.

Theoretical orientation

I align my research with the *Knowledge in Pieces (KiP)* perspective (diSessa, 1993). *KiP* models knowledge as a complex system of elements that are cued in response to the sense-making demands of a particular context. Consistent with constructivism, the development of more sophisticated knowledge involves the reorganization and refinement of existing knowledge. This view stands in contrast with perspectives on conceptual change that view knowledge as a unitary theory that is cued consistently across contexts, and learning as a process in which a naïve theory is replaced by a more expert one (Clement, 1982; McClosky, 1983).

Motivated by the *KiP* model of learning, I investigate elements of students' prior knowledge that are potentially productive for their construction of a scientific model of equilibration. I have therefore analyzed student contributions to a relevant class discussion and identified potentially productive knowledge elements, characterized those elements, and considered how they might foster students' construction of a *difference drives rate* model of equilibration. I compare the elements that I identify with a class of intuitive knowledge called *phenomenological primitives*, or *p-prims* (diSessa, 1993). *P-prims* model the smallest units out of which a larger knowledge system may be comprised. They are deeply intuitive elements of explanations, providing a learner with an *intuitive sense of mechanism*. An example of a previously documented *p-prim* that is also identified here is *Ohm's p-prim*: the intuition that *more effort begets more result*.

This study is embedded in a larger program of design-based research (Collins, Joseph & Bielaczyc, 2004). The present analysis considers data from a class discussion that occurred during an iteration of a yearlong course in which students practiced identifying and articulating models of process patterns including equilibration. The fundamental assumption of *KiP* that guided my design of instruction is that prior knowledge can play a productive role in students' construction of scientific knowledge. Designing instruction is therefore about creating opportunities for the learner to activate productive elements of their prior knowledge and engage those elements in their construction of new knowledge (Hammer, 2000). This stands in contrast with instruction designed from the *misconceptions* perspective, which focuses on identifying incorrect knowledge and replacing it with correct knowledge (McClosky, 1983). Instruction was therefore designed to elicit and engage students' prior knowledge in their construction of more sophisticated pattern models.

To identify productive elements of students' prior knowledge, I drew on a set of strategies for qualitative analysis connected with the *Knowledge in Pieces* framework and organized under the name *Knowledge Analysis (KA)*. *KA* is characteristically focused on knowledge, analyzing the *ideas* that learners internalize, rather than the means or modes through which they communicate those ideas. *Knowledge in Pieces* is firmly committed to the complexity and idiosyncratic nature of individual knowledge systems. *Knowledge Analysis*, therefore, begins with a grounded characterization of the data. It then compares emergent knowledge with the existing *KiP* model in order to contribute to its expansion and refinement.

Methodological approach

This study investigates the prior knowledge of a group of 8th grade students that is potentially productive for their construction of a scientific model of equilibration. It considers data taken from a class discussion that occurred during an iteration of design-based research. The instructional design under test was a middle school science course called Patterns Class. Patterns Class met for 40 minutes on Monday, Tuesday, and Thursday mornings. The class met both fall and spring semesters, totaling approximately 60 hours of instruction over the course of the school year. The researcher was the primary instructor and undergraduate research assistants doubled as teaching assistants, attending class about one morning a week with moderate consistency over the school year.

Patterns Class curriculum was designed to guide students through the systematic exploration of four patterns: threshold, equilibration, exponential growth and oscillation (though these names were never formally introduced to the class participants). The results presented here are taken from analysis of data from the equilibration unit. The target model of the equilibration pattern can be characterized as *difference drives rate*, in which the rate of a system's equilibration is directly proportional to the difference between its current state and its equilibrium state. An example that is explored by students during the unit is the equilibration of a glass of cold milk with a warm room. At the start, the temperature of the milk is farthest from the temperature of the room and it is observed to warm at the greatest rate. As it warms, the difference between the milk's temperature and the temperature of the room decreases and the milk is observed to warm at a progressively slower rate until it reaches the temperature of the room.

In each of the units, instruction was designed to support students in modeling the pattern by activating and engaging their prior knowledge in a general sequence of activities that alternated between exploring prototypical examples and generating and refining models of the patterns those examples followed. The equilibration unit was comprised of 7 core activities and ran for approximately 20 instructional hours. The sequence of core activities consisted of: 1) investigating the thermal equilibration of a glass of cold milk, 2)

investigating the thermal equilibration of a glass of hot tea, 3) constructing a model of the general pattern of behavior common to both examples, 4) investigating the equilibration of beans in a partitioned box through a simulation of diffusion, 5) revising pattern models, 6) generating additional examples that followed the pattern of behavior exhibited by the three examples, and 7) revising pattern models.

Participants

Twenty-one 8th grade students participated in the focal iteration of the Patterns Class. The majority of the students were children of families that had immigrated to the U.S. from Mexico and Central America. Several students identified as African American and European American. English was a second language for most, Spanish being the primary language spoken at home. The majority of students attending the school were designated as English Language Learners, and the majority qualified for free and reduced lunch. The group of students participating in Patterns Class was selected on the basis of availability and willingness to participate. The particular school was selected because the science teacher there was amenable to sharing her elective period students with our group for both fall and spring semesters. Her elective period had traditionally been used as a science enrichment period for students that had scored proficient or higher on tests of basic skills in English and math.

Data collection

Data relevant to the present report were collected in three different forms: 1) video footage, 2) field notes, and 3) teacher reflections. Two digital video cameras recorded the activities of every Patterns Class. One camera was positioned at the middle of one side of the room and pointed at an angle out across the classroom toward the front board. This camera captured the activity of the teacher and/or student(s) speaking at the front of the room and the artifacts recorded on the front board. A second camera was positioned at the front of the room just to the side of the front board. It was pointed out across the tables at which the students sat and captured the activity of the students as they attended to the front of the room or engaged in small group work. Field notes taken by research assistants and reflections written by the teacher were used to identify segments of video for more careful analysis. Video footage of a whole-class discussion was selected as data for the present study, as it was particularly generative of prior knowledge. The discussion was transcribed and investigated through a Knowledge in Pieces lens. Potential knowledge elements were identified, characterized and compared with previously documented phenomenological primitives, and considered in terms of their potential productivity for students' construction of a *difference drive rate* model of equilibration.

Certain conventions were used to transcribe students' verbal utterances captured on videotape. The names of students contributing to the class discussion have been replaced with pseudonyms. I omit certain contributions that are interruptions or part of productive conversations that are not temporally linked to the contributions that are immediately relevant to my analysis. I use the symbols defined below to indicate the flow of speech and gesture:

- // - break in speech
- /.../ - interruption or parallel speech
- <...> - gesture

Major findings

My analysis addresses the research question: "What prior knowledge emerged as a resource for students' construction of a *difference drives rate* model of equilibration?" I will present elements of prior knowledge invoked by the students during a class discussion to explain the *changing* rate of temperature change exemplified by the results (Figure 1, below) of a thermal equilibration investigation. Facilitated by the teacher, students shared their explanations for why the temperature of a glass of cold milk would warm up quickly at first and then slow down as the milk reached room temperature. Below, I present the knowledge elements in their order of appearance and describe their context of emergence, character, and potential productivity for students' construction of a *difference drives rate* model of equilibration.

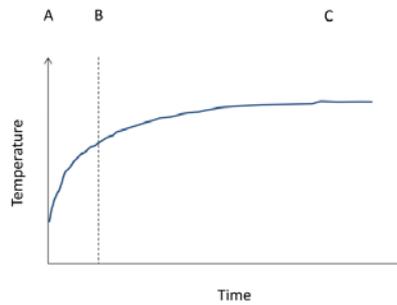


Figure 1. Approximate graph of temperature over time for cold milk equilibrating with a warm room.

Knowledge element 1: *Slowing Down to Stop*

This element emerged near the beginning of the whole class discussion. The teacher had asked students to respond to an explanation for the milk's changing rate of temperature change written by one student at the end of the previous lesson: "Because it was getting to room temperature at the end so it was slowing down. It's like a race, when you're getting to the destination you start to slow down."

- Leo: um//why would you slow down when you're about to finish a race? It doesn't make sense
- Alvaro: say there's a wall// are you going to run straight into it Leo?
- Leo: well I'm not gonna go slower though// because then I'll lose
- Alvaro: like no no no no no like// say you're winning cause you're going as fast as you can// then when you're gonna reach the wall// don't you start to like <stomps feet on the ground> kinda/
- Michelle: /is there a wall?
- Alvaro: yes there's a wall
- Michelle: there is no wall
- Alvaro: there is a wall
- Michelle: where?
- Alvaro: the room temperature// huh? huh Michelle what? This time when you're running// you're racing// you run as fast as you can// but then you're gonna go// you're going to hit a wall so you start to like slow down// *you have to slow down to stop*

The discussion begins when one student, Leo, challenges the internal consistency of the race analogy by questioning the sensibility of slowing down at the end of a race. Alvaro responds to Leo by explaining that this is a special case in which the race ends at a wall. Michelle challenges Alvaro's wall, questioning the mapping between the wall and the milk. Alvaro connects the wall with room temperature and presents his idea once more. His explanation for the decreasing rate of change is: "*you have to slow down to stop.*"

Slowing down to stop appears to be a notion that belongs to Alvaro's *intuitive sense of mechanism*. The wall, in Alvaro's analogy, serves to reinforce the sense that there is a hard line beyond which neither the runner, nor the temperature, can go. Trying to stop after moving quickly is a common experience of many middle school-aged students (and many people, in general). Attempting an abrupt stop after activities like running, skating or driving are several possible origins of this intuition. If one attempts to stop abruptly, inertia causes one to overshoot and continue beyond the intended stopping point. An abstraction of the experience of inertial overshooting may be the previously documented *p-prim overcoming* – the sense that one influence wins over another (diSessa, 1993). One may intuitively know that they can avoid this overshooting by slowing down before they stop. It is reasonable to conjecture that this is an intuitive element of knowledge that Alvaro is drawing on when he says: "*you have to slow down to stop.*"

Though it has not been previously documented, *slowing down to stop* is a candidate for intuitive knowledge roughly the grain size of a *p-prim*. The connection between *slowing to a stop* and the experience of overshooting suggests that the knowledge element may in some cases be cognitively linked (and have high cuing priority) with the *p-prim overcoming*. *Slowing to a stop* seems close, in character, to the *p-prim slowing equilibration* – the notion that things slow as they approach equilibrium (diSessa, 1993). I would argue, however,

that it is different in three important ways. The first difference is the *degree of urgency*. While *slowing equilibration* is described by a natural and gradual easing to a stop, the slowing described by Alvaro is marked by an urgent need to stop. The second difference is in the location of the *impetus* for slowing. In the case of *slowing equilibration* the object slows because that is its natural internal tendency. In the case of Alvaro's slowing, the object is pressured to stop by an external entity or demarcation. The third difference is related to the *stopping point*. In the case of *slowing equilibration* the object is returning to its natural state or state of balance. In the case of Alvaro's slowing, the object is moving to a new destination.

Slowing to a stop is potentially very productive for the construction of a scientific model of equilibration. While it only explains the latter half of the equilibration curve (Figure 1, segment BC), *slowing to a stop* maps well to that part of the curve and, moreover, it is analogous with *difference drives rate*, though in a spatial incarnation as *distance drives rate*. As the distance to the endpoint decreases, so does the speed of the equilibrating entity, until, at the endpoint, the rate of change is zero. The productivity of the element is limited in that it cannot be generalized to explain the decrease in speed at any distance from the endpoint. Instead it only explains the change in speed at a point where the entity is near enough to the endpoint that it needs to begin to slow down. This limitation will be clear in the way Alvaro thinks about the first half of the equilibration curve (Figure 1, segment AB).

Knowledge elements 2 and 3: *Energy Drives Rate* and *Energy is Greatest at the Start*

Noticing that the students have focused on the second half of the equilibration curve (Figure 1, segment BC), the teacher turns their attention to the task of constructing an explanation for the first half (Figure 1, segment AB).

Teacher: let me ask you this: what if we need to do the other half of that// going really fast at the start? Why would the water start warming up really fast at the start?

Leo: because *at the start you have a lot of energy to run so you run!*

Continuing to reason within the general context of the race analogy but shifting from the *end* of the race to its *beginning*, Leo offers an idea that is completely unrelated to the "wall" at the finish line. He suggests: "at the start you have a lot of energy to run." Because it is given in response to the teacher's question, Leo's idea could be interpreted as meaning "at the start you have a lot of energy to run, *so you run fast*." At the beginning of the race the runner's speed is the greatest because they have the most energy. As the race unfolds, the runner uses up energy and slows down as a result. It seems as though there are two underlying intuitions here. The first is *energy drives rate*. This is a potential abstraction of the experience of moving with greater speed when one feels energized and moving more slowly when one feels less energized. The second intuition is that *energy is greatest at the start* of an activity. This implies that one has a fixed amount of energy to devote toward activity, and as one is active, one depletes that fixed amount of energy. The second intuition is potentially fruitful for constructing a scientific understanding of conservation of energy. It is probably grounded on the very common physical experience of moving quickly at the beginning of an activity (such as a race) when one is fresh and has not yet exerted oneself. As one goes along, one experiences increasing fatigue (which one might attribute to the expenditure of energy). Really, a person uses up chemical potential energy as their body converts it into the kinetic energy of their movement, so the explanation that Leo suggests is not unscientific in the context of the race.

Leo's ideas might also be connected with the impetus conception of force, documented by early *misconceptions* researchers (Clement, 1982; McClosky, 1983). The conception explained a moving object as being driven forward by an internal impetus *force* that perpetuated its motion until it gradually died out. Here Leo is calling the impetus force *energy* as opposed to *force*, but the idea appears to be much the same. At the beginning, the object (the runner in the case of his analogy) could be said to have a lot of force, but the force dies away over time and results in the slowing and eventual stop of movement. An important distinction between Leo's idea and the impetus force conception is that Leo's idea is invoked to explain human motion (though ultimately he may have meant it to explain the warming of the milk, as he shared the idea in response to the teacher's question about the milk). Impetus force is a documented explanation for the slowing of inanimate objects. While it is not scientific to explain an object's slowing as the result of running out of impetus force, it *is* scientific to explain the slowing of a human runner as the result of running out of energy.

Leo's intuition about having more energy to run fast at the start may in fact produce a curve that matches the equilibration curve and therefore provide a working explanation (possibly even one with predictive power) for the changing rate of temperature change. Its productivity is ultimately limited, however, because it does not map to *difference drives rate*, which is the target model of equilibration. While a race of a particular length might

afford a run that begins as a sprint and slows as the runner loses energy, it is possible to imagine that a shorter race would yield a sprint of a consistent rate, or that a longer race would yield a run of variable pace, or one in which the runner stopped entirely before reaching the finish line. A main limitation of Leo's reasoning is that rate is not dependent on the amount the entity has to change (position or temperature). It therefore does not map to the scientific model of equilibration as *difference drives rate*.

Knowledge elements 4 and 5: *Space Allows Speed and Ohm's P-prim*

Leo has hardly shared his idea when Alvaro interjects to voice his disagreement and provide an alternative explanation.

Alvaro: /no// I disagree// Because *you have more space* <gestures spreading hands apart>// you're not going to crash into a wall <moves whole body forward> so *that's why you run faster!*

Leo: /that's not really true/

Alvaro: /you *try to win!*//So you can try to like// win// *that's why you're running faster!*// but then// when you're like approaching the wall// you're gonna start to like slow down// so you don't want to like crash into it

Continuing to reason in the context of the analogy of a race to a wall, Alvaro produces an alternative explanation for the high speed of the runner at the beginning of the race. This is that, unlike the end of the race where the wall constrains the runner's speed, at the start, the runner can safely run as fast as they want. His final contribution suggests that the runner will go as fast as possible at the start because they are trying to win the race. The goal of winning, together with the safety of the large space, drives the high speed of the runner at the start of the race. It is possible that Alvaro's conception of *space* is invoked as a result of the way he is thinking about the race with respect to the wall: with both the start of the race and the wall in mind *the space between the starting line and wall is made salient*. At the start, there is a great deal of space between the two and it is therefore *safe to run fast*. Near the finish there is less space and the runner is in *danger of crashing into the wall* and must therefore slow to a stop.

Both driving factors - space and effort – appear to be separate intuitions. Alvaro's explanation has nice internal consistency and it is likely that his intuition that *space allows speed* is cognitively connected with *slowing to a stop*. When there is space to run quickly, one runs quickly, as the space decreases one must slow to a stop. Both knowledge elements are connected with, and probably mediated by, a perception of safety. The experience of running to a wall is a possible origin of the intuition that *space allows speed*. Alvaro's intuition that effort drives rate is essentially an instantiation of the previously documented intuition *Ohm's p-prim*: greater effort begets greater result. *Ohm's p-prim* has previously been shown to be highly productive in students' construction of a *difference drives rate* model of equilibration (diSessa 2014). As well, Alvaro's idea of *space allows speed* might be useful for constructing a *difference drives rate* model of equilibration. It does map to *difference drives rate*, if the entity is motivated to go as fast as it is allowed *and* responds proportionally to the decrease in space. This conceptualization of the relationship between space and rate is limited, however, in that it is possible to imagine that the entity would not have a proportional relationship with the distance. Rather, it may go fast at a constant rate until it needs to slow down to avoid crashing into the wall.

Knowledge element 6: *Difference Drives Desire to Eliminate Difference*

The teacher summarizes the ideas shared by Leo and Alvaro and frames them as competing explanations. She adds another explanation to these two, one that had been written by a student the previous day: "Because when it first goes in, the temperature is so different so it changes faster." She invites the students in the class to respond to any of the ideas put forth by their classmates thus far. She pulls Sofia's name at random and asks her which of the ideas make sense to her.

Sofia: they all make sense// because// the last one// because the temperature is super different// it wants// I guess it warms up faster

Teacher: why does it go fast and slow// based on what you said?

Sofia: I guess// It wants to get warmer? I don't know

Teacher: <writes on board> because it wants to/

Sofia: /get warmer and reach room temperature

The original idea read by the teacher maps directly to *difference drives rate*, however it is not clear what intuition, if any, underlies it. In explaining why this idea makes sense to her, Sofia invokes an intuition about the relationship between the difference in temperatures and the *desire* of the cold milk to reach room temperature. She appears to be suggesting that the milk wants to be at the same temperature as the room. The temperature difference sparks that desire in the milk, to "get warmer" and "reach room temperature." Though her explanation is fragmented, it seems that for Sofia, the milk's desire to eliminate the difference in temperature drives the rate at which it warms. It is plausible that she is invoking *Ohm's p-prim*, connecting desire and rate through effort. In this case, a more complete version of her explanation would be comprised of three elements of intuitive knowledge: *difference drives desire to eliminate difference*, *desire drives effort*, and *effort drives rate (Ohm's p-prim)*.

With so little data, it is difficult to say for certain what knowledge Sofia invokes. If she is in fact thinking that the milk's desire to warm is directly proportional to its difference with room temperature and invoking *Ohm's p-prim*, this would be highly productive. On this view, when the temperatures are very different, the milk has a very strong desire to decrease that difference, and, as a result of that strong desire works harder and decreases the difference more quickly. As the difference decreased, so would the desire and effort of the milk, causing the difference to decrease more slowly. In this way, logic relating the difference in temperature to the rate of temperature change is highly productive for constructing a *difference drives rate* model of equilibration.

Knowledge element 7: *Difference Drives Rate*

The teacher turns to the class to invite the other students to respond to the ideas that have come up during the class discussion. She draws Mateo's name at random and asks him to explain the ideas that have made sense to him during the discussion.

Mateo: there's a big difference in the temperature// um like it has a lot to cover so it wants to do it fast

Teacher: what about this from here to here? <points to the second half of the warming curve>

Mateo: like Alvaro said// it slows down 'cause there's a wall

Mateo's intuition that "it has a lot to cover so it wants to do it fast" appears to map directly to *difference drives rate*. It is a sensible instinctive reaction to having a great distance to cover, or a great amount of difference to decrease. Mateo's intuition seems to be the symmetric reflection of *Ohm's p-prim*. If we interpret "has a lot to cover" as implicating a large result to produce and "wants to do it fast" as implicating a desire to put forth a great amount of effort, "it has a lot to cover so it wants to do it fast" can be interpreted as "greater result motivates greater effort." It is important to note that this is only meant to explain the beginning of the equilibration curve and Mateo invokes Alvaro's notion of *slowing to a wall* to explain the rate at the end of the equilibration process. The knowledge element introduced by Mateo is very powerful. It can be used to both explain and predict the equilibration curve and maps smoothly to the idea of *difference drives rate*. His explanation in fact precedes the emergence of an articulation of *difference drives rate* that is free of anthropomorphic language and independent of Alvaro's "slowing down to stop" explanation.

Discussion

Seven elements of prior knowledge were identified as potential *resources* for students' construction of a *difference drives rate* model of equilibration. Six of these elements were previously undocumented. The first element appeared to be connected with the previously documented p-prim *overcoming*. Several elements appeared to be connected with a previously documented phenomenological primitive, *Ohm's p-prim*. These results verify existing elements of the *Knowledge in Pieces* model and suggest possible extensions to it.

The broader study in which this work is situated makes a practical contribution through the design of classroom instruction that supports students' construction of a *difference drives rate* model of equilibration. In addition to pattern knowledge and science practices (such as explanation and modeling) outlined by the Next Generation Science Standards, Patterns Class presents students with a novel picture of the scientific enterprise. Much of the science curriculum that has been designed to emulate the practices of professional science has focused on empirical activities, such as inquiry-based learning. While certain versions of inquiry have focused on the theoretical side (White, Frederiksen, Collins, 2009; Lehrer, Schauble & Lucas, 2008), much inquiry curriculum is focused on engaging students in practices of observation, data collection, and representation. Patterns Class includes empirical investigations but the theoretical half of the inquiry process is more heavily weighted.

Moreover, the theoretical half is not focused on engaging students in activities concerned with *existing theories*; rather, it is focused on scaffolding students' construction of *their own theories*. Much instructional time is devoted to theory building activities (such as the whole class discussion that was the focus of this report).

Patterns Class features a strength-based curriculum that values and leverages the prior knowledge that individual students bring to their learning. The responsive nature of the instructional design does not privilege *formal knowledge* over *everyday knowledge*. It supports students in the task of knowledge construction by building on their strengths. Though instruction has been designed with target scientific models in mind (e.g., equilibration as *difference drives rate*), it is meant to support students' construction of their *own* pattern models and strives to treat students as autonomous agents of knowledge construction. The general nature of patterns makes them accessible to students from a variety of backgrounds. Equilibration, for example, can be explored in physical phenomena such as the warming of a glass of cold milk, or psychosocial phenomena such as the dissipation of a strong emotion over time. Students can think carefully about an example with which they personally resonate and construct their model of equilibration on the basis of that example. Because they afford a multitude of diverse entry-points to their exploration, patterns are an excellent object of thought with which to engage students from different backgrounds and levels of academic preparation in abstract thinking and authentic practices of science.

References

- Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of physics*, 50(1), 66-71.
- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the learning sciences*, 13(1), 15-42.
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and instruction*, 10(2-3), 105-225.
- disessa, A. A. (2014). The construction of causal schemes: Learning mechanisms at the knowledge level. *Cognitive science*, 38(5), 795-850.
- Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics*, 68(S1), S52-S59.
- Lehrer, R., Schauble, L. & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development*, 23, 512-529.
- McCloskey, M. (1983). Naïve theories of motion. *Mental models*, 299-324.
- Schweingruber, H., Keller, T., & Quinn, H. (Eds.). (2012). *A Framework for K-12 Science Education:: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.
- White, B., Frederiksen, J., & Collins, A. (2009). The interplay of scientific inquiry and metacognition: More than a marriage of convenience. *Handbook of metacognition in education*, 175-205.

Acknowledgements

I would like to thank Andy diSessa and the Patterns research group at U.C. Berkeley for their formative contributions to this work. I would like to thank Allan Collins for his critical insights into my analysis and feedback on this work. This research was supported by the Spencer Foundation under grant 201100101.

A Knowledge Analytic Comparison of Cued Primitives When Students Are Explaining Predicted and Enacted Motions

Victor R. Lee, Utah State University, victor.lee@usu.edu

Abstract: The Knowledge in Pieces theoretical perspective posits p-prims as an important knowledge element in intuitive reasoning. Because p-prims are a class of knowledge elements developed and abstracted from everyday physical experiences, it seems plausible that immediate physical experiences, both in terms of sensations and actual observations of motion, would cue knowledge in different ways than when those experiences are just discussed as hypotheticals. This paper presents two cases to show that immediate embodied experiences with everyday objects does change which p-prims are cued and how they are deployed by students to explain situations involving motion. These cases come from a corpus of videorecorded interviews with high school students who were asked to explain predicted and enacted motions that involved deliberate sensory engagement with their bodies. Findings suggest that a connection is indeed present, and learners' embodied experiences should be leveraged in future work to support conceptual change in science.

Keywords: conceptual change, embodied cognition, p-prims, knowledge in pieces, intuitive knowledge

Introduction

While there are still various frameworks actively used for analyzing students' intuitive knowledge, particularly in science, several learning scientists have found a "complex systems" approach to modeling knowledge to be especially fruitful (Smith, diSessa, & Roschelle, 1993). This approach, often associated with diSessa's "Knowledge in Pieces" (KiP) theoretical framework (diSessa, 1988) is especially apt for explaining why student reasoning exhibits discernable variability and change in response to changes in the immediate context (Sherin, Krakowski, & Lee, 2012). KiP posits that such dynamical knowledge and its interactions can be productively and precisely modeled as a diverse set of elements of varying form and content. These elements are activated under different situational circumstances to produce explanations, expectations, or predictions of natural phenomena. Conceptual change results from continual exposure to such situations and transforming what set of knowledge elements gets cued in the desired contexts.

My goal in this paper is, with a KiP orientation in mind, to demonstrate that deliberate introduction of embodied, kinesthetic, or sensory experience is an important feature for cuing knowledge elements. Consistent with growing interest among learning scientists (e.g., Hall & Nemirovsky, 2012; Lee, 2015) and cognitive psychologists generally (e.g., Barsalou, 1999), I take seriously the notion that embodied actions and experiences have detectable consequences on how students think and learn. As they do not map well onto existing knowledge representational schemes, they have been underrepresented in intuitive knowledge research. However, I expect that for students who are still developing technical understandings of physics, we should see notable changes in students' explanations when they begin to consider and integrate immediate embodied experiences in their reasoning. Stated by way of example, I hypothesize that describing principles about how an imagined object will move across a surface when a force is applied will yield different thinking than explaining what happens when one is doing the work of actually physically moving an object across a surface. A more robust understanding of whether and how student thinking responds to immediate embodied experience would better inform how we could design instruction that will reshape students' intuitions.

This paper has modest aims. Specifically, I focus on showing the phenomenon of explanation change when students experience specific motions. The descriptive analyses describe what knowledge elements are cued before and after those motions are enacted. In the next section, I briefly elaborate on the KiP perspective. This is necessary to provide a vocabulary for describing knowledge. Then I will describe the methodological approach I used, *knowledge analysis* (diSessa, Sherin, & Levin, 2016), and the data I obtained. Two cases of student reasoning about motion are presented. In both cases, the students cued new knowledge and articulated the importance of immediate embodied experience in shaping how they thought about the motions.

Knowledge in Pieces

"Knowledge in Pieces" (KiP) is used to describe a complex knowledge system consisting of diverse knowledge elements that are dynamically cued and activated to help people function in the world. While the underlying knowledge system posited in KiP has elements that are diverse in both form and content, one kind of knowledge

element that has been identified and heavily discussed, particularly in physics understanding, is the *phenomenological primitive* or “p-prim” (diSessa, 1993). These elements are primitive in the sense that they are thought to behave like atomistic units of conceptual reasoning; they are a specification of structurally simple relational or causal schemas that are often treated as self-evident when invoked. For instance, if a physics novice were asked why standing farther from a bell makes the sound of the bell softer than if they were standing close to it, we would expect a response to be something to the effect of “because that’s just what happens.” In this case, the observation that the sound of a bell seems to diminish or soften as one stands farther away from it does not need nor does it have any further intuitive explanation. This is because sounds exhibit a behavior of *dying away* over time and space, where *dying away* is a p-prim that asserts the eventual diminishing and ending of some effect over time or space. This particular primitive is also applicable to a range of other common situations: standing farther from a heat source, observing the back-and-forth motion of an empty swing eventually coming to a stop, or a rolling ball eventually coming to a stop on a flat surface. Such origins are part of why p-prims are considered phenomenological. They are developed from everyday experiences and observations in the physical world. Other commonly noted p-prims include *Ohm’s p-prim*, which schematizes a set of qualitative relationships connecting increased effort necessary to bring about a comparable effect in the face of increased resistance and *cancelling*, which is schematized as opposing influences that negate one another’s influence. These primitives will appear later in my analyses.

It is important to note that p-prims are not the only knowledge element posited within the KiP perspective. Others have identified “e-prims”, a larger category of *explanatory primitives* that function similarly to p-prims but are not schematizations in the same way (e.g., “Gravity pulls things downward”) (Kapon & diSessa, 2012); *symbolic forms*, which link aspects of external representations to conceptual schemas (e.g., equal sign understood as balance) (Sherin, 2001); *nominal facts*, which are accurate but relatively shallow, general assertions (e.g., “for every action there is an equal and opposite reaction”) (diSessa, 1996); *narratives* such as potential energy converting to kinetic energy as a car moves up and down a roller coaster (diSessa, 1996); and several others. Several of these will also appear in the analyses below.

Embodiment and embodied cognition as relevant to KiP

Embodied cognition serves as counterpoint to earlier models of cognition that have emphasized “disembodied” computational symbols for representing knowing and understanding. While there are several features and variations in what is considered to be embodied cognition (Wilson, 2002) and also some concern that embodiment need not be linked to cognition specifically, the common presuppositions are that we have bodies that have sensory and motor capabilities and that experiences in the world with those bodies fundamentally shape knowledge and action. Stated simply, bodily experiences matter for why we think in the ways that we do.

Of relevance to KiP and to p-prims is the posited basis of how p-prims are developed and identified. In diSessa’s list of 15 heuristics for recognizing p-prims, he states “P-prims are likely to be abstracted in internally evident terms, especially early in development. Thus agency, muscle tension, and so on are likely to be represented in important base vocabulary for p-prims.” (diSessa, 1993, pp. 122-123). This statement and others published elsewhere that tie sensory information to the formation of p-prims (e.g., Kapon & diSessa, 2012) suggest that while the p-prims are knowledge elements abstracted from everyday experience, they still have strong genetic ties to sensory experience. Forces, for example, come to be understood because we feel exertion and tension in our bodies. My proposal is that the connection extends beyond genesis. Rather, ties still exist that lead to cuing of p-prims during immediate sensory experiences.

One additional connection between KiP approaches and embodied approaches is in how information is extracted from readings of a situation immediately at hand. In proposing *coordination classes*, a KiP-based construct for more precisely characterizing the structure and composition of entities we colloquially call concepts, diSessa and Sherin (1998) also identify *readout strategies*, or what is now referred to as *extractions* (diSessa, Sherin, & Levin, 2016). These are “readings” of the present situation, such as visual appraisals of distance travelled used to infer speed or observations of reversal of motion to infer change in force. While visual appraisals have been most often featured in published research, I expect *extractions* can operate from other sensory modalities. For instance, we “readout” weight from pressure against one’s hand or feel impacts against our body. Those should cascade to different knowledge being cued than before such information was extracted.

Methodological approach

The primary method used was knowledge analysis (KA) of video-records (diSessa, 1993; Sherin, 2001). A recent and detailed articulation of the history of and central principles guiding KA appear in diSessa, Sherin, & Levin (2016). Briefly stated, it involves detailed and iterative review of curated video records to identify verbal or behavioral markers that could be taken as evidence of one or more cued knowledge elements as they are used,

with the assumption of dynamic on-line reasoning is being recorded and that reasoning can be modeled. It bears resemblance to microgenetic analysis (e.g., Siegler & Crowley, 1991), but emphasizes articulation of specific knowledge elements and the relations between them.

Data collection

For this study, 11 high school students (all in their junior or penultimate year of secondary education) who had already completed a year of high school physics participated. These students largely came from one public high school in the US that had an explicit STEM focus and a statewide reputation for excellence in STEM performance and STEM extracurricular activities (e.g., robotics clubs, science fair participation, etc.).

Three interview protocols exploring different physics topics were designed using everyday objects that the students likely had encountered in their lives at some point (i.e., athletic balls, toy trains, and a bicycle). The overarching interview structure was for students to predict and explain an object's motion, to enact and experience that motion, and finally to immediately explain why those motions appeared the ways that they did. Interviews were done after school at a university campus. All interviews were videorecorded and lasted between 30 and 60 minutes each. There was some participant attrition over time. In total, 24 interviews were collected. The two interview protocols relevant to the cases in this paper are summarized below.

Interview 1: Projectile motion with athletic balls

For this interview, the students met with the interviewer in a university gymnasium that was used for a variety of indoor sports (such as basketball, volleyball, and badminton). As such, there were a large number of lines painted on the ground, which were used as reference points in conversations later. Several athletic balls were provided and visible to the student. The task for the student was to explain what forces would be involved when each of the balls were thrown underhand "as hard and as far" as the student could throw them across the gymnasium. The three athletic balls most relevant to the case below and the order in which they were presented include: baseball (diameter = 2.9 in, weight = 143 g), tennis ball (diameter = 2.6 in, weight = 57 g), and a yellow foam ball (diameter = 2.9 in, weight = 47 g).

Interview 2: Collisions with a toy train

The second interview involved multiple collision situations for a toy train that was released from the top of a wooden track (Figure 1). Two are most relevant for the case below. For one, a brick was placed in the middle of the track and the interviewer asked what would happen when the train was released and what forces would be involved, with a specific focus on the moment that the train and the brick were interacting with one another. For the other, they were asked to imagine their flat hand was placed with the palm facing the train and to explain what would happen and why when the hand and train interacted. After these were discussed as predictions with explanations, the activities were all enacted with the train being actually released.



Figure 1. The toy train track used in Interview 2 with a brick (left) and with a hand (right) placed on it.

Analysis

Interview records were all transcribed and iteratively reviewed. While "coding and counting" is not a standard expectation of KA work, all interviews were coded in terms of p-prims and selected other knowledge elements comparable in approach to those documented in Sherin, Krakowski, & Lee (2012). The purpose of this coding was not to generate frequency counts (as frequency can be affected by follow-up questions and student speaking style) but rather to serve as annotation for identifying when knowledge elements were cued, to mark linguistic indicators that served as evidence for inferring particular knowledge elements, and to aid in case selection.

Results

Case 1: Discussing how three balls would differ when thrown

The first case involves Carina (a pseudonym), a Latina student at the STEM high school who reported doing "fine" in physics class, although it was not her favorite subject. For this case, I focus on her discussion of forces involved in the motion of the three aforementioned athletic balls. The argument I make is that both her tactile and sensory experience with the balls led her to shift in how she was reasoning about force.

Before throwing the balls

To begin, Carina was asked about the forces involved in producing the baseball's expected parabolic motion. When given the baseball to hold and asked about the motion that would come from throwing it, Carina immediately talked about a *potential and kinetic energy conversion narrative*. Following an initial discussion of potential and kinetic energy, the interviewer asked Carina about force and to focus on what is happening as the baseball is ascending. (Pauses are marked in parentheses, gestures are described in square brackets.)

C: It's cause, it's losing force because it's going up [motions hand upward] and not just going down [motions hand downward], cause when it goes down it gets more force because gravity is pulling on it [grasps at the air and pulls downward], but as it goes up [motions hand upward] things are like, it's losing force because things are trying to pull it back down [makes a grasping shape with hand and pulls it downward].

Int: What is pulling it back down?

C: Like gravity and, like things are making it stop, like the air [extends arm and makes a pulling motion with her hand toward her body] kind of like makes it draft, have a draft so it is slowing it down [places hands near each other and extends arms outward].

In this excerpt, Carina articulates several different ideas. Many features of her emergent explanation are correct. Carina is aware that gravity is involved in causing a decrease in the upward movement, although she is describing force as being something akin to an impetus that is facing resistance. As far as a p-prim activation goes, it appears that because "things are trying to pull it [the ball] down", *overcoming* is a fine candidate as gravity has some agency and is causing the upward force to decrease and gravity eventually brings the ball down. Other p-prims that could be implicated include *force as a mover*, which would explain why upward movement is expected to be associated with a force (even though the only force being applied is gravity). She also notes how "air kind of like makes it draft" and the draft "is slowing it down". This is one of the first mentions of air and some form of resistance that it creates, potentially pointing toward *Ohm's p-prim*, *overcoming*, and related p-prims.

Following this, the interviewer handed Carina the tennis ball so that she held the tennis ball in one hand and the baseball in the other. She was asked which, if either, would go farther when thrown as hard and as far as she could throw them. She responded that the tennis ball would go farther. Her justification for this was as follows:

C: Since it is lighter [raises hand with tennis ball slightly] gravity is pulling down on it less so it can go farther because gravity [makes downward pulling motion with tennis ball] doesn't pull down on it as much because the baseball is heavier.

Carina has ascertained that the tennis ball was lighter and would travel farther. She had honed in on muscle tension (as suggested by her arm movements) to keep the ball upright as an extraction, which she articulated as an inference that the tennis ball was lighter. However, gravity is discussed in terms of producing some resistance on the upward motion of the tennis ball, but by virtue of having extracted lighter weight from holding the tennis ball, she inferred the tennis ball should be less affected. While it may look similar to *Ohm's p-prim* in that she described gravity as giving the same amount of influence, I contend it is a different element that was cued because the lighter ball received less of that influence due to its reduced weight ("gravity is pulling down on it less..."). While not previously named as a p-prim, this qualitative proportionality (Forbus, 1984) of *smaller is less affected* is consistent with being akin to an explanatory primitive. Here, different elements were cued in response based just on extracted weight.

Carina then returned the tennis ball and was handed the yellow ball, which she then held in one hand while holding the baseball in another hand. She squeezed the ball and then held both balls in front of her. When asked which, if any ball, would go farther and why. Carina answered:

C: The yellow ball, again it's lighter. And (6.0) [holds baseball and yellow ball next to one another]. It is a little smaller than the baseball too.

Int: Does being smaller help it?

C: Yeah because it gets- [motions with her hand in front of the yellow ball and pushing fingers in the direction of the ball] air doesn't like press onto it more. Like it gets less drag because it's smaller.

She stated that the yellow ball should go even farther than the tennis ball. While she did not talk about gravity again, her mention of "again it's lighter" suggested that the same justification she offered immediately prior for

the tennis ball should apply. Gravity would not push the yellow ball down as quickly because *smaller is less affected* (applied to weight). However, she held the two balls together and extracted size information based on visual appraisal, also concluding that the yellow ball is visually smaller. When asked if that helped, air is revisited as resistance. Again, it resists what appears to be forward motion based on her gestures, and again involves the *small is less affected* primitive, applied to size. Here, extraction of more information, namely apparent size, led to a slightly different articulation of what behavior was predicted. In total, simply by being given different objects to hold and examine, Carina exhibited some changes in what primitives were cued. This becomes more pronounced and changed more abruptly after throwing the balls.

After throwing the balls

Prior to each throw, Carina was reminded to throw every ball as hard and as far as possible each time she used a new ball. The baseball did not go as far as the tennis ball. The interviewer asked what had happened and why.

- C: The baseball didn't go as far because it was heavier than the tennis ball [pulls downward with both hands] so gravity was pulling down on it more to make it fall to the ground sooner [continues to pull downwards with hand]. And, um, the tennis ball is also-is also smaller so it has less of an air drag than the baseball.

In this explanation, much of what Carina said was consistent with what she had said earlier in her explanations for the three balls above. Gravity was pulling the tennis ball down less than the baseball because *smaller is less affected* (applied to weight) by gravity's pull. She also added that the size, which she had not extracted before but noted this time as being salient, was different in that the tennis ball was smaller. Therefore, *smaller is less affected* was applied to size, which would be less affected by air drag. Given less resistance from both air and gravity, it was sensible for her to conclude why the tennis ball went farther. She then threw the yellow foam ball, which did not land as far as the other balls. When asked about what happened, she replied as follows.

- C: Well, since it was lighter – um, the, it wasn't pushing on the air as much as the tennis ball and baseball. Also, um (4.4) well, that is like the major difference. (2.1) The tennis ball (3.0) I can't really think of it now that they're out of my hands, but um, (3.5) the yellow ball went shorter because ...it had more air [opens and closes hands] it couldn't go through the air as easy as the tennis ball and the baseball because it was lighter than them [brings hands together] and it caused more of an air drag and it fell to the ground sooner [pulls hand downward] because it couldn't push [makes swiping motion with hand] through the air as easier.

Of note is her explicit statement, "I can't really think of it now that they're out of my hands," indicating that the immediate sensory information mattered to her. Moreover, in trying to reason through it, Carina had extracted a weight difference. Size was no longer a high priority extraction, and she said the lighter weight made it less able to move through air. This conflicts with what she had said for the tennis ball just prior and for the yellow ball before any throwing. After throwing, *Ohm's p-prim* was cued to explain the reduced distance. In this moment, the sensations she experienced handling and throwing the yellow ball, combined with it landing behind the other balls, led to a shift in what information was extracted and what p-prim was cued to explain the motion. Before the throw, the yellow ball should be less affected because it was lighter. After, the yellow could do less to affect things because it was lighter.

Case 2: Discussing differences in toy train collisions with a brick or a hand

Another student from Carina's school, Isaac, was noted by his peers as an exceptionally strong physics student and science student generally. This case examines how this student responded to questions in interview 2, involving the toy train and collisions with a brick and with his hand (a difference in direct sensory experience).

Before releasing the train

When first shown the brick and asked about what would happen, it was quite evident that *bouncing* was an active p-prim because he explicitly said so: "it [the train] might bounce and stop." Follow up questions were about force and what happened when the train and the brick were interacting with one another at the bottom. (// denotes overlapping speech)

- I: Um, the train is exerting force on the brick. The brick is exerting an equal amount of force. The brick might be (0.8), yeah they are exerting an equal amount of force on each other [shrugs]

- Int: So this is exerting force on the brick [lifts train and hold it near brick] and the brick is exerting a force on the train.
- I: Mm hmm.
- Int: Would that, umm..(1.0) should we expect to see anything happen as a result of that?
- I: The train might bounce back or it might just stop.
- Int: Does one of those seem more likely or less likely to you?
- I: Bouncing back.
- Int: And can you explain how force is involved in making that bouncing back //if it is involved?
- I: //Um...(2.0) The brick isn't completely firm [makes a fist and brings it in toward his body, then releases and lowers] so it would give way, but, er, the-it's actually probably the wood on the train. The train would give away a bit and spring back to its original position exerting whatever force caused it to compress back on the brick [makes fist], pushing the train back [unclenches fist and motions sideways away from brick].

In this transaction, Isaac immediately gave a nominal fact, *equal and opposite reactions*, (diSessa, 1996) consistent with what is frequently said in physics classes in relation to two objects coming in contact with one another. However, when Isaac was pressed about what he thought was a more likely outcome from those two forces, *bouncing* seemed to have a slightly higher activation. Part of the *bouncing* mechanism was unpacked as involving *springiness*, a p-prim involving an object compressing and then returning to its original shape. This was assigned to the wooden train through a perceptual extraction and comparison of firmness between brick and wood. Implicit in Isaac's responses are also the e-prim that *gravity pulls things downward* (Kapon & diSessa, 2012) and also a p-prim for *guiding* that created an expectation of the train staying on track. As will be discussed later, *guiding* was violated when the motion is enacted and the train falls off the track.

For the situation involving predicting what would happen with a hand being placed where the brick was, Isaac had a different expectation even though he invoked the *potential and kinetic energy conversion narrative* and *equal and opposite reactions* nominal fact again.

- I: The train would slide down the track, hit my hand. The train would transfer all of its kinetic energy into my hand [opens hand and jerks it quickly away from the train]. The train would stop, my hand probably wouldn't move [lifts open hand, glances at it, and returns it to lap] because the train isn't very big.
- Int: Okay. Um-and then what's going on with force when the train and your hand [points two open and separated hands toward the center] are interacting with one another?
- I: The train is exerting a force on my hand [extends open hand on lap], my hand is exerting an equal amount of force on the train [jerks open hand forward], which is what causes it to stop.

Starting with the "hand probably wouldn't move because the train isn't very big" implies extractions and comparisons of size of colliding objects and potentially some cuing of *overcoming* or *Ohm's* as a p-prim where the train cannot overcome the size and entailed resistance of his hand. When asked to talk about force specifically, his expectation was for the train to stop – a different expected outcome than what would happen with the brick. The *equal and opposite reactions* nominal fact led to cuing of two forces that *cancel* one another, thus leading the train to stop.

After releasing the train

What happened after the train was released and collided with the brick and the hand was that the behaviors of the train after collisions were quite different from what he had expected. With the brick collision, the train toppled over. Isaac gave the following explanation for what happened.

- I: Ok. It run [sic] down the track collecting kinetic energy and then it hit the brick, um, putting all its kinetic energy into the brick. The brick didn't move, um, and the train stopped moving and fell sideways.
- Int: Okay, um, so what was happening in terms of force at the bottom right when they're in contact? [releases train from top again, it collides with the brick and falls again]

- I: The train is exerting a force on the brick, the brick is exerting an equal amount of force on the train, um, and I guess the reason it fell off is because the train was still trying to exert and go forward but it couldn't go forward so it went sideways.

Following the *energy transformation* narrative, Isaac observed that the brick had no obvious movement, but he noted that the train had fallen over. When asked about the involvement of force, Isaac again gave the nominal fact of *equal and opposite reactions*. Expectations set by the unarticulated *guiding* p-prim were violated since the train was no longer on its track. Surprisingly, even though he had talked about equal amounts of force, he attributed some agency to the train as “still trying to exert and go forward”, suggesting that there was some sort of *deflection* – something that reroutes a moving object following contact – that had taken place. This is different from his earlier explanation in terms of *bouncing* and *springiness*. Ultimately, the unanticipated event led him to activate a new set of primitives to explain the deviation, suggesting that while he still cued the nominal fact of *equal and opposite reactions*, a different set of elements were brought to bear. Similar shifts took place when the hand replaced the brick. When the train collided with his hand, it stayed on the track and rolled backward a few centimeters before stopping.

- I: The train came down the track collecting kinetic energy. When it hit my hand it transferred that force into my hand. My hand exerted an equal amount of force back. I guess my hand compressed or transferred some of that energy back into the train causing it to slide backwards. My hand, I felt the impact, I shifted back a little bit. Not very much.

Int: Your hand shifted back?

- I: Yeah, or I felt it trying to shift it back. I was trying not to move my hand so it didn't move. The force from the train exerted force on my hand and my hand tried to respond and move in the direction the force was going.

Int: Ok. Can you feel the force from your hand pushing back the train? Is that something that-

- I: Yeah. I can feel my hand pushing back.

Int: Like-how?

- I: You just kind of feel the impact and the compression I guess. It is just something that is there.

Int: Ok. And that is just in the palm of your hand?

- I: Yeah I feel something bouncing off of my hand.

Equal and opposite reactions was cued yet again, although he seemed to be exploring two possibilities. The first was that his hand had compressed, suggesting cuing of *springiness*. Note that he had not previously talked about nor explicitly registered *springiness* as being related to his hand before. He also considered that the energy *transferred* back to the train from his hand, where his hand was some sort of conduit (“transferred some of that energy back”). Isaac also offered a description of feeling his hand “trying to shift” and how his hand “tried to respond.” In follow up questioning, he struggled to articulate what “pushing” felt like beyond the “impact and the compression.” However, thinking about those sensations seemed to favor greater cuing of *springiness* and potentially also *bouncing*. In sum, when he physically experienced the collision and reflected on what he had felt, along with the observed reverse motion of the train, Isaac cued a new set of primitives to think through the situation than when he had been simply predicting the motion before. Before the hand collision, he expected *cancelling*. After the hand collision, he thought in terms of *bouncing*.

Discussion

Through these two cases, of which others also exist in the larger interview corpus, I have sought to demonstrate that for students who have already had formal exposure to physics, the knowledge that is cued can shift depending on what is physically experienced and enacted. In some respects, changes in knowledge activation and cuing are to be expected for relative novices in physics learning according to a KiP perspective. However, this paper has been a deliberate and newly focused inquiry into some of the dynamics of how immediate embodied experience and sensory extraction can shift what knowledge is cued.

While there is much more to do in the future to understand the broad range of knowledge dynamics involved, we can begin to speculate on the importance of such findings. As much research has shown, students can perform well in formal instruction and assessment tasks but struggle with intuitive reasoning that involves hypothetical and everyday situations. That happened here. If engagements with everyday situations are where and how the basic elements of our physical intuitions develop, it makes sense for us to understand how formally taught

knowledge is used, if at all, in such situations. At home, toy trains collide and fall off of wooden tracks and we feel objects pressing against our hands. It is not clear that those experiences are being brought into coordination with targeted understandings from school physics. Also unclear is how scientific epistemologies associated with repeated and designed experimentation interact with what students immediately perceive and feel as routine bodily experience.

Still, we can and should consider how we might bridge familiar everyday experiences encountered outside of designed experiment with understandings and models derived from science. While they try, traditional materials do not succeed with this (Lee, 2010). Supportive efforts could be made to deliberately help students to situate science in messy, lived, and sensed experiences. Thus far, a few notable efforts are beginning to be developed in this area (Pauw, et al. 2015), and some inquiry-oriented curricula attempt versions of this as well. However, a way to transform learning of disciplinary content could just involve taking mundane daily experiences and gradually unpacking the physics involved. While intuitive physics conversations attempt to do this, letting bodily sensations be recognized and discussed would be a novel addition. Orchestrating discussions of science content and everyday sensations, experiences, and observations is still a formidable task. However, as more research and development concerned with embodied experience proceed, we will hopefully see efforts that transform what students encounter and feel in everyday life into more robust learning of scientific ideas.

References

- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(04), 577-660.
- diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age*. Hillsdale, N. J.: Lawrence Erlbaum.
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10(2&3), 105-225.
- diSessa, A. A. (1996). What do "just plain folk" know about physics? In D. R. Olson & N. Torrance (Eds.), *The handbook of education and human development : new models of learning, teaching, and schooling* (pp. 709-730). Cambridge, MA: Blackwell.
- diSessa, A. A., & Sherin, B. (1998). What changes in conceptual change? *International Journal of Science Education*, 20(10), 1155-1191.
- diSessa, A. A., Sherin, B., & Levin, M. (2016). Knowledge Analysis: An Introduction. In A. A. diSessa, M. Levin, & N. J. S. Brown (Eds.), *Knowledge and Interaction: A Synthetic Agenda for the Learning Sciences* (pp. 30-61). New York, NY: Taylor & Francis.
- Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence*, 24, 85-168.
- Hall, R., & Nemirovsky, R. (2012). Introduction to the Special Issue: Modalities of Body Engagement in Mathematical Activity and Learning. *Journal of the learning sciences*, 21(2), 207-215. doi:10.1080/10508406.2011.611447
- Kapon, S., & diSessa, A. A. (2012). Reasoning Through Instructional Analogies. *Cognition and Instruction*, 30(3), 261-310. doi:10.1080/07370008.2012.689385
- Lee, V. R. (2010). Adaptations and continuities in the use and design of visual representations in US middle school science textbooks. *International Journal of Science Education*, 32(8), 1099-1126. doi:10.1080/09500690903253916
- Lee, V. R. (Ed.) (2015). *Learning technologies and the body: Integration and implementation in formal and informal learning environments*. New York, NY: Routledge.
- Pauw, D., Clegg, T., Ahn, J., Bonsignore, E., Yip, J., & Uchidiuno, J. *Navigating Connected Inquiry Learning with ScienceKit*. Paper presented at CSCL 2015. Gotheburg, Sweden.
- Sherin, B. (2001). How students understand physics equations. *Cognition and Instruction*, 19(4), 479-541.
- Sherin, B., Krakowski, M., & Lee, V. R. (2012). Some assembly required: How scientific explanations are constructed in clinical interviews. *Journal of Research in Science Teaching*, 49(2), 166-198. doi:10.1002/tea.20455
- Siegler, R. S., & Crowley, K. (1991). The Microgenetic Method: A direct means for studying cognitive development. *American Psychologist*, 46(6), 606-620.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the learning sciences*, 3(2), 115-163.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625-636.

Acknowledgments

This study was made possible by a National Academy of Education/Spencer Foundation Postdoctoral Fellowship. Thanks go to Joel Drake for his assistance with data collection, Michelle Berry for transcription, and the participating students who graciously agreed to meet during their free time for this research study.

Teacher Learning in a Professional Learning Community: Potential for a Dual-layer Knowledge Building

Seng Chee Tan, Nanyang Technological University, sengchee.tan@ntu.edu.sg
Shien Chue, Nanyang Technological University, shien.chue@ntu.edu.sg
Chew Lee Teo, Ministry of Education, Singapore, TEO_chew_lee@moe.gov.sg

Abstract: This study is situated in the field of knowledge building and teacher professional learning community (PLC). It describes a case study of eight elementary school teachers working on lesson design using knowledge building pedagogy to enhance student learning. The research question is: “How is knowledge building accomplished among teachers within a PLC?” The main method employed is the analysis of teacher’s discourse in the PLC. Results indicate that although not cognisant of it, these teachers functioned as a reciprocal layer of knowledge builders over their students’ work. Critically, teachers’ knowledge building was enabled by working with authentic classroom problems, embracing idea diversity and demonstrating epistemic agency in their knowledge advancement. Future work can focus on building teacher’s awareness of this layer of knowledge building and scripting creation of knowledge artefacts among teachers to mediate and record their collaborative inquiry.

Keywords: knowledge building, teacher learning, professional learning community

Introduction

This paper is derived from a study on teacher learning as the teachers engaged in advancing their students’ knowledge building. It uncovers a tight coupling of teacher and student knowledge building – teachers engaging in building knowledge on their professional practices while attempting to design for their students’ learning using knowledge building pedagogy. This study is an attempt to address two intertwining research gaps in knowledge building and in teacher’s professional learning community.

Knowledge building is generally understood as a discursive activity intended to enhance collective understanding (Bereiter, 2002). Members in a knowledge building community construct and progressively improve on ideas through knowledge artefacts. Knowledge building is based on the assumption that members engage in specific discourse activities where content of the discourse is related to the construction of coherent and consistent knowledge (Greeno, 2003). Several principles, or what Scardamalia (2002) regarded as socio-cognitive determinants, are needed to support knowledge building. Idea-centric approach and participants’ epistemic agency, are principles relevant to this study. In knowledge building, teachers focus on helping students learn through collaborative improvement of ideas on a topic and following the trajectory of idea development, rather than trying to cover topics in a pre-determined sequence. To achieve that, students need to work on authentic knowledge problems that arise from their attempts to understand the world; they need to identify knowledge problems and articulate them as ideas. In a group, there will usually be diversity of ideas among members because of their prior experience and their different strengths and expertise. This provides the natural impetus for discussion about the ideas, leading to collective improvement of their ideas. By improvement of ideas, we mean improving the coherence, quality and utility of ideas. Continual engagement in idea improvement process may lead to rise above, when students are able to integrate several ideas or frame the inquiry using a higher level principle or theory. This process of negotiation requires enabling conversational moves such as asking questions, and making statements elaborative or regulatory in nature. It also requires the participants to take collective responsibility for learning what they need to know as they engage in deep discussions centered on problematized content. Students show epistemic agency when they display ownership of their collective inquiry: using knowledge criteria to negotiate a fit between their own ideas and those of others and use the differences to catalyse knowledge advancement. Much research has been conducted with respect to students’ experiences in knowledge building (Scardamalia & Bereiter, 2010), yet this cannot be said for teachers doing knowledge building as part of their professional learning.

In terms of research gap in teacher professional development, traditionally, the predominant mode involves formal courses, workshops or seminars offered by external agencies, which has limitations in terms of the real impact on how teachers can address problems encountered in their classrooms (Lieberman & Mace, 2008). This mode of professional development may lack nuanced sensitivity to the classroom challenges and does not provide contextualised learning opportunities for teachers to reflect on and improve their classroom practices (McLaughlin & Talbert, 2006). In recent years, professional learning community (PLC) (DuFour et al., 2006) is

gaining traction for its affordances for teachers to engage in collaborative investigation of classroom practices specific to a school setting so as to enable more direct impact on student learning (McLaughlin & Talbert, 2006). However, research is relatively silent on how teachers learn in PLCs that can transform student learning. Using cultural historical activity theory as a lens, Lee (2015) studied two PLC activities in a Singaporean elementary school where 13 teachers engaged in book study and lesson study over two semesters. Lee reported several disturbances to teacher learning in the book study sessions, which include the lack of preparation for assigned readings, teachers' reservations about the ideas in the selected book, teacher's struggle with competing activities in the school, and extensive focus on content coverage and recall rather than application of the ideas. Teachers were more engaged in the lesson study sessions, but there were still disturbances attributed to challenges in logistical arrangement for observation and post-lesson colloquia.

There are potential values in integrating teacher's knowledge building and professional development. Currently, existing cultures and discourse communities in many schools do not value or support critical and reflective examination of teaching practice (Putnam & Borko, 2000). Educational reformers have argued that for teachers to be successful in constructing new roles, they need a platform to participate in a professional community that "discusses new teacher materials and strategies and that supports the risk taking and struggle entailed in transforming practice" (McLaughlin & Talbert, 1993, p. 15). PLCs offer fertile ground to study knowledge building practices in teacher community. When teachers come together in a discourse community, they draw upon and incorporate each other's expertise to create rich conversations that offer new insights into teaching and learning. As teachers share their expertise in a PLC, they construct new knowledge about instruction and content (Little, 2003; Stoll et al., 2006). In fact, collaborative work in teaching involves problem posing and the articulation of practice (Horn & Kane, 2015). Teachers are positioned to learn from talking with colleagues and there are opportunities for learning constituted in teacher workgroups.

This study examines teacher learning opportunities and possibilities residing in teachers' PLC meetings and addresses the research question: "How is knowledge building accomplished among teachers within a PLC?" There were studies about how teachers and facilitators work together for knowledge building (Orland-Barak & Tillema, 2006), but less is written about how knowledge building proceed when teachers come together in a PLC to improve classroom practices for enhancing students' learning. Importantly, the teachers in this study adopt knowledge building pedagogy (Scardamalia & Bereiter, 2015) that focuses on engaging students in solving scientific problems and in writing. With the current gap about knowledge building in PLCs, our work aims at contributing to the understanding of how PLCs afford opportunities for teacher learning and innovation in knowledge building practices. Second, focusing on the content of teacher talk, it offers insights into how conversations may be a source of knowledge building for teachers. Overall, this study has the potential to update collective repertoire of practices (Allaire, Laferrrière, & Gervais, 2011) as we surface how principles of knowledge building can supply intellectual, social and materials resources for teacher learning and innovations in practice.

Methods

Instrumental case study (Stake, 1995) approach was used in this study as the focus was on gaining insights into how teachers learn as they participated in a PLC rather than an intrinsic interest in the PLC per se. Examination of the interactions is critical, thus, video recording was used as the main data collection method. We video recorded the PLC meetings by directing a video camera at the teachers seated around the meeting table to capture their talk. In this way, we were also able to zoom in on teachers' gestures as well as images projected on a projector screen in front of the meeting table if any. This resulted in 20 hours of video-recorded meetings that captured the interactions between the teacher participants of the PLC meetings. Pseudonyms are used in this report as a preemptive measure to protect the identities of the participants.

Participants

The research team worked with eight elementary teachers from Future State Primary School to incorporate knowledge building principles in the design of their elementary science and English lessons. The teachers have been teaching from 2 to 15 years. The teachers met once weekly for two hours to design lessons. These meetings were facilitated by Cindy, a former secondary school teacher, who has cumulated more than 10 years of experience implementing knowledge building pedagogy, as well as working with multiple schools on implementing knowledge building principles. As part of the field support, Cindy first provided training for teachers on the use of Knowledge Forum and its associated learning environment before they met weekly to craft their lesson designs. Importantly, Cindy catalysed collaborative knowledge building for the teachers; and as teachers became more proficient with the pedagogy, her scaffolding faded to an observer role. As a result, the PLC meetings were participant driven as the teachers engaged in the tasks of teaching, assessment, and reflection. Their meetings

were collaborative and interactional in nature (Palinscar & Brown, 1984; Bereiter & Scardamalia, 1989) which further supported the work of the teachers as situated within a PLC.

Analysis

Video recorded data of teachers' weekly PLC meetings were transcribed verbatim. We first read and classified the content of transcripts according to procedural and conceptual discourse. Procedural talk includes discussion on logistics of lessons, venues, and distribution of work. Conceptual talk encompasses teachers' individual reflection, group discussion on how to teach particular a Science or English topic. The start and break of talk segments are marked by teachers announcing their intentions or listing their agenda for the meetings. Focusing on language in use, that is, how teachers used talk to thread lesson design for classroom implementation, we analysed the conceptual discourse that was available for making sense of teachers' knowledge building process.

Next, we created event summaries of the conceptual talk occurring during each PLC meeting to document the instructional topics discussed. This provided a macro view of how the teachers designed Science and English lessons using content topics drawn out from students' knowledge building posts. These event summaries documented teachers' design of knowledge building lessons for their students over time. Importantly, these event summaries identified episodes of interactions which positioned the function of talk against the instructional plans of the teachers (Edley, 2001). Following the principles of interaction analysis (Jordan & Henderson, 1995), we met repeatedly to discuss emergent meanings about the purpose of the teacher talks. We focused on how an idea from one teacher would influence the next teacher who speaks. We were motivated to understand how knowledge building principles may be present in the talk of the teachers. With such a focus, we read the weekly transcripts to familiarize ourselves with the many content topics that were covered. We also examined how a particular topic was developed and built up over the weeks. During the intensive group discussion of the data, we asked questions such as: What is the trajectory of learning across the topic of system? How did the teachers plan on starting the topic of "System"? How did the discussion on "Systems" change mid-way? Did they begin with definitions or questions or case examples to trigger questions from their students? What happened after each teacher shared their individual teaching idea? How did the English teacher respond to the other English teacher's teaching ideas? Did the teachers revisit the questions post by their students? Why are the English teachers focusing on stimulus based conversations as a start for the use of knowledge forum? What is the progression? Were more topics introduced along the way? Such questions were useful in focusing our attention on the discourse of teachers during repeated readings of the transcripts. As our comments and assertions evolved, so did our analytic approach. Ultimately, we applied Scardamalia's (2002) socio-cognitive determinants of knowledge building to identify how the teachers engage with knowledge building during their PLC meetings and all extracts were analyzed by the researchers until a common agreement about the interpretation was established. We then chose representative episodes of knowledge building where instructional topics were designed and subsequently implemented in the classroom to present teachers' work of knowledge building during their professional learning meetings.

Findings

We selected two vignettes of teachers' discourse that could illustrate progression in teacher's pedagogical approach across two subjects: Science and English.

Idea-centric rather than topical approach

This vignette illustrates how Cindy coached the teachers to focus on advancing students' ideas, a key principle of knowledge building, rather than sequential progression across topics to cover the official syllabus.

Relating students' posting about states of matter arising from the topics of life cycles, Patty highlighted the learning trajectory of her students and waited for reactions from other teachers.

Patty: I have managed to complete the comparison between life cycle of frog and mosquito. But I notice the children are coming to a point where they are talking about matter, changes in state in KF such as water becoming water vapour.

However, Cindy had other concerns. She had wanted to understand the sequence of Patty's classroom events leading to the posts about water and water vapour.

Cindy: Turn back a bit, with the cloud and water, what happens next?

Patty: They had managed to see water droplets forming during the experiment. Some said it is gas, students said the hot air goes up and then the water vapour is formed inside. That was when I

told them to go on to KF to ask their questions and the students have done that. But I told them to also hold it there as I want to develop the other section of their postings which is related to life cycles rather than the cloud experiments that they have done.

Patty was referring to the “Cloud in the Bottle” experiment that aimed at simulating formation of cloud in a plastic bottle using warm water. Cindy highlighted her concern with students’ understanding of clouds and waters under the bigger topic of system. This difference in knowledge problems persisted as Patty announced her intention to tap on students’ posts about cloud and water for the bigger topic on States of Matter in future lessons. Countering Patty’s ideas that developing students’ understanding could wait, Cindy reminded Patty that the original purpose of the “Cloud in the Bottle” experiment was to help students understand the components of a system. Cindy feared missing the opportunity to develop students’ understanding of a concept even though it was not required at such depth in the P4 science syllabus. She surfaced the problem more explicitly by asking if students had provided any conclusion about whether cloud was not a system.

Cindy: Did the students come to a conclusion? Is it important for them to have some sense of whether the cloud is a system after all the explanation?

Patty: They talked about life cycle and told them to move on about life cycles. They did not come to a conclusion about whether cloud is a system.

Cindy: But it is important for them to have some sense after their explanation.

Before Cindy could complete her sentence, Patty counteracted to explain that while students did not come to a conclusion, she was satisfied that her students were able to see the process of water cycle as being embedded in the cloud experiment. Not giving up on her idea, Cindy suggested to Patty that she could ask her students to write about whether cloud was a system so as to reveal students’ understanding of system. The momentum of discussion continued towards improving of ideas when Patty acknowledged Cindy’s concern and suggested a better way of actualising Cindy’s suggestion:

Patty: But of course, on the other end, I want students to also know that system is made of many parts working together. Rather than asking students if cloud is a system, to me, I will ask them why do you say cloud is a cloud system inside the water cycle system. Because these students think that it is a system within another system.

Sensing Patty’s receptive attitude toward her idea, Cindy quickly suggested that they could pull out a few notes on the Knowledge Forum and to get students to support their writing of their notes with more evidence. At this juncture, Emma, another teacher in the PLC group, revealed that her students had been posting their ideas about whether clouds belong to a system on the Knowledge Forum. Turning their focus on the students’ posts in the forum, the three teachers read the notes of students together and noticed that while some students concluded that clouds do not belong to a system, many more students were still holding on to the misconception that cloud is a system. It is interesting to note that in this process, Emma and Patty noted that students would finally realise that cloud belongs to the water system once they proceed to the topic of State of Matter. This helped Patty to find her justification to move on the topic of State of Matter.

Emma: It (cloud) is part of the water system. So that is why we (students) have not got it. Because I think once they do the topic of change of state, then, they will realise that water is after all the same thing.

Patty: So that is why I want to move into matter using these notes. It is changing from one solid to a liquid or liquid to gas. And slowly students will see that it is the same water going up and coming back down and up and it is not many parts.

From resisting Cindy’s idea to accepting and improving on Cindy’s idea, Patty and Emma moved the conversation towards development of knowledge advancement. Aligning their diverse ideas over the problem of students being uncertain with the concept of system, the teachers made advancement in their lesson design by focusing on students’ posts and offering concrete ways of helping students clarify the components of a system. The concerted conversation of these teachers thus became a collective knowledge which resulted in Patty asking students to write out their thoughts about whether cloud was a system in the following day’s lesson. Following

knowledge building principles, the teachers offered diverse ideas to deal with the authentic problem of student's understanding and they collaboratively improved on the lesson design ideas.

Rising above current practices

During the 8th PLC meeting for English lessons design, three teachers first recalled verbally how they had in the previous week helped students on a continuous writing task on the topic of a snatch thief through the Knowledge Forum. Kenny took the lead to share how his students managed to conceive greater details supporting the building up of the story:

Kenny: Students are supposed to come up with one idea first and the other groups are supposed to look at ideas and ask questions based on that particular idea. So we have some examples such as this robber was caught by the police in the end and some students would ask how the thieves were caught by the police. Students also came up with ideas about the type of punishment the robbers would get.

These ideations were concluded by Kenny as useful in helping students focus on their writing task. However, this initial knowledge building effort was negated when Kenny revealed that his students were given a sample writing subsequently so that they could submit an individual piece of writing. In short, the lesson reverted to the traditional mode of scaffolding with model essay.

Next, Priscilla shared a different model for teaching the writing task. She had first paired up with another teacher (Tim) to set up a thread in the Knowledge Forum to demonstrate to students how they can build on their peers' notes. For example, when Tim suggested on the Knowledge Forum that the thieves were getting caught, Priscilla would ask how and where the thieves were captured and at the same time offer the suggestion that the police should be roped in for the plot. Guiding students through examples of how they can post their thoughts and questions on the forum, Priscilla modelled for students how they could develop their story plot. Unfortunately, after this modelling, it was back to the usual task of writing as Priscilla gave students sample stories to read after their discussion and students were instructed to submit their individual written work. Finally, Patty shared her method of teaching. She announced that her class was not expected to produce writing with different introductions. However, students were "expected to start with dialogue." Patty justified that within the framework of a dialogue, students could think and articulate in Knowledge Forum how their dialogue could be more exciting and fun. Patty also shared that within the constraint of working only with dialogues for story starters, students were able to generate more vocabulary and scenarios.

As these teachers replayed their instructional acts during the PLC meeting, they seemed to reveal a common trend that students all ended up with individual writing tasks. The messiness of students' posts was not refined and the diverse ideas were treated as alternative ideas. Students were in fact guided through deeply entrenched traditional ways of writing.

The opportunity to break away from current practices came about when these teachers attended a conference on Knowledge Building in Hong Kong. Upon their return, the teachers shared new insights into the ways of doing knowledge building in the classroom during the 10th PLC meeting. As revealed by Patty:

Patty: We were too stuck in looking at mechanics of writing such as the types of words used. There is no in-depth exploration of the writing. The Hong Kong representatives commented the pictures we used for composition were so restrictive. They proposed we think about doing KB with a thematic slant. For example, we can do KB on a story hand out, so that they can build on their knowledge in the theme, and the subsequent writing piece can be on the theme. In this way, when students do the writing, they can reflect on the themes which they had ideas built upon.

With knowledge gained from the conference, these teachers worked towards higher-level formulation of the writing problem. This was visible as Kenny next announced that the English teachers were going to plan for a writing task on the abstract theme of *Friendship* as mandated by the lower primary curriculum. Critically, the lesson designs were planned to help students build on their own knowledge about friendship:

Kenny: My initial lesson design requires students to make a stand regarding the story 'The Four Friends'. They will read and reason if the story is a good story about friendship and why. In this way, we will have all kinds of ideas from the students and after that we hope there can be one question where they can build on and then they can tie up with their postings. This would help them in their writing and build on their own knowledge of friendship in their writing.

While Kenny was hopeful that students may pose question such as “I want to know what friendship is all about”, which has scope for other students to build on, he was also aware that such a scenario may not pan out. Hence, he suggested that after having students’ questions up on the KB wall, they can revisit the theme with another reading about friendship and teachers can ask the same question “is this story a good story about friendship?” As students were expected to give their ideas, teachers subsequently can help them do comparison of ideas and await questions that can tie up with the theme of friendship.

Aligning her ideas with Kenny’s, Patty added her own views of how the theme of friendship can be further developed:

Patty: We intend to come up with the question, ask them an argumentative kind of question whereby they take a stand on a story, the wolf story that they had read last term. The story takes a different perspective of the actual story, so we want our children to take a stand, to decide is the wolf is as innocent as it seems or if the three little pigs are at fault. We hope to tap on the moral of the story which is to be truthful. After that students can go on to the composition about the taxi driver who had passengers leaving their wallets behind in his taxi after alighting. I mean students would have gained the knowledge of being truthful and use the vocabulary generated during posting of their views about the wolf story to write their stories.

Kenny immediately added that Patty could even give students different snippets of stories on truthfulness which will help give students even more ideas for discussion. This exchange captured how teachers rally around the improved idea of providing students with thematic writing tasks. As evident from Patty’s comments above, she had extended the writing task to tackle more complex issues such as truthfulness. Subsequently, realising the challenge this new way of teaching may pose for her students, Patty voiced:

Patty: I actually need suggestions because based on this argumentative task that we are thinking about, students may not come up with questions or examples. This is their first time and they are only at K3.

Not afraid of voicing her fear as a more experienced teacher, Patty felt safe in revealing her uncertainty about this new approach. Supporting Patty’s emerging instructional goals, Kenny promptly suggested they could devise a template with scaffolding to help students along in writing down their stand about the stories read. In this way, students could also exchange their writings on the template as notes to help each other along. By offering such a solution that requires a change of the existing scaffolds in the Knowledge Forum, new syntheses was derived at the KF front. By moving to a higher plane of how writing can be taught, the teachers abandoned old methods of picture style writing tasks (the snatch thief) as they embraced a more complex way of teaching thematic writing that can potentially lead to different instructional outcome for different teachers. In the vernacular of knowledge building, the teachers achieve *rise above* from their initial ideas about writing, and display collective cognitive responsibility in contributing to the joint effort in lesson design.

Discussions

The results of our analysis locate knowledge building principles as embedded within conversations of teachers during their weekly PLC meetings. Moving from dealing with authentic problems to embracing idea diversity, Patty and Cindy were initially observed to be reacting to each other’s ideas. However, by subsequently taking a collaborative mode to anchor their discussion on how students’ posts about clouds and rain can be addressed, Patty improved on her initial idea. As a result, diversity of ideas surrounding students’ posts enabled Patty as well as Emma to collaboratively construct an instructional plan for the topic of system different from the initial lesson design plan for the topic of matter. Similarly, over a span of two weeks, Kenny, Priscilla and Patty demonstrated attempts at rising above their initial conception of cookie-cutter type of writing lessons to embrace a thematic approach. This was especially salient when teachers improved on each other ideas on how the theme of friendship could be developed by their students on the Knowledge Forum. Ultimately, they engaged in knowledge building discourse that moved students from simplified writing lessons to a thematic one.

In terms of the addressing the research gap in knowledge building, the two examples in this study illustrate the principles of Knowledge Building (Scardamalia, 2002) in action leading to production of new knowledge for teaching during PLC meetings. These new knowledge include innovation as seen from the way thematic writing tasks were implemented as well as permanent advancement of ideas as evident from instructional designs for topic of *Systems*. Critically, these teachers engaged in the trajectory of knowledge building as they

attempted to design lessons for students to engage in knowledge building. Using the lens of socio-cognitive determinants of knowledge building (Scardamalia, 2002) to analyze the teachers' talks, we uncover a reciprocal layer of knowledge building in action (Table 1). The teachers were engaging in building knowledge on their professional practices while attempting to design for their students' learning using knowledge building pedagogy.

Table 1: Dual-layer of knowledge building

	1 st Vignette	2 nd Vignette
Teacher's reciprocal layer of knowledge building while designing for students' knowledge building	<ul style="list-style-type: none"> • Authentic problem of advancing student's knowledge of "Systems" • Diversity of ideas leading to better ideas on how to enact idea-centric pedagogy • Teachers showed epistemic agency in reasoning and rationalizing their actions 	<ul style="list-style-type: none"> • Authentic problem of enhancing students' essay writing skills • Diversity of ideas on how to scaffold students in essay writing • Rise above from cookie-cutter framing of students' ideas to thematic brainstorming of ideas as resources for essay writing • Teachers showed epistemic agency in learning new approach and suggesting ideas to design lessons based on the new approach
Student's knowledge building	<ul style="list-style-type: none"> • Authentic problem on whether cloud is a system 	<ul style="list-style-type: none"> • Ideation for essay writing

However, a critical question remains: Can teachers continue to distill their learning during PLCs and translate it into the classrooms and stay faithful to knowledge building principles? First, the teachers may not be aware of this tight coupling of their knowledge building on their professional knowledge while working on lesson design for their students. In short, they are not cognisant that they can apply the same knowledge building principles to enhance their own professional learning. For example, the teachers may not be aware that they can in fact use Knowledge Forum to augment their discussion. By creating their knowledge artefacts in the online forum, these artefacts can mediate their discussion and serve as record of their innovation. Thus, in further iteration of this research, participation of these teachers in the online Knowledge Forum platform will be considered. Examining the posts of these teachers will also allow for inquiry in the innovation challenges faced by the individual teachers as they reflect upon their teaching practices.

In terms of addressing the research gap in PLC, this study suggested a way of engendering teacher learning while making a real difference in student learning. Unlike the lesson study sessions reported in Lee's study (2015), the teachers in this study were deeply engaged in their weekly discussions; they were not rushing to complete the task of "meeting". Similar to lesson study sessions reported by Lee, the teachers in this study made significant progress in lesson design that can impact student learning. The collaborative knowledge building work among the teachers helps them focus on the needs of the learner and work relentlessly to improve pedagogy so those needs are effectively met (Harris & Jones, 2010). Indeed, the knowledge building PLC provides intellectual, social and materials resources for teacher learning and innovations in practice by anchoring collaborative reflective practice on students' real ideas and authentic problems. The students' posts in the Knowledge Forum provide the material resources as referents for teacher's discussion. The common goal in designing lessons to help students advance their ideas provides a natural collaborative task. This arrangement brings meaning to teacher's contribution during their meetings, thus extends taken-for-granted weekly meetings as a means to support their professional thinking instead of a simple requirement task of meeting to work out the logistics of following week's lessons.

Conclusion and implication

The knowledge building practices of teachers in this report highlights the continuity of the work practice and individual teacher's needs to realise their professional development goals. We propose that consideration of teacher learning in a professional learning community be conceptualised in terms of participatory practices as facilitated by knowledge building principles. A helpful approach is to encourage and facilitate teacher learning through work and provision of an environment where such learning and associated teacher professionalism can flourish. Foregrounding knowledge building as a self-organizing concept that engages teachers in continuous improvement of lesson plans for implementation, we concur with Barber and Mourshed's (2009) suggestion that this mode of localized teacher professional development is effective for being authentic, situative, contextualized

and practice-oriented. Moving forward, we suggest enhancing the awareness of teacher participants in terms of the reciprocal layer of knowledge building with more intentional leverage on creating knowledge artefacts to mediate their collective advancement in pedagogical knowledge.

References

- Allaire, S., Laferrière, T., & Gervais, F. (2011). Enhancing pre-service teachers' knowledge building discourse with a hybrid learning environment. *Research on Education and Media*, 3(1), 67-83.
- Barber, M., & Mourshed, M. (2009). *Shaping the future: How good education systems can become great in the decade ahead*. Report on the International Educational Roundtable, 7 July 2009, Singapore. McKinsey & Company.
- Bereiter, M. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum.
- Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. *Knowing, Learning, and instruction: Essays in honor of Robert Glaser*, 361-392.
- DuFour, R., DuFour, R., Eaker, R., & Many, T. (2006). *Learning by doing: A handbook for professional learning communities that work*. Bloomington, IN: Solution Tree.
- Edley, N. (2001). Analysing masculinity: Interpretive repertoires, ideological dilemmas and subject positions. In M. Wetherell, S. Taylor & S. Yates (Eds.), *Discourse as data. A guide for analysis* (pp.189-228). The Open University: Sage.
- Greeno, J. G. (2003). On claims that answer the wrong questions. *Educational Researcher*, 26, 5-17.
- Harris, A., & Jones, M. (2010). Professional learning communities and system improvement. *Improving Schools*, 13(2), 172-181.
- Horn, I. S., & Kane, B. D. (2015). Opportunities for Professional Learning in Mathematics Teacher Workgroup Conversations: Relationships to Instructional Expertise. *Journal of the Learning Sciences*, (just-accepted).
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The journal of the learning sciences*, 4(1), 39-103.
- Lee, L. H. J. (2015). *Teacher learning in a professional learning team: Affordances, disturbances, contradictions, and action possibilities*. Doctoral thesis. Nanyang Technological University.
- Lieberman, A., & Mace, D. H. P. (2008). Teacher learning: The key to educational reform. *Journal of Teacher Education*, 59(3), 226-234.
- Little, J. W. (2003). Locating learning in teachers' communities of practice: Opening up problems of analysis in records of everyday work. *Teaching and Teacher Education*, 18(8), 917-946.
- McLaughlin, M. W., & Talbert, J. E. (2001). *Professional communities and the work of high school teaching*. University of Chicago Press.
- Orland-Barak, L., & Tillema, H. (2006). The 'dark side of the moon': a critical look at teacher knowledge construction in collaborative settings. *Teachers and Teaching: theory and practice*, 12(1), 1-12.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction*, 1(2), 117-175.
- Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational researcher*, 4-15.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. *Liberal education in a knowledge society*, 97, 67-98.
- Scardamalia, M. & Bereiter, C. (2010). A brief history of knowledge building. *Canadian Journal of Learning and Technology*, 36(1). Retrieved from <http://www.cjlt.ca/index.php/cjlt/issue/view/70>
- Scardamalia, M., & Bereiter, C. (2015). Knowledge building: Theory, pedagogy, and technology. In Sawyer, R. K. (Ed.). *The Cambridge Handbook of the Learning Sciences* (2nd Ed., pp. 397-417-118). NY: Cambridge University Press.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: SAGE Publications.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of Educational Change*, 7, 221-258.

Acknowledgments

Appreciation goes to the teachers for their conscientious effort in designing and enacting knowledge building lessons. This work is funded by MOE Academies Fund NRF2012-EDU001-EL008 and the National Institute of Education, Singapore, project RS 7/13 TSC.

No Lives Left: How Common Game Features Can Undermine Persistence, Challenge-Seeking, and Learning to Program

Laura J. Malkiewich, Alison Lee, Stefan Slater, Chenmu Xing, and Catherine C. Chase
ljm2172@tc.columbia.edu, alison.lee@tc.columbia.edu, slater.research@gmail.com,
xing@exchange.tc.columbia.edu, cc3663@tc.columbia.edu
Teachers College, Columbia University

Abstract: Persistence is necessary for learning, yet students have difficulty persisting at academic tasks, especially when faced with challenge. One context where students seem to persist is gameplay. This paper investigates whether an educational computer programming game can enhance student persistence and learning. Students were assigned to play either the Full version of a game, or a Minimal version of the game that ablated several common game features. Contrary to our hypotheses, we found that students who played the Full Game persisted less at challenging tasks, wrote less challenging code during gameplay, and were less willing to learn more about coding in the future. Ultimately, players of the Full Game also learned less than students who played the Minimal Game. Results suggest that certain design features of educational games can negatively impact a player's approach to challenging learning tasks, which in turn can negatively impact learning.

Keywords: games, computer programing, persistence, self-efficacy, challenge-seeking

Introduction

Persistence is critical for learning, yet students frequently fail to persist enough for real learning to occur, often giving up when academic tasks become too challenging (Chase, 2013; Diener & Dweck, 1978). Though students struggle to persist in academic contexts, they seem more willing to persist while playing video games. Children and adolescents spend an average of 90 mins/day playing video games (Rideout et al., 2010), and over half of adolescents find online games "addictive" (Yee, 2006). These statistics suggest that children are persistent during gameplay even though failure occurs often during gameplay, and people find failure uncomfortable (Juul, 2013). Furthermore, a study by Ventura, Shute and Zhao (2013) showed that adult subjects who reported playing video games frequently spent more time attempting to solve an impossible task. These results suggest that educational video games might be a context where students could develop persistent behaviors, which in turn could benefit learning. Still, while educational games might promote student persistence, there is little direct evidence that games enhance persistence during challenging tasks. Furthermore, little empirical research explains how games promote persistence more generally, making it difficult to ascertain what game-like elements, if any, might help students persist during or outside of gameplay.

One way in which games might promote player persistence is by increasing motivation. Many argue that games foster high intrinsic motivation or enjoyment (Habgood & Ainsworth, 2011; Lepper & Malone, 1987; Dickey, 2007; Fernandez, 2008), so children want to keep playing even when the game gets challenging. For instance, work on self-determination theory (Ryan & Deci, 2000) suggests that games promote intrinsic motivation because they satisfy needs for competency, relatedness, and autonomy through common game elements like the use of fantasy, structured rules and goals, sensory stimuli, challenge, mystery, and control (Ryan, Rigby, & Przybylski, 2006; Garris, Ahlers, & Driskell 2002; Birk & Mandryk, 2013). Other work suggests that games can launch players into a state of flow where they feel completely immersed in gameplay (Csikszentmihalyi, 1990), perhaps by providing clear goals, immediate feedback, and a sense of autonomy (Cowley, Charles, Black, & Hickey, 2008; Chen, 2007; Sweetser, Wyeth, 2005). However, these motivational theories focus mostly on how game elements make gameplay enjoyable. Persistence in the face of challenging tasks may be more related to students' perceptions of their own competence.

We take a novel viewpoint in this paper by proposing that games promote persistence by increasing self-efficacy (Bandura, 1993; Ketelhut, 2007), a feeling of competence for a task that is enhanced by experiences of success and diminished by experiences of failure. While some findings are mixed (Schunk, 1989, 1991), many studies have demonstrated the link between high self-efficacy and persistence in problem solving (Relich, Debus, & Walker, 1986), writing (McCarthy, Meier, & Rinderer, 1985; Pajares & Johnson, 1994; Schunk, 2003), coping behavior (Bandura, 1977), and even the pursuit of scientific majors (Lent, Brown, & Larkin, 1984). However, the literature has yet to review how self-efficacy affects persistence during video gameplay. We argue that games

enhance self-efficacy by altering players' perceptions of success and failure. For instance, games may enhance self-efficacy by providing challenges that gradually increase in difficulty over time, in line with the player's ability and progress, providing learners with a series of more and more impressive successes (Przybylski, Rigby, and Ryan, 2010). Moreover, players may view failure in games as acceptable and normal since it is "just a game," whereas failure in school contexts where the stakes are higher can be discouraging (Litts & Ramirez, 2014). Finally, games provide functional feedback in the form of metrics like badges and points. Past work shows that performance feedback after success and failure can improve student self-efficacy (Bandura, 1993; Vallerand & Reid, 1984). Therefore, there seem to be many common game features that could boost players' self-efficacy for gameplay, which in turn could promote student persistence at game challenges. However, limited empirical research links self-efficacy to persistence in the context of an educational game.

By promoting persistence, games could lead students to learn deeply. While still a hotly contested topic (Hilton & Honey, 2011), several recent meta-analyses have concluded that educational games enhance learning over standard instructional activities (Wouters et al., 2013; Clark, Tanner-Smith, & May, 2013). Perhaps these games are more effective than traditional instruction because they promote students' self-efficacy at in-game learning tasks, which enhances persistence at challenging learning tasks and ultimately improves learning.

We believe that educational games are particularly well suited for teaching academic content that is highly iterative, where students face a lot of failure during the learning process and persistence is necessary for success. One such domain is computer programming, as coding requires students to effectively debug code that fails to function properly (Papert, 1980; Pea & Kurland, 1984b). The plethora of recent commercially developed games that aim to teach computer programming such as CodeSpark, Code Combat, Space Chem, and many others supports the notion that games might be a good avenue for teaching computer programming.

In this paper, we ask whether an educational game can enhance persistence at challenging, iterative learning tasks in comparison to a learning environment with few game features. Students played one of two versions of a game designed to teach computer programming: a Full Game and a Minimal Game (Figure 1). The Full Game was a commercially available educational puzzle game about computer programming with typical game features; the Minimal Game contained identical scaffolding and content but was stripped of many game features that we hypothesized would affect student self-efficacy and therefore student persistence in turn.

We hypothesized that the Full Game would promote more persistence, particularly in the face of challenge, which would lead to greater learning. We also predicted that Full Game players would have higher self-efficacy for coding and that self-efficacy would be associated with more persistence at a challenging task. Finally, we hypothesized that players of the Full Game would have greater intrinsic motivation for gameplay. Our results supported the link between persistence and learning but did not support the link between self-efficacy and those constructs. Results also suggested that some features of the Full Game were detrimental to both persistence and learning, perhaps by discouraging challenge-seeking behaviors.

Methods

Participants

Thirty-seven fifth grade students (54.1% female) were recruited from an urban charter school after-school program. Prior to the study, 89% of participants had no programming experience; the rest had less than a month's worth of programming experience.

Design

Students were randomly assigned to the Full Game condition ($n=18$) or the Minimal Game condition ($n=19$). Both games used block-based code, similar to Scratch (Maloney et al., 2010). In each game level, students solved problems by writing code to navigate an agent over obstacles and reach a goal. Both games contained identical problems and hints. However, the Full Game contained additional standard game features such as a narrative story, high quality graphics, and sound. Additionally, the Full Game highlighted successful or failed attempts at a game level by displaying messages like "OOPS" or "Missed It" after a failure or having the character smile or jump after a success. The Full Game also provided performance metrics in the form of points and stars at the end of each level. Students could earn one, two, or three stars depending on how many hints they asked for, and how many coding blocks they used (where fewer is better) to complete a problem. Finally, the Full Game provided purely fun "bonus" levels where students could collect jellybeans or shoot down asteroids instead of solving coding problems.



Figure 1. Screenshot of a typical level in the Full Game (left) and the same level in the Minimal Game (right).

Procedure

Each student participated in the study for five, 40-minute sessions: a pretest session, two gameplay sessions, a challenge session, and a posttest session. In the pretest session, students took a survey of their self-efficacy for coding followed by a brief paper pretest of their coding knowledge. In the two gameplay sessions, students played the game individually on iPads. In the next session, students attempted an impossible coding challenge (the challenge task). Finally, in the posttest session, all students were given a survey assessing their self-efficacy for coding and their intrinsic motivation for the game, followed by a paper posttest and a then a measure of persistence at future coding.

Measures

Prior programming knowledge and ability was measured by a paper-based pretest consisting of four test items on constructing code, interpreting code, and planning, a relevant skill to programming (Pea & Kurland, 1984a). Learning outcomes were measured by a 16-question paper-based posttest that assessed programming skills such as conceptualizing, writing, interpreting, and debugging code.

Persistence in the face of failure was measured by the time students voluntarily spent on the challenge task without giving up. The challenge task was an impossible, novel level embedded in the game the students were playing, and therefore had all the challenge level had all the same features of a typical level from each respective game (Full or Minimal). The challenge level was locked prior to the challenge task day, so students could not access the challenge level during the first two days of gameplay. Students were told that to complete the challenge level, they had to solve the challenge task using no more than a certain number of blocks. Unbeknownst to the students, this number of blocks was less than the minimum required to successfully solve the problem. As a result, students experienced failure at every attempt; even when students thought that they solved the problem, they were told by the researchers that they used too many blocks in their solution and they could try again. At any time, students were given the option to continue trying to solve the challenge level (for up to 40 minutes), or quit to explore computer-based science simulations.

Students' self-efficacy for programming was measured by a 7-point Likert-scale survey before and after gameplay, in which students rated their confidence in their ability to write, interpret, debug, and choose correct code. Students were asked questions like, "If something is wrong with your code, how confident are you that you can fix it?". Student intrinsic motivation for the game was also measured using a 7-point Likert-scale survey that asked students how interesting, exciting, enjoyable, and fun they found the game in general, across the two days of gameplay and the challenge level. To measure persistence in future coding, students were given the option of taking a handout that referred them to various websites where they could learn more about coding.

Gameplay data was collected using screen recordings and embedded log capture. Due to technical difficulties, some video and log data was lost, so all measures of student coding behaviors and time spent coding only involve a portion of the total study sample (n=9 for each group). We did an exploratory analysis of the gameplay data to find in-game behaviors that indicated students' approach and response to challenge during gameplay. This analysis helped us identify one measure of in-game persistence, and one measure of challenge-seeking during gameplay. In-game persistence was measured by time spent playing coding levels vs. bonus levels. Challenge-seeking during gameplay was measured by how often students chose to engage in challenging learning behaviors by contrasting the percentage of levels that students completed with either basic,

intermediate, or complex code. Students had access to two types of coding blocks: “action” blocks make the character walk or jump, while “parameter” blocks include repeating loops and if/then conditional commands that need to be filled with action blocks in order to work. Codes that only used action blocks were labeled “basic”, codes that implemented rote usage of parameter blocks were labeled “intermediate”, and codes that adapted parameter blocks to suit the given problem were labeled “complex”. We coded each game level’s difficulty by indicating how parameter or action blocks could be used in that level. Easy levels required only using action blocks, medium levels allowed students to use action blocks and one parameter block, and hard levels allowed students to use action blocks and multiple parameter blocks.

Findings

Learning outcomes

Students in both conditions started with equivalent prior knowledge, and learned from pretest to posttest, but the Full condition acquired less programming knowledge from gameplay. T-tests found no significant difference between conditions on the coding skills pretest, $t(35) = 0.12, p = .91$, or mean GPA scores, $t(35) = 1.63, p = .11$. An ANCOVA with pretest as covariate revealed that students in the Full condition performed worse on the posttest than the Minimal condition, $F(1,34) = 6.50, p = .02, \eta_p^2 = 0.16$. When controlling for prior knowledge, Full Game students scored 22% lower than Minimal Game students at posttest (Table 1). These results imply that, contrary to our hypotheses, students who played the Full Game learned less than students who played the Minimal Game.

Persistence measures

Also contrary to our hypothesis, measures of persistence indicated that students who played the Full Game were less persistent at coding tasks.

On the challenge task, where students were given an impossible coding problem that they could quit at any time, students who played the Full Game persisted about 8 mins or 30% less than those who played the Minimal Game (Table 1), $t(35) = 2.61, p = .01; d = 0.88$. So students in the Full condition gave up sooner on a hard challenge set within the context of the game than students who were in the Minimal condition.

Table 1: Student learning and persistence scores

**Significant difference in scores, $p < .05$. Challenge times are in minutes.*

Game	<i>n</i>	Pretest		Posttest		Challenge Time	
		<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Full	18	1.97	0.18	7.58*	0.61	19.08*	2.31
Minimal	19	1.99	0.15	9.76	0.60	27.01	1.98

While persistence on the challenge task shows students’ willingness to persist within their respective game environments, we also wanted to assess whether students would persist at coding in general, beyond the confines of the study. We found that students who played the Full Game were less likely to want to pursue future coding activities. A chi-square test demonstrated that fewer students in the Full Game condition voluntarily took a handout explaining where they could learn more about coding, $X^2(1, N = 37) = 4.56, p = .03$.

Students in the Full Game also failed to persist at coding during gameplay. Students who played the Full Game spent more time playing bonus levels and less time on coding levels during the second gameplay session. A repeated measures ANOVA of time on coding levels showed that there was a significant interaction between condition and gameplay session, $F(1,16) = 5.18, p = .04$. On the first day of gameplay, students across both conditions spent the same amount of time playing coding levels ($M_{Full} = 35.47, SE_{Full} = 0.90; M_{Minimal} = 34.95, SE_{Minimal} = 0.90$) even though students in the Full version of the game had access to bonus levels with no learning content. However, during the second play session, students who played the Full Game spent less time on coding levels than students who played the Minimal Game ($M_{Full} = 26.95, SE_{Full} = 2.11; M_{Minimal} = 33.90, SE_{Minimal} = 2.11$), $t(16) = 2.33, p = .03, d = 1.17$. This implies that by the second day of gameplay, when the novelty of the game had worn off, students who played the Full Game persisted less at coding and instead spent more time on non-coding, bonus levels.

Students in the Full condition were also less likely to engage in challenging coding behaviors, particularly in challenging situations. For instance, students in the Full condition tended to shy away from writing

complex code on hard problems. A repeated measures ANOVA showed that condition had a significant effect on the level of complexity of the code that students used in their first attempt to solve hard problems, $F(2,32) = 4.29$, $p = .02$, $\eta_p^2 = 0.21$. Specifically, students who played the Full Game were more likely to use intermediate code in their first attempt at a hard problem, $t(16) = 2.23$, $p = .04$, $d = 1.12$, and less likely to use complex code in their first attempt at a hard problem, $t(16) = 2.45$, $p = .03$, $d = 1.23$. These differences were significant after applying Scheffé's method for alpha correction for family-wise Type 1 error rate control. The same pattern appeared when we looked at the type of code students used to complete a problem. A repeated measures ANOVA showed a significant interaction between condition, coding complexity (basic, intermediate, complex), and problem difficulty (medium vs. hard) on percentage of problems completed $F(2,32) = 10.50$, $p < .001$, $\eta_p^2 = 0.40$. Relative to the Minimal condition, the Full condition solved a larger percentage of hard problems with intermediate code, $t(16) = 2.66$, $p = .02$, $d = 1.33$, and a smaller percentage of hard problems with complex code, $t(16) = -3.22$, $p = .01$, $d = 1.61$ (Table 2). Again, these differences were significant after applying Scheffé's method for alpha correction for family-wise Type 1 error rate control. These results indicate that students who played the Full Game were less willing to take risks in challenging situations; they chose to attempt less complicated code when first trying to solve hard problems and when completing them.

Table 2: Percent of levels completed with various sophistications of code

	Medium Levels		Hard Levels	
	Minimal Game	Full Game	Minimal Game	Full Game
Complex	15%	22%	34%*	9%
Intermediate	18%	16%	31%*	46%
Basic	67%	62%	36%	45%

*Significant difference in scores, $p < .05$. Note: all easy levels are solved with basic code.

To test the relationship between persistence, risk taking, and learning, we explored the association between persistence measures and learning outcomes. A regression of persistence times on post-test scores, controlling for condition and pretest score, revealed that persistence on the challenge task trended towards predicting learning, $\beta = 0.46$, $t(32) = 1.952$, $p = .06$. The model predicted a significant amount of the variance in posttest scores $R^2_{adj} = 0.26$, $F(4,32) = 4.15$, $p = .01$. There was no interaction between challenge time and condition and there was no main effect of condition. So regardless of condition, the more students persisted on the challenge task, the better they did on the posttest. Challenge time and percent of hard levels completed with complex code were highly correlated $r(16) = 0.54$, $p = .02$, but willingness to take risks and use complex code on hard problems was more predictive of learning. A linear regression controlling for condition and pretest score found that percent of hard levels solved with complex code predicted a significant amount of the variance in posttest scores $R^2_{adj} = 0.53$, $F(4,13) = 5.84$, $p < .01$. Again, there was no interaction with condition; regardless of condition, students who used challenging code to solve hard problems did better on the posttest. Specifically, every 10% increase in the proportion of hard levels students completed with complex code was associated with a 0.87 point increase in posttest score, $t(13) = 3.27$, $p < .01$. In other words, the more students attempted challenging coding behaviors on hard levels, the more they learned.

Challenge-seeking behaviors also predicted student persistence on the challenge task. A linear regression showed that the percent of hard problems that students completed with complex code predicted the amount of time students spent on the challenge task before giving up, $\beta = 0.54$, $t(16) = 2.59$, $p = .02$. Percent of hard problems completed with complex code also predicted a significant amount of the variance in challenge time $R^2_{adj} = 0.25$, $F(1,16) = 6.731$, $p = .02$. So students who risked getting hard problems wrong by using more complex code, also spent more time persisting at the challenge task. One possible explanation for this finding is that students who engaged in more complex coding during the game persisted longer on the challenge task, because they had developed more strategies (both intermediate and complex codes) to attempt before giving up on the challenge task. In other words, we suspect that the effect of persistence on posttest scores was mediated by challenge-seeking behavior, however we did not have the statistical power to detect these effects, since we only had in-game behavior data for half of our sample ($n=9$ per condition). Another possibility is that the features of the Full Game are having two independent effects by leading players to engage in both less challenge-seeking and less persistence in the face of challenge, which both negatively affect learning.

Self-efficacy and intrinsic motivation

Conditions did not differ in their coding self-efficacy before game play, $t(35) = 0.939$, $p = .35$. A repeated measures ANOVA on pre and post self-efficacy ratings found a main effect of time $F(1,35) = 84.01$, $p < .01$, $\eta_p^2 = 0.71$ and a significant interaction $F(1,35) = 4.30$, $p = .05$, $\eta_p^2 = 0.11$. This suggests that while students in both conditions gained self-efficacy from pre to post, the Full condition made significantly smaller gains (Pre-Post Gains: $M_{Full} = 1.90$, $SE_{Full} = 0.40$; $M_{Minimal} = 3.01$, $SE_{Minimal} = 0.36$). Contrary to our hypothesis, self-efficacy was not correlated with persistence measures or posttest scores, p 's $> .09$.

There was no significant difference between conditions in their intrinsic motivation for playing the game. Both conditions found the game equally enjoyable, giving fairly high average ratings on the 7-point intrinsic motivation survey $t(35) = 1.06$, $p = .30$, ($M_{Full} = 5.25$, $SE_{Full} = 0.35$; $M_{Minimal} = 5.75$, $SE_{Minimal} = 0.32$). Intrinsic motivation was not related to persistence measures or posttest scores, p 's $> .15$.

Conclusions and implications

Contrary to our hypotheses, we found that students who played the Full version of a coding game, with many standard game features, showed less persistence at challenging coding tasks compared to students who played a Minimal version of the game, with many game features removed. When given an impossible coding problem within the context of the game environment, students in the Full condition gave up more quickly. Fewer students in the Full condition wanted to pursue future coding after the study ended. During the second gameplay session, after the novelty of the game had worn off, students who played the Full Game were more likely to spend time playing content-free "bonus" levels and less time playing coding levels. In addition to demonstrating less persistence at challenging tasks, exploratory analyses revealed that students who played the Full Game chose to engage in less challenging coding behaviors by using less complex code on hard levels, where the risk of failure was great. This suggests that players of the Full Game engaged in fewer risk-taking behaviors. It is interesting to note that while we had originally focused on persistence at challenging tasks as a key learning behavior, we found that the Full Game discouraged risky learning behaviors during challenging tasks. Overall, this suggests that players of the Full Game had a less productive approach to challenge – they chose to engage in risky coding behaviors less frequently and persisted less.

The Full Game group's less productive approach to challenge may explain why they learned less. Use of complex code on hard problems significantly predicted learning and persistence in the challenge task marginally predicted learning. This was true regardless of condition. So students who had a productive approach to challenge learned more from the game. While the relationship between productive approaches to challenge and learning outcomes is correlational (not causal), it provides a likely explanation for why the Full condition learned less from the game. The Full Game could have hindered learning by hindering a players' use of complex code on hard problems and a player's willingness to persist at the challenge task, both of which seem to be independently pathways for learning. Given the correlational relationship between challenging coding behaviors and persistence on the challenge task, another possibility is that learning more complex code during gameplay enabled learners to persist longer in the challenge task by giving students a larger repertoire of coding behaviors to draw from. So students who were risk-taking and tried using more complex code to solve hard problems during the game might have learned more coding techniques to try while attempting to complete the challenge task. In this case, use of complex code during the game could have mediated the relationship between persistence during the challenge task and learning. We unfortunately do not have enough power from this study to test these two theories.

What caused the Full condition to show a less productive approach to challenge by persisting less at challenging tasks and not choosing to engage in writing complex code on challenging problems? The main mechanism we hypothesized, self-efficacy, does not seem to explain this phenomenon. Although the Full Game students gained less self-efficacy for coding by playing the game, self-efficacy was not predictive of persistent behaviors or challenge-seeking coding behaviors. These findings suggest that other motivational mechanisms were affected by the Full Game, which in turn discouraged challenge-seeking and persistence. One possibility is that the Full Game was not engaging for students. However, both conditions rated their intrinsic motivation for the game equally highly. Another possible explanation is that the Minimal Game may have felt less "game-like" and more like a serious learning environment, causing the Minimal condition students to persist more at learning-relevant behaviors. However, when asked to rate how much the learning environment "felt like a game," both conditions gave equally high ratings.

A more plausible explanation for the Full condition's less productive approach to challenge is that the Full Game made failure more salient. When players failed in the Full Game, a huge "OOPS" pulsed on screen. At the end of each level, star counts and points indicated whether the level was solved with the optimal code, further marking failure when the students' code was non-optimal. The Minimal condition did not contain these standard

game features. While meant to provide clear feedback and encourage revision, these features may have inadvertently intensified perceived student failure. We are currently running a second study where we are manipulating failure feedback in the game to test how salience of failure impacts persistence and learning. Even though in the first study the rate of failure was equivalent across conditions, if the Full condition felt more affected by failure, then it stands to reason that they would both persist less at challenging tasks and choose to engage less in challenging coding behaviors, where the risk of failure is great. In fact, the Full condition did show lower gains in self-efficacy in comparison to the Minimal condition, which would be a fitting outcome since self-efficacy is primarily boosted by feelings of success. Though we did not find a significant relationship between self-efficacy and persistence, it is possible that more proximal measures of self-efficacy (e.g. given immediately after failure in the game) would better predict persistence.

Finally, the Full condition may have shown a less productive approach to challenge because the Full Game's feedback and reward system dis-incentivized students from tinkering. Feedback and reward for students was operationalized by their star count, shown at the end of each level and on the main level screen so that students could see their performance on completed levels. Even if students solved a level with the optimal code, their star count would decrease if they took too many attempts while trying to solve a level. This reward mechanism could have discouraged students from trying more complex code, as using more complicated code makes it harder to get a problem right on the first try. This could explain why students who played the Full Game used less complex code when beginning and solving hard problems, while students who played a game without this reward system were more willing to attempt using complex code on hard levels. This may also explain why Full Game players were less persistent at the challenge task. If students playing the Full Game were discouraged from tinkering, they may have developed fewer coding strategies to try on the challenge task, causing them to run out of ideas and quit sooner.

We do not mean to imply that all games are bad for persistence and learning. Obviously, this study is limited in that it only investigated a single game that encapsulated multiple standard game features. Rather, we interpret these results to mean that *some* commonly-used game design elements *can be* implemented in such a way that they negatively affect persistence and challenge-seeking behaviors, which can in turn hinder learning. Future work can focus on isolating which game features discourage productive learning behaviors, and under what conditions.

References

- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.
- Birk, M., & Mandryk, R. L. (2013, April). Control your game-self: effects of controller type on enjoyment, motivation, and personality in game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 685-694). ACM.
- Chase, C.C. (2013). Motivating expertise: Equipping novices with the motivational tools to move beyond failure. In J.J. Staszewski (Ed.), *Expertise and skill acquisition: The impact of William G. Chase*. New York: Psychology Press.
- Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, 50(4), 31-34.
- Clark, D. B., Tanner-Smith, E. E., & May, S. K. (2013). *Digital Games for Learning: A Systematic Review and Meta-Analysis (Executive Summary)*. Menlo Park, CA.
- Cowley, B., Charles, D., Black, M., & Hickey, R. (2008). Toward an understanding of flow in video games. *Computers in Entertainment (CIE)*, 6(2), 1-28.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal performance*. New York: Cambridge University Press.
- Dickey, M. D. (2007). Game design and learning: A conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development*, 55(3), 253-273.
- Diener, C. L., & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 36, 451-462.
- Fernandez, A. (2008). Fun Experience with Digital Games: A Model Proposition. In O. Leino, H. Wirman & A. Fernandez (Eds.), *Extending Experiences: Structure, Analysis and Design of Computer Game Player Experience* (pp. 181-190). Rovaniemi, Finland: Lapland University Press.

- Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & gaming*, 33(4), 441-467.
- Habgood, M. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *The Journal of the Learning Sciences*, 20(2), 169-206.
- Hilton, M., & Honey, M. A. (Eds.). (2011). *Learning science through computer games and simulations*. Washington: National Academies Press.
- Hsu, C. L., & Lu, H. P. (2004). Why do people play on-line games? An extended TAM with social influences and flow experience. *Information & Management*, 41(7), 853-868.
- Juul, J. (2013). *The art of failure: An essay on the pain of playing video games*. Cambridge, MA: MIT Press.
- Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology*, 16(1), 99-111.
- Lepper, M. R., & Malone, Th. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction: Vol. 3. Conative and affective process analyses* (pp. 255-286). Hillsdale, NJ: Lawrence Erlbaum.
- Litts, B., Ramirez, D. (2014) Making people fail: Failing to learn through games and making. Proceedings of the Games Learning Society (GLS) Conference. Madison, Wisconsin.
- Maloney, J., Resnick, M., Rusk, N., Silverman, B., & Eastmond, E. (2010). The scratch programming language and environment. *ACM Transactions on Computing Education (TOCE)*, 10(4), 16:1-16:15.
- Pajares, F., & Johnson, M. J. (1994). Confidence and competence in writing: The role of self-efficacy, outcome expectancy, and apprehension. *Research in the Teaching of English*, 28, 316-334.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books, Inc.
- Pea, R. D., & Kurland, D. M. (1984a). *Logo programming and the development of planning skills*. (Technical Report No. 16). New York: Columbia University.
- Pea, R. D., & Kurland, D. M. (1984b). On the cognitive effects of learning computer programming. *New Ideas in Psychology*, 2(2), 137-168.
- Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology*, 14(2), 154-166.
- Relich, J. D., Debus, R. L., & Walker, R. (1986). The mediating role of attribution and self-efficacy variables for treatment effects on achievement outcomes. *Contemporary Educational Psychology*, 11(3), 195-216.
- Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). Generation M²: Media in the Lives of 8-to 18-Year-Olds. *Henry J. Kaiser Family Foundation*.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68-78.
- Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, 30(4), 344-360.
- Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review*, 1(3), 173-208.
- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26(3-4), 207-231.
- Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading & Writing Quarterly*, 19(2), 159-172.
- Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, 3(3).
- Vallerand, R. J., & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. *Journal of Sport Psychology*, 6(1), 94-102.
- Ventura, M., Shute, V., & Zhao, W. (2013). The relationship between video game use and a performance-based measure of persistence. *Computers & Education*, 60(1), 52-58.
- Voiskounsky, A. E., Mitina, O. V., & Avetisova, A. A. (2004). Playing Online Games: Flow Experience. *PsychNology Journal*, 2(3), 259-281.
- Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249-265.
- Yee, N. (2006). Motivations for play in online games. *CyberPsychology & Behavior*, 9(6), 772-775.

Acknowledgments

We thank participating teachers and students. We would also like to thank Joan Cejudo for his programming work on this project.

Writing Analytics for Epistemic Features of Student Writing

Simon Knight, University of Technology Sydney, simon.knight@uts.edu.au
Laura Allen, Arizona State University, laurakallen@asu.edu
Karen Littleton, Open University, k.s.littleton@open.ac.uk
Bart Rienties, Open University, bart.rienties@open.ac.uk
Dirk Tempelaar, Maastricht University, d.tempelaar@maastrichtuniversity.nl

Abstract: Literacy, encompassing the ability to produce written outputs from the reading of multiple sources, is a key learning goal. Selecting information, and evaluating and integrating claims from potentially competing documents is a complex literacy task. Prior research exploring differing behaviours and their association to constructs such as epistemic cognition has used ‘multiple document processing’ (MDP) tasks. Using this model, 270 paired participants, wrote a review of a document. Reports were assessed using a rubric associated with features of complex literacy behaviours. This paper focuses on the conceptual and empirical associations between those rubric-marks and textual features of the reports on a set of natural language processing (NLP) indicators. Findings indicate the potential of NLP indicators for providing feedback regarding the writing of such outputs, demonstrating clear relationships both across rubric facets and between rubric facets and specific NLP indicators.

Keywords: epistemic cognition, literacy education, writing analytics, learning analytics

Introduction

Literacy, including the abilities to comprehend rich multimedia, and effectively communicate through written texts, is key to learning, and full participation in society across age ranges (OECD, 2013; OECD & Statistics Canada, 2010). Indeed, the 2009-2015 PISA definition of reading indicates that: “Reading literacy is understanding, using, reflecting on and engaging with written texts, in order to achieve one’s goals, to develop one’s knowledge and potential, and to participate in society” (OECD, 2013, p. 9). Thus Rouet (2006) suggests that literacy in the context of rich-multimedia environments (such as the internet) involves the skills of: *integration* of prior knowledge and across documents (including competing claims); *sourcing* of features that identify the provenance, genre, etc. of the information; and *corroboration* to check information across multiple sources.

Multiple document comprehension

One class of research into this area of literacy has explored multiple document processing, the ability to read, comprehend and integrate information from across sources, (see, for examples, Bråten, 2008; Bråten, Britt, Strømsø, & Rouet, 2011; Ferguson, 2014; Foltz, Britt, & Perfetti, 1996; S. R. Goldman et al., 2011; Hastings, Hughes, Magliano, Goldman, & Lawless, 2012; Kobayashi, 2014; Rouet & Britt, 2011), with some research specifically viewing these behaviours through the lens of epistemic cognition – characterised as beliefs about the certainty, simplicity, source, and justification of knowledge (see, for examples, Bråten, 2008; Bråten et al., 2011; Ferguson, 2014). Such tasks provide a context in which to explore the ways in which different sources are treated and drawn on in subsequent tests of knowledge or argumentation; the language participants use in these subsequent contexts may relate to the particular documents implicitly or explicitly drawn on. In this vein, in recent work on literacy and epistemic cognition (Anmarkrud, Bråten, & Strømsø, 2014; Bråten, Braasch, Strømsø, & Ferguson, 2014) students were asked to produce written outputs, which were then scored for:

1. Presence of explicit or implicit sourcing (i.e. explicit reference to the source, or indirect reference such as “one article spoke of [specific detail]” but without direct use of source information);
2. References to trustworthiness of the source or information from that source (coding separately for negative and positive evaluations);
3. Finally, whether connections were made between content-source trustworthiness (for example, whether content was trusted more because of the properties of the document from which it was sourced)

That research found that, approximately half of sourcing references were explicit (with the other half implicit) and students did not make reference to the full list of sources (approximately 3 of 6 references). In other multiple document processing research, Goldman, Lawless, Pellegrino and Gomez (2012) identified three clusters of students from their written outputs: *satisficers*, who selected few sources; *selectors* who selected many sources but did not connect them; and *synthesisers* who selected sources and integrated them.

Developing language technologies

Given these prior research findings, the development of multiple document processing tasks provides opportunity to explore relationships among psychological constructs, outputs, and learning process data (Knight & Littleton, 2015). Emerging language technologies raise potential for such research into relationships between features in output texts and score-descriptors on rubric facets grounded in theorized constructs such as epistemic cognition. For example, analysis of the written outputs for: rhetorical moves that are indicative of claims making, evaluation, and connecting (or synthesis) (see, for example, de Waard, Buitelaar, & Eigner, 2009; Groza, Handschuh, & Bordea, 2010; Simsek, Buckingham Shum, Sandor, De Liddo, & Ferguson, 2013); text cohesion (McNamara, Louwerse, McCarthy, & Graesser, 2010); and topic coverage and integration (see, for example, Hastings et al., 2012).

In the last of these studies (Hastings et al., 2012), students were asked to use three texts with relatively little semantic overlap to answer the inquiry question “In 1830 there were 100 people living in Chicago. By 1930, there were three million. Why did so many people move to Chicago?” They compared three methods to match source material to student writing outputs: Pattern matching approaches (i.e. looking for common text-strings); latent semantic analysis (LSA) to compare semantic-content at a sentence level across student outputs and assigned texts; and machine learning (using support vector machines) assigning student sentences to topic-classes assigned by human-raters. They found that LSA performed best in identifying explicit use of the assigned texts, while pattern-matching approaches were superior for detecting intra and inter-textual inferences (which could be characterised as synthesis or integration of information).

The varied approaches to text analysis are of potential interest in developing approaches to a rich understanding of literacy and writing encompassing not only topic coverage, but also the ways information from multiple sources are integrated (synthesis), which sources information is drawn from (source diversity), and markers of evaluation and contrast (source quality and evaluation). The potential of language technologies, then, is to connect particular types or styles of language to epistemic characterisations; further work to connect computational outputs to human interpretable scores or feedback would then be required. Throughout this prior work the use of key-content, implicit and explicit citation, evaluation of those citations and (separately) their content, and the synthesis of information are foregrounded.

Current study

In the research reported in this paper, a conceptual alignment is drawn between the key literacy considerations of topic and source coverage, synthesis, and sourcing (including evaluation and justification), and a set of language-technology indicators. Examples of the kinds of language of interest to these features of written text are given, with preliminary results indicating relationships between outcomes on a rubric, and language technology indicators. The paper thus addresses the research aim to produce alignment between epistemic features in writing based on multiple document processing, and automated textual analysis of features that align with those key epistemic considerations.

Methods

Participants and ethics

270 students at Maastricht University (Netherlands) enrolled in a 1st year business and economics course took part in the study from which this data is drawn. Participants worked in pairs ($n = 135$), assigned by the researcher in the class environment. The research followed British Psychological Society ethics guidelines (British Psychological Society, 2014), with all students consenting to take part in the research and informed of their right to withdraw at any time. As part of this consent, students agreed to the Terms of Use of the Coagmento tool (described below), including that data obtained could be used for research purposes. Data is shared among the collaborators such that the Open University team holds all data, with other collaborators having access to subsets of data (anonymized for non-Maastricht researchers). The research was conducted on university PCs, and the browser cache was cleared between sessions to ensure no personal data (or active logins) was exposed either between participants, or to the researcher.

Materials and procedure

Participants completed the tasks in a computer lab using the Firefox browser (Mozilla, 2014) along with a customised add-on called Coagmento (Shah, 2010, 2014). Coagmento is a collaborative information-seeking (CIS) tool designed to support people in collaborating on shared information tasks. It includes a chat tool, and an

integrated ‘etherpad’ environment – a shared document editor, such that collaborators can write together in real time and share ‘snips’ (copied text) via the browser add-on tool.

In this research, participants were given a short warm-up task to familiarise themselves with the respective tools and their paired collaborator, following which they were requested to write collaboratively using the etherpad, to create a report of the “best support claims” around the risks of a substance (a herbicide called glyphosate). This topic selection follows the work described in the introduction; topics with conflicting perspectives and a variety of source-quality were sought to foreground participant’s commitments to varying source-content qualities, and abilities to integrate a variety of perspectives. The topic thus:

1. Provides a focussed topical research area which can be studied in isolation, within a 1 hour session;
2. Is not a topic that was high profile or/and a large scale controversy (such as climate change, or genetically modified crops, both of which receive large amounts of press coverage);
3. Has a variety of source-types and qualities regarding it, from varying perspectives.

For the task, a set of eleven documents was collated from materials that discussed a herbicide, the safety of which has been questioned in terms of human health implications and agricultural risks. A simplified document-model (building on Rouet’s work 2006, Chapter 3) is given in Figure 1, depicting the three key themes identified in this document set (the presence of glyphosate markers in human urine; the risks to human health of glyphosate; and the agricultural risks of glyphosate use), the document stance (broadly negative – orange; critique/broadly positive – green; largely neutral/scientific – blue) and the relations among them (+ - support; - - critique; note document 11 relates to 2 primary themes). Documents were presented via a webpage, with titles and short ‘snippets’ given (mirroring a search engine results page presentation of website titles and excerpts).

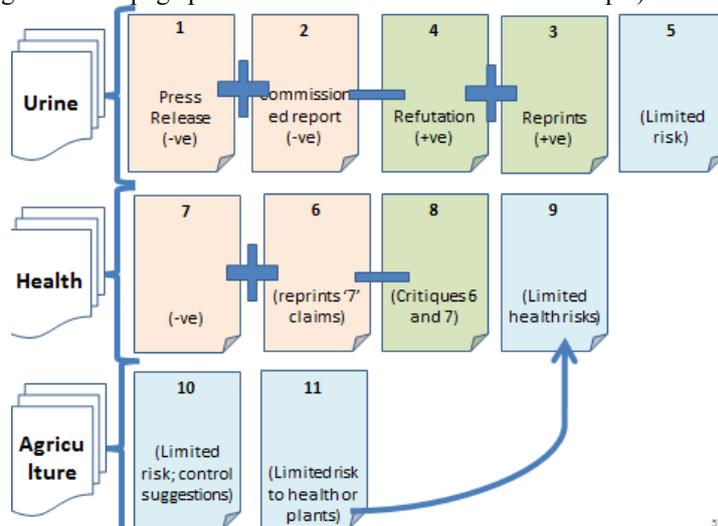


Figure 1. Simplified MDP Document Module.

For each document, the original source (HTML or PDF) was saved and formatted for presentation to students (to ensure it would load without scripts, etc., and would load in html without need to use a PDF viewer or other external reader). The documents were also cleaned, to remove extraneous detail and to reduce them to core claims around glyphosate. Only the abstracts of 5, 7 and 9 were given, while 10 was reduced to the abstract and first section of the introduction; 11 was reproduced in an abridged form, it was also the most comprehensive document in terms of coverage of potential risks.

Note in particular that the set of documents compiled is rather complex. For example, the author of 7 is criticised in 8. Ostensibly, 7 is more trustworthy because it is in a peer review journal and republished by Reuters (6), while 8 is a blog. However, the critique provided in 8 (and the evidence referred to) is strong and the source features of the blog (also based at MIT) are also strong. Furthermore, the author of 7 has been criticised for publishing in an area they are not an expert in, (including praising the discredited Andrew Wakefield on autism), and while the journal is peer reviewed, it is primarily a physics journal not a health-sciences one. We also see in documents ‘3’ and ‘4’ a reprint on a trade website (Farmers Weekly; 4) of an independent critique (3) – something students might identify as raising concerns of bias in ‘4’, although the content is identical. Thus the selection of documents provides a set of conflicting sources, of varying quality, with a range of sub-topics present. As such, the topic and selected documents provide good source material for probing students’ abilities to extract, integrate and evaluate information from across sources.

Participant outputs were assessed by the first author using a rubric. The rubric, based on the particular task design, and the MDP work described above, consisted of:

1. Topic coverage – The text covers a range of different topics and relates them to the question (the risks of the substance)
2. Range of sources coverage – The text uses a range of sources
3. Quality of sources – The text evaluates the quality of sources cited
4. Synthesis of information – The text synthesizes information from across sources

In this rubric, a score of ‘1’ indicates content coverage, ‘2’ the sourcing of that content, ‘3’ evaluation of the source features and content, and ‘4’ the ways intertextual ties that are identified and made in the text. Each rubric was assessed on a 1-3 scale. The same rubric was used for a parallel task based on different materials, for which a second rater assessed the student outputs, with acceptable reliability (>.8 Cohen’s Kappa on all indices except the synthesis score which had a .58 Kappa).

Analysis

Quantitative analysis

Quantitative analysis was conducted to assess the variance in rubric scores for the written outputs, intending to explore whether or not epistemic differences in writing are captured by the task design. Following a conceptual alignment (highlighted in the next section) between the rubric and quantifiable textual indicators in the written outputs, correlation analysis was conducted to explore these relationships empirically. The potential of this approach is to design models (for example, using regression analyses) to align the rubric facets with particular textual features that may be identified automatically using natural language technologies.

Qualitative analysis

Qualitative analysis was conducted to ground the conceptual alignment drawn between the rubric facets and anticipated textual features. The texts were analysed with regard to their epistemic and textual properties, with key features identified. Across the rubric facets variations in outcome were characterized by, for example:

1. **Synthesis:** Use of lists and extracts from individual articles, versus integrated text organized thematically and drawing from across multiple sources
2. **Topic Coverage:** A sparse use of topically salient keywords, or/and a focus on individual subtopics rather than drawing from the full range of themes and their keywords
3. **Source Diversity:** A focus on sourcing information from ‘one best’ article, versus the discussion and integration of claims from multiple sources
4. **Source quality:** Uncritical citation of claims, even where claims disagreed, versus identification and connection of disagreements and a critical balancing of claims based on source features

Natural language processing analysis

To assess the linguistic properties of the students’ writing, we utilized the *Tool for the Automatic Analysis of COhesion* (TAACO). TAACO is an automated text analysis tool that calculates 150 classic and recently developed indices related to both the local and global cohesion of a text, in contrast to other tools that either do not assess cohesion or focus solely on local cohesion (e.g. Coh-Metrix, McNamara, Graesser, McCarthy, & Cai, 2014). An additional strength of the tool is that it incorporates part-of-speech (POS) tags and synonym sets. The POS tags in TAACO are identified using the POS tagger developed as part of the Natural Language Tool Kit (Bird, Klein, & Loper, 2009), and the synonym sets are taken from the WordNet lexical database (Miller, 1995). In the sections below, we provide descriptions of the categories of TAACO indices most relevant to the current paper (i.e., *basic indices*, *sentence overlap*, *paragraph overlap*, and *connectives*). For more specific information on all of the indices provided by TAACO, see Crossley, Kyle, and McNamara (2015).

Basic indices. TAACO provides indices related to basic information about a text, such as the number of words (i.e., tokens), number of word types (i.e., unique words), and the type-token ratio (TTR). TTR calculates word repetition by dividing the total number of words in a text (tokens) by the number of individual words (types). Therefore, this index describes the amount of given information in a particular text. TAACO calculates a number of different TTR indices. These include simple TTR (the ratio of types to tokens), content word TTR (TTR using only content words such as nouns, verbs, adjectives, and adverbs) and function word TTR (TTR using only function words such as pronouns, preposition, and determiners). TTR indices have demonstrated positive relations with measures of cohesion in previous studies (Crossley & McNamara, 2014; McCarthy & Jarvis, 2010), but

generally demonstrate negative relations with measures of text coherence (Crossley & McNamara, 2010; McNamara, Crossley, & McCarthy, 2010).

TTR indices may account for both local and global characteristics of cohesion. Because they are measured at the level of the overall text, it is difficult to determine whether word repetition is occurring between sentences or larger portions of the text. In the current study, we calculated the *total number of words* and the *total number of word types*. Additionally, we used the basic *type token ratio* index to provide basic information about the lexical diversity in the students' texts.

Sentence overlap. The TAACO tool provides multiple sentence overlap indices to assess *local* text cohesion. These indices calculate lemma overlap between two adjacent sentences and among three adjacent sentences. TAACO also provides average overlap scores across the text for lemma overlap, content word lemma overlap, and lemma overlap for POS tags, such as nouns, verbs, adjectives, adverbs, and pronouns. Finally, TAACO calculates binary overlap scores for all features; these scores indicate if there is or isn't (i.e., 1 or 0) any overlap between adjacent sentences. Overall, overlap indices tend to be positively related to measures of cohesion (see, e.g., McNamara, Louwerse, et al., 2010); however, they typically are unrelated to measures of text *coherence* (Crossley & McNamara, 2010, 2011).

In the current study, we assessed local cohesion in students' texts using three measures of sentence overlap: *adjacent sentence overlap (all words)*, *adjacent sentence overlap (content words)*, and *adjacent sentence overlap (function words)*. These three indices all provide information about the local cohesion established between sentences in a text. We included the content and function word indices to provide more specific information about the type of local cohesion found (or not) in a given text. Content word overlap would indicate that similar topics are being discussed in adjacent sentences, whereas function word overlap would be more indicative of similar rhetorical information and sentence structures.

Connectives. The TAACO tool also includes a number of connective indices to measure local cohesion. A number of these indices are based on indices found in the Coh-Metrix tool (McNamara et al., 2014) and can be theoretically described according to two dimensions. The first connective dimension differentiates between positive and negative connectives, whereas the second dimension relates to previously defined categories of cohesion (Halliday & Hasan, 2014; Louwerse, 2001), such as temporal, additive, and causative connectives. Previous research has found negative or no correlation between these indices and measures of writing quality and coherence (Crossley & McNamara, 2010, p. 20, 2011).

TAACO includes multiple new connective indices that are based on the rhetorical purposes of connectives (see Crossley et al., 2015 for more thorough descriptions). Some of these indices have demonstrated positive relations with measures of cohesion in previous studies (McNamara, Louwerse, et al., 2010), but typically do not significantly relate to measures of coherence (Crossley & McNamara, 2010, 2011).

We analyzed three connective indices in the current study: *basic connectives*, *sentence linking connectives*, and *reason and purpose connectives*. These three indices all provide information about the local cohesion established between sentences in a text. However, they differ from basic sentence overlap because they describe how links are being established at the sentence level. We included the basic connectives index to provide an overall measure of the connectives present in a students' text. Additionally, we included the sentence linking (e.g., *nonetheless*) and reason and purpose (e.g., *hence*) connectives to indicate local cohesion that is being established for specific rhetorical purposes.

Paragraph overlap. Finally, in addition to the local cohesion indices, TAACO calculates multiple paragraph overlap indices to assess *global* cohesion. These indices include lemma overlap between two adjacent paragraphs and among three adjacent paragraphs. These indices are based on the same features as the sentence overlap indices (i.e., average and binary lemma overlap, content word lemma overlap, and lemma overlap for POS tags). Previous research suggests that paragraph overlap indices are positively related to text coherence (Crossley & McNamara, 2011).

We assessed the global cohesion of students' texts using three measures of paragraph overlap: *adjacent paragraph overlap (all lemmas)*, *adjacent paragraph overlap (content lemmas)*, and *adjacent paragraph overlap (function lemmas)*. These three indices provide information about the global cohesion established between paragraphs in a text. Similar to the sentence overlap indices, we included the content and function word indices to provide more specific information about the type of global cohesion found (or not) in a given text. Additionally, we investigated overlap among *lemmas*, as opposed to explicit words because we were interested in the degree to which paragraphs overlapped in their general meaning (i.e., global overlap/cohesion), as opposed to specific words.

Findings

Rubric scores were correlated with the TAACO indicators as indicated in Table 1. The significant relationships between rubric facets and TAACO indicators vary across rubric facets, indicating that along with some general relationships, there were distinct textual features associated with each facet of the rubric design, as discussed further below.

Table 1: TAACO Indicators

	§1: Basic Indices		
Rubric Feature	Number of Words	Number of Types	Type/Token Ratio
Synthesis	-.117*	-.177**	-.071
Topic Coverage	.447**	.487**	-.271**
Source Diversity	.395**	.456**	-.220**
Source Quality	.137*	.089	-.243**
Rubric Feature	§2: Sentence Overlap Indices		
	Adjacent Overlap (All)	Adjacent Overlap (Content)	Adjacent Overlap (Function)
Synthesis	.260**	.231**	.059
Topic Coverage	-.161**	-.134*	-.029
Source Diversity	-.120*	-.117*	.011
Source Quality	.164**	.170**	-.002
Rubric Feature	§3: Connectives Indices		
	Basic Connectives	Sentence Linking	Reason and Purpose
Synthesis	.133*	.251**	.147*
Topic Coverage	-.137*	-.164**	-.089
Source Diversity	.038	-.161**	-.168**
Source Quality	-.043	-.003	.118*
Rubric Feature	§4: Paragraph Overlap Indices		
	Adjacent Overlap (All)	Adjacent Overlap (Content)	Adjacent Overlap (Function)
Synthesis	.035	.113	-.027
Topic Coverage	.119*	.101	.126*
Source Diversity	.150*	.110	.173**
Source Quality	.130*	.187**	.080

Synthesis. Across relationships to the synthesis facet, there were a number of significant relationships, which can be aligned with the theorized account above. Longer texts tended to include less synthesis (§1), with more sentence-level (but not paragraph-level) overlap (§2 & 4) and associations to sentence-level and purposeful connectives (§3), indicating higher levels of local cohesion establishing links between connected ideas through use of overlapping words (between sentences) and explicit links (connectives), but perhaps thematic shifts between paragraphs (with no relationship to paragraph level indices).

Topic coverage. Across relationships to the topic coverage facet we see strong relationships (§1) to number of words and types but negative relationship to type/token ratio indicating that lexical diversity, or information given – rather than number of words per se – is related to this facet. The facet is negatively related to local cohesion (§2) and connectives (§3), indicating that higher topic scores perhaps tended to involve more ‘listing’ of claims from sources, with less integration of those claims on a local level (a feature observed in the scoring exercise) although there is some paragraph level (§4) cohesion perhaps indicating general themes threaded through participant outputs.

Source diversity. Across relationships to the source diversity facet we see similar relationships to those indicated in topic coverage. In both cases, a level of global cohesion is indicated at the rhetorical level; that is, they are maintaining similarities in the way that they are talking, even where talking about different content. However, differences can be seen in task approach to developing cohesion such that those scoring high on topic coverage tended to jump between listed ideas, while source diversity scores might be more related to establishing clusters of claims for which multiple sources were cited (and logical connectives used to link these).

Source quality. Across relationships to the source quality facet we again see relationships to lexical diversity (or information given) in the type/token ratio (§1), and (as in synthesis) a relationship to sentence overlap (§2) indicating that local cohesion was being built (suggesting local argumentation focused on specific topics). We also see associations to paragraph overlap (§4) indicating that those who evaluated tended to build a cohesive

argument through their text, making purposeful connections (§3) between sentences. Thus while connectives generally are associated with synthesis (as above), evaluation (in source quality writing) is associated only with purposeful connectives.

Conclusions and implications

Being able to read, select information from, critique, and synthesis from multiple – oft competing – documents is an important skill, which often manifests in the writing of reports. Understanding the textual features of such written outputs is important for developing techniques to support writing tasks, including the potential of automated or semi-automated feedback to students through the use of NLP technologies. The potential, then, is for the assessment of written texts along an epistemic rubric, in which particular textual moves or features are associated with particular epistemic stances or cognitions. This work has given a preliminary demonstration of the potential conceptual and empirical alignment between such features. The potentials drawn highlight some significant relationships, although their sizes are generally small, but does not discuss the range of non-significant relationships or relationships between textual features that mediate rubric scores; further work should analyse this issue. We demonstrate that different cohesion categories varied in their relation to scores on an epistemically-aligned rubric, in line with previous work finding that variability in cohesion can be indicative of differences in cognitive processes (Allen, McNamara, & McCrudden, 2015). Overall, this work demonstrates the importance of investigating the fine-grain properties of students' writing. Further work should examine these differences using deeper analyses and study designs to probe varying cohesion types at a fine grain.

References

- Allen, L. K., McNamara, D. S., & McCrudden, M. T. (2015). Change your Mind: Investigating the Effects of Self-Explanation in the Resolution of Misconceptions. Presented at the Society for Text & Discourse, Minneapolis, Minnesota, USA.
- Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences, 30*, 64–76. <http://doi.org/10.1016/j.lindif.2013.01.007>
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. O'Reilly Media, Inc.
- Bråten, I. (2008). Personal Epistemology, Understanding of Multiple Texts, and Learning Within Internet Technologies. In M. S. Khine (Ed.), *Knowing, Knowledge and Beliefs* (pp. 351–376). Dordrecht: Springer Netherlands.
- Bråten, I., Braasch, J. L. G., Strømsø, H. I., & Ferguson, L. E. (2014). Establishing Trustworthiness when Students Read Multiple Documents Containing Conflicting Scientific Evidence. *Reading Psychology, 0*(0), 1–35. <http://doi.org/10.1080/02702711.2013.864362>
- Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48–70. <http://doi.org/10.1080/00461520.2011.538647>
- British Psychological Society. (2014). Code of Human Research Ethics. British Psychological Society.
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2015). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods, 1*–11.
- Crossley, S. A., & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 984–989). Austin, TX: Cognitive Science Society.
- Crossley, S. A., & McNamara, D. S. (2011). Text coherence and judgments of essay quality: Models of quality and coherence. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 1236–1241).
- Crossley, S. A., & McNamara, D. S. (2014). Developing Component Scores from Natural Language Processing Tools to Assess Human Ratings of Essay Quality. In *The Twenty-Seventh International Flairs Conference*.
- de Waard, A., Buitelaar, P., & Eigner, T. (2009). Identifying the epistemic value of discourse segments in biology texts. In *Proceedings of the Eighth International Conference on Computational Semantics* (pp. 351–354). Stroudsburg, PA, USA: Association for Computational Linguistics.
- Ferguson, L. E. (2014). Epistemic Beliefs and Their Relation to Multiple-Text Comprehension: A Norwegian Program of Research. *Scandinavian Journal of Educational Research, 0*(0), 1–22. <http://doi.org/10.1080/00313831.2014.971863>

- Foltz, P. W., Britt, M. A., & Perfetti, C. A. (1996). Reasoning from multiple texts: An automatic analysis of readers' situation models. In G. W. Cottrell (Ed.), *Proceedings of the 18th Annual Cognitive Science Conference* (pp. 110–115). Lawrence Erlbaum, NJ.
- Goldman, S. R., Lawless, K., Pellegrino, J. W., & Gomez, K. (2012). A Technology for Assessing Multiple Source Comprehension: An Essential Skill of the 21st Century. In M. Mayrath (Ed.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research*. Information Age Publishing (IAP).
- Goldman, S. R., Ozuru, Y., Braasch, J. L. G., Manning, F. H., Lawless, K. A., Gomez, K. W., & Slanovits, M. (2011). Literacies for learning: A multiple source comprehension illustration. In N. Stein L. & S. Raudenbush (Eds.), *Developmental science goes to school: Implications for policy and practice* (pp. 30–44). Abingdon, Oxon: Routledge.
- Groza, T., Handschuh, S., & Bordea, G. (2010). Towards automatic extraction of epistemic items from scientific publications. In *Proceedings of the 2010 ACM Symposium on Applied Computing* (pp. 1341–1348). New York, NY, USA: ACM. <http://doi.org/10.1145/1774088.1774377>
- Halliday, M. A. K., & Hasan, R. (2014). *Cohesion in english*. Routledge.
- Hastings, P., Hughes, S., Magliano, J. P., Goldman, S. R., & Lawless, K. (2012). Assessing the use of multiple sources in student essays. *Behavior Research Methods*, 44(3), 622–633. <http://doi.org/10.3758/s13428-012-0214-0>
- Knight, S., & Littleton, K. (2015). Developing a multiple-document-processing performance assessment for epistemic literacy. Presented at the The 5th International Learning Analytics & Knowledge Conference (LAK15): Scaling Up: Big Data to Big Impact, Poughkeepsie, NY, USA.
- Kobayashi, K. (2014). Students' consideration of source information during the reading of multiple texts and its effect on intertextual conflict resolution. *Instructional Science*, 42(2), 183–205.
- Louwerse, M. (2001). An analytic and cognitive parametrization of coherence relations. *Cognitive Linguistics*, 12(3), 291–316.
- McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392.
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic Features of Writing Quality. *Written Communication*, 27(1), 57–86. <http://doi.org/10.1177/0741088309351547>
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, UK: Cambridge University Press.
- McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing Linguistic Features of Cohesion. *Discourse Processes*, 47(4), 292–330. <http://doi.org/10.1080/01638530902959943>
- Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38(11), 39–41.
- Mozilla. (2014). *Firefox*. Retrieved from <https://www.mozilla.org/en-US/firefox/new/>
- OECD. (2013). PISA 2015: Draft reading literacy framework. OECD Publishing.
- OECD, & Statistics Canada. (2010). Literacy in the Information Age - Final Report of the International Adult Literacy Survey. OECD.
- Rouet, J.-F. (2006). *The Skills of Document Use: From Text Comprehension to Web-Based Learning* (First Edition edition). Mahwah, NJ: Routledge.
- Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). Information Age Publishing (IAP).
- Shah, C. (2010). Coagmento—a collaborative information seeking, synthesis and sense-making framework. *Integrated Demo at CSCW*, 6–11.
- Shah, C. (2014). *Coagmento*. Retrieved from <http://coagmento.org/>
- Simsek, D., Buckingham Shum, S., Sandor, A., De Liddo, A., & Ferguson, R. (2013). XIP Dashboard: visual analytics from automated rhetorical parsing of scientific metadiscourse. Presented at the 1st International Workshop on Discourse-Centric Learning Analytics, Leuven, Belgium.

Acknowledgments

This work was conducted as part of the first author's PhD at the Open University, UK. We thank participating students at Maastricht University, and Dr Dirk Tempelaar there for his collaboration on the work. Our thanks also to Dr Chirag Shah and Matthew Mitsui at Rutgers for their collaboration on the use of Coagmento, and study design.

Visual Augmentation of Deictic Gestures in MOOC Videos

Kshitij Sharma, Faculty of Business and Economics, University of Lausanne, kshitij.kshitij@unil.ch
Sarah D'Angelo, Northwestern University, sdangelo@u.northwestern.edu
Darren Gergle, Northwestern University, dgergle@northwestern.edu
Pierre Dillenbourg, École Polytechnique Fédérale de Lausanne, pierre.dillenbourg@epfl.ch

Abstract: We present an eye-tracking study to compare different modalities for visual augmentations of the teacher's explicit deictic gestures on a video lecture. We compared three visualizations: 1) hand gestures with a pointer, 2) gaze overlay, and 3) no-augmentation baseline. We investigate the teacher-student pair in a video-based learning context as an abstraction of an expert-novice pair where the goal is to attain a high level of shared understanding. The key phase of having a shared understanding is to have a common ground between the pair. Previous studies showed that explicit deixis plays a major role in initiating and maintaining common ground. This led us to hypothesize that augmenting videos with teacher's deictic gestures might help students perform better. We found that augmenting the video with teacher's gaze results in higher learning gain than no visualization. Moreover, gaze visualization also helped students in maintaining longer attention spans than hand gestures.

Keywords: eye-tracking, MOOCs, learning analytics, online deixis

Introduction

Lecture material can be displayed in a variety of formats to engage students and maintain their attention. Common augmentations include overlaying a pointer controlled by the teacher on the slides (Jermann, 2014) or including a video of the teachers face in the corner (Kizilcec et al., 2014). Now that eye tracking is becoming more accessible for educational applications, displaying teacher's gaze is another way to provide visual aids for students (Sharma et al., 2015).

Student's attention has been found to be correlated with performance in many studies related to visual tasks (Yantis and Jonides, 1984; Prinzmetal et. al., 1986; Juola et. al., 1991). In a visual comparison of two line segments Prinzmetal et. al. (1986) and Juola et. al., (1991) showed that the more attentive participants were more often correct in selecting the longer line segment. Yantis and Jonides (1984) found similar results in visual perception tasks. In a classroom, attention is: "*listening, sitting and working on assigned tasks*" -Homes et. al. (2006). In the context of academic performance previous research has shown strong association between students' attention and academic performance (Finn, 1989). In the context of video-based learning, one could use eye-tracking data to measure the amount of time the learner is paying attention to the elements that the teacher is referring to, verbally or through deictic.

In this study, we are particularly interested in maintaining attention in complex visual environments in which deictic references become important. We designed a video lecture on cloud identification to evaluate the presence of visual aids with highly visual and linguistically complex content. For this paper we will only consider the visual aspects of the content. In the case of a teacher-student dyad, following teacher's deictic references was correlated with the performance (Sharma et al., 2014). The main objective for any teacher-student dyad is to create a shared understanding of the content (a teacher-student dyad is a special case of the expert-novice dyad). Usually in dyadic interactions the basis of shared understanding is the common ground between the participants. Explicit deictic gestures play a key role in initiating and maintaining this common ground (Clark and Brennan, 1991).

In the present contribution, we use eye-tracking to capture student gaze patterns and compute their perceptual "with-me-ness" or the extent to which they follow along with the teacher's explicit deictic cues (Sharma et al, 2014); as well as their attention distribution across the lecture material. Additionally, we use posttest scores as an indicator of learning gain. Eye-tracking provides unprecedented access to the students' attention in video-based learning environments. Previous dual eye-tracking studies (Cherubini et al., 2008; Jermann and Nüssli, 2012; Richardson et al., 2007) have shown that eye-tracking could be used as an evaluation tool for the effectiveness of dyadic interaction. Cherubini et al (2008) found that the misunderstandings in a collaborative problem-solving task were correlated to the difference in participants' gaze patterns. Additionally, Richardson et al (2007) found that the cross-recurrence of a speaker-listener pair is correlated to the listener's comprehension. Furthermore, Jermann and Nüssli (2012) found that the cross-recurrence was correlated with the pair's collaboration quality. In this work we evaluate the presence of different deictic visualizations (pen pointer or gaze overlay) in a video lecture on learning gain and perceptual with-me-ness. We found that showing teacher's gaze

to students improves learning gain compared to no visual aid. Additionally, the presence of gaze information increased perceptual with-me-ness compared to the pen pointer visualization.

Related work

The use of eye tracking in online education has provided researchers with insights about students' learning processes and outcomes. For example, Van Gog et al. (2005) used eye-tracking data to differentiate the expertise levels in the different phases of an electrical circuit-troubleshooting problem and concluded that experts focused more on the problematic area than the novices. In another experiment, where the participants had to learn a game, Alkan and Cagiltay (2007) found that the good learners focused more on the contraption areas (areas that appeared strange or unnecessarily complicated) of the game while they think about the possible solutions. Additionally, Mayer (2010) summarized the major results of research on eye tracking in online learning with graphics and concluded that there was a strong relation between fixation durations and learning outcomes; and visual signal guided students' visual attention. Understanding novice and expert gaze patterns in online education has informed a number of interventions to improve student learning. For example, Van Gog et al. (2009) found that displaying an expert's gaze during problem solving guided the novices to invest more mental effort than when there was no gaze displayed.

We know from previous eye-tracking research that speakers looked at the objects they refer to just before pointing and verbally naming the objects (Griffin and Bock, 2000). Listeners on the other hand, looked at the referred objects shortly after seeing the speaker point and refer to the objects (Allopenna et al., 1998). D.C. Richardson and Kirkham (2007) showed that the listeners who were better at attending the references made by the speaker were also better at understanding the context of the conversation. One way to aid the listeners attending the reference in a better way could be to display where the speaker is looking at. This might help the listeners in a better disambiguation of the complex references (Gergle and Clark, 2011, Hanna and Brennan, 2007). In the case of complex stimulus displaying the gaze of speaker made the disambiguation of the references even easier (Prasov and Chai, 2008).

Gaze contingent experiments are at the proactive side of the eye-tracking technology. These experiments consist in displaying the gaze of collaborating partners to each other; or displaying the gaze of an expert to a novice in order to teach the novice (Chetwood et al., 2012). Another modality of gaze contingency is using gaze as a mode of communication. In a collaborative "Qs-in-Os" search Brennan et al. (2008) showed that the sharing gaze information between collaborating partners resulted in a strategy of division of labor as effective as if the partners were talking face to face. Displaying the gaze of speaker helped the listener in deciphering the references (Gergle and Clark, 2011, Hanna and Brennan, 2007). Moreover, gaze of speaker made it easier for the listener in deciphering the references in situations with high ambiguity (Prasov and Chai, 2008).

Current study

Research question

Previous work has shown that with-me-ness is correlated with the learning gains (Sharma et al, 2014). When students were provided feedback on their with-me-ness and it improved their learning gains (Sharma, 2015). Additionally, it has been shown that putting teacher's gaze online correlates with a specific video navigation pattern (fewer and less frequent pauses, fewer backward jumps) that signifies the low perceived difficulty by the students (Sharma et al, 2015). However, these two results had two different modalities for augmenting the video with teacher's explicit deictic references: displaying the pointer and displaying the gaze. This study compares those two modalities with a baseline of no visual augmentation. Specifically, we address the following research question: How does deixis visualization affect the learning gain and students' attention?

Participants

Forty-three university students participated in the study; the mean age was 19.18 (sd = 2.07). Six of the participants were females and thirty-seven were males, all students were bachelor status. Informed consent was obtained for all participants and they received an equivalent of 20\$ compensation for their participation.

Procedure

Participants were informed that they would be listening to a lecture on cloud identification and that their gaze would be recorded using a remote eye tracking system (SMI RED 250). Participants were asked to use a chin rest to keep their head stable and wear headphones. Before the experiment participants eye gaze was calibrated using a 5-point calibration. The start of the study began with ten-question cloud identification pretest, which was

followed by the cloud identification lecture and concluded with ten-question cloud identification posttest. Subjects were informed that they should complete the tests and listen to the lecture at their own pace and they were able to pause the lecture and go back and forward in time as much as they wanted.

Task

The cloud identification pre and posttests consisted of ten multiple-choice questions. The ten cloud types covered in the lecture appeared on each test once, five of the cloud types were graphically represented and five were represented with photographs. The graphical representations and photograph representations were swapped for the pre and post test so each cloud type was represented both graphically and pictorially but no stimuli were repeated. Participants were asked to select the correct cloud type from four choices, they were instructed not to guess on the answers and if they did not know the correct answer to skip the question. The cloud lecture was 11 minutes and 37 seconds and consisted of seventeen slides. Each of the ten cloud types had an individual slide that contained the cloud name, two descriptors based on altitude and feature, and two representations a photograph and graphical depiction. The average time for a cloud content slide was 43 seconds (sd = 6.47 seconds). The lecture started and concluded with summary slides containing the ten types of clouds. The lecture content explained how to identify ten cloud types based on their distinguishing characteristics such as altitude in which they occur (i.e. high, medium, low) and describing features (i.e. puffy or layered). For example, the altocumulus cloud was described as a mid altitude cloud composed of puffy grey and white patches that are most likely to be seen with other clouds.

Measures

The participant's gaze was recorded for the duration of the study including the pretest, lecture, and posttest. All responses to the tests and interactions with the video (i.e. pauses) were recorded. Additionally, the gaze patterns for the teacher were recorded for the duration of the lecture, while the content was being recorded.

Independent variable

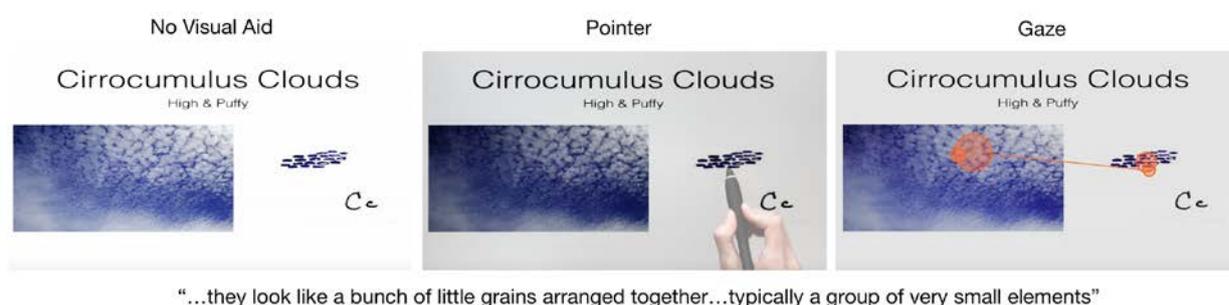


Figure 1. Visual aid conditions.

Deixis visualization

We manipulated the availability of deixis visualization as a between subjects variable. There are three conditions of deixis visualizations: a pen pointer representation, gaze representation, or no additional visual aid (Figure 1). The lecture contained 242.6 seconds of pointer information; we replaced the same exact video segments with the gaze representation. The control condition does not contain any visual augmentation. Slide and lecture content were identical for all three conditions.

To determine the appropriate amount of gaze information to display, we conducted a small pilot study to evaluate how long the gaze trails should be visible on screen. Five participants viewed the lecture with both a 5 second gaze trail and a 2 second gaze trail alternating every 60 seconds of deixis visualization matching the pointer condition. Participants were asked if they noticed a difference in the gaze representation and if they did to state their preference for which gaze representation was most appropriate. The majority of the participants preferred the 2 second gaze trail, stating that the 5 second trail was disruptive and occluding too much content, therefore we used 2 second trail in the gaze condition.

Dependent variables

Learning gain

Learning gain was evaluated by the post test scores. Participants were not familiar with the content before the study and we observed a floor effect in the pre-test scores (median = 0, mean = 0.83/10, 27 out of 36 participants scored 0 or 1 in the pretest) therefore we do not consider pre-test scores in our analysis of learning gain. Since identifying clouds by name is a complex task, we developed a scoring rubric in order not to be restricted by the requirement of memorizing the cloud names. Students received one point for marking the correct name of the cloud. A half point was given to answers that had one of the correct characteristics of the cloud. For example if the correct answer is altocumulus students receive a half point for answers containing the prefix “alto” (indicating a mid-altitude cloud) or answers containing “cumulus” (indicating a puffy cloud shape). Zero points were given to incorrect answers that did not share correct characteristics.

Table 1: Example scoring grid for altocumulus cloud

	Cumulus or “Puffy”	Stratus or “Layered”	“Exceptions”	
High Altitude	Cirrocumulus (.5)	Cirrostratus (0)	Cirrus (0)	Cumulonimbus (0)
Middle Altitude	Altocumulus (1)	Altostratus (.5)	Nimbostratus (.5)	
Low Altitude	Cumulus (.5)	Stratus (0)	Stratocumulus (0)	

With-me-ness

With-me-ness is the extent to which the students succeed in following the teacher’s dialogues and deictic gestures on the screen. In eye-tracking terms, with-me-ness captures: “how much time a student spent looking at the part of the display that the teacher is talking about?” With-me-ness is defined at two levels: perceptual and conceptual. There are two ways a teacher may refer to an object: with deictic gestures, generally accompanied by words (“here”, “this variable”) or only by verbal references (“the counter”, “the sum”). Perceptual with-me-ness measured if the students looked at the items referred to by the teacher through deictic acts. Conceptual with-me-ness was defined using the discourse of the teacher: did students look at the object that the teacher was verbally referring to, i.e., that the teacher was referring to a set of objects that were logically or semantically related to the concept he was teaching. In the present study, since we are only interested in the effect of the deixis visualization on the students’ gaze patterns, we will use the perceptual level of with-me-ness only. The perceptual “with-me-ness” has 3 main components: entry time, first fixation duration and the number of revisits. (a) Entry time was the temporal lag between the times a referring pointer appeared on the screen and stops at the referred site (x,y) and the time student first looked at (x,y). (b) First fixation duration was how long the student gaze stopped at the referred site for the first time. (c) Revisits were the number of times the student’s gaze came back to the referred site.

Results

Time on video lecture: We observe no effect of time spent on the video-lecture on learning gain ($r(43) = -0.03$, $p = .84$). Moreover, we do not observe a difference in the time spent on the video between experimental conditions ($F[2, 37] = 2.67$, $p = .10$)

Learning gain: The learning gain in the gaze condition is significantly higher than the learning gain in no visual aid condition. However, there is no difference in learning gain across the pen pointer and the no visual aid condition. Table 2 shows the pairwise ANOVA results.

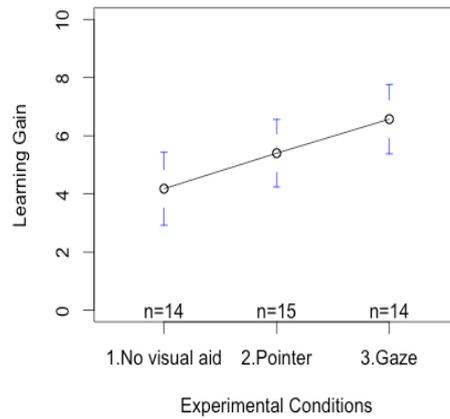


Figure 2. Learning gain.

Table 2: Pairwise ANOVA for learning gain

Condition pair	ANOVA effect size	p-value
Pointer vs. No visual aid	1.22	.28
Gaze vs. No visual aid	2.39	.01
Gaze vs. Pointer	1.17	.31

Perceptual With-me-ness: A one-way ANOVA, without the assumption for equal variances, shows that the perceptual with-me-ness is significantly higher in the gaze condition than that in the per pointer condition (gaze mean = 0.25, sd = 0.24, pointer mean = 0.09, sd = 0.09, no visual aid mean = 0.09, sd = 0.1, Figure 3). Table 3 shows the pairwise ANOVA results. In all the cases, i.e., pointer, gaze and no visual aid condition, we logged the exact point of the reference and then defined a circular area of 50 pixel diameter to compute the perceptual wit-me-ness.

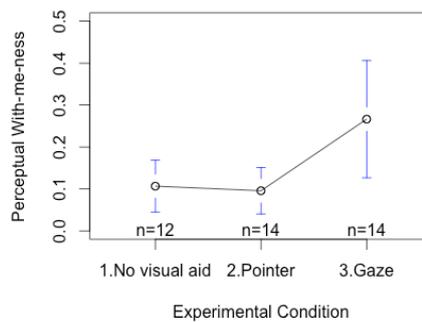


Figure 3. Perceptual With-me-ness.

Table 3: Pairwise ANOVA for perceptual with-me-ness

Condition pair	ANOVA effect size	p-value
Pointer vs. No visual aid	0.01	.98
Gaze vs. No visual aid	0.15	.05
Gaze vs. Pointer	0.17	.02

Gaze on video slides: We observe no difference in amount of time spent on the slides explaining individual cloud types, however, we do see a difference in the time spent looking at relevant content in the summary slides based on visualization condition. Participants who were shown the teacher’s gaze overlay spend significantly more time looking at the specific cloud types in the summary slide compared to the no visualization condition and the pen pointer condition. There is no significant difference between the pen pointer condition and the no visualization condition. Table 4 shows the pairwise ANOVA results.

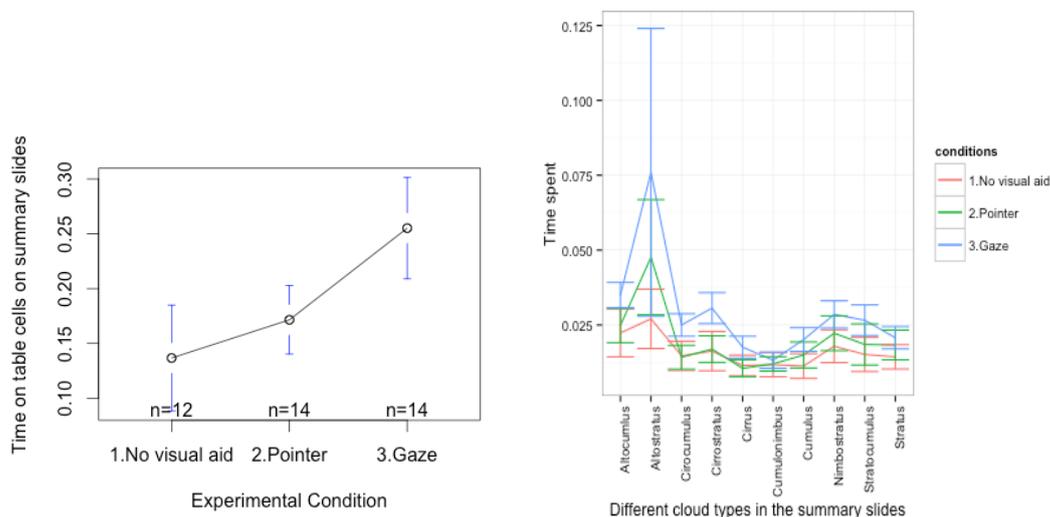


Figure 4. Time spent on summary slide and each cloud type.

Table 4: Pairwise ANOVA for time spent on summary slides

Condition pair	ANOVA effect size	p-value
Pointer vs. No visual aid	0.46	.50
Gaze vs. No visual aid	7.09	.01
Gaze vs. Pointer	6.02	.02

Discussion

The results of this study suggest that showing the teacher’s gaze to students when making explicit references to information on the slides can be useful for students. As a visual aid, gaze highlights important areas on the slides that the teacher is explaining to students. Additionally, gaze provides more information than a pen pointer based representation, which may have contributed to the positive effects of sharing gaze information. Although we controlled for the time both visual aids were displayed on screen, the two second gaze trail allowed for multiple points of reference to be displayed at once while the pen was limited to a single point. For example, in one frame of the gaze condition the teacher can visually compare multiple areas on the slides. Additionally, gaze captures potentially unintentional signals that the teacher uses such as looking back at the name of cloud that may help students connect different knowledge points. This could have been less likely in the pen pointer condition, as the teacher intentionally controls the pointing behavior.

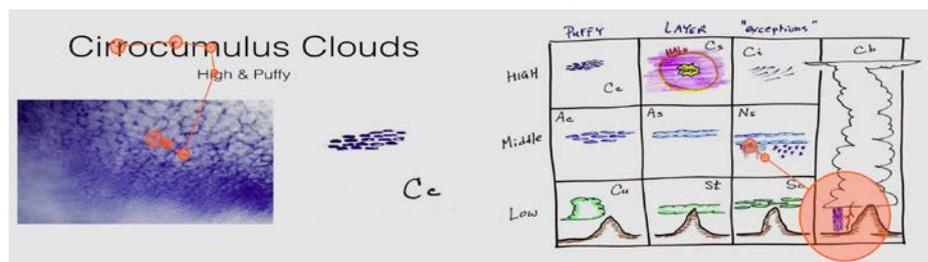


Figure 5. Examples of gaze use cases.

Since we do not see an effect of time spent on video, the increase in learning gain is most likely to be a result of where students were paying more attention in the lecture. We see that students spend more time looking at the cloud representations in the summary slides. This suggests that they may have spent more time comparing the distinguishing characteristics for cloud types, which could have contributed to improved performance on the posttest. The gaze representation may have been particularly useful, for the summary slides, because it is a more targeted representation and contains more temporal information about teacher's references. Whereas, the pen pointer enters from the bottom of the screen, which may distract attention to irrelevant areas of the slide. Another plausible explanation for the higher learning gain in the gaze condition could be higher levels of perceptual with-me-ness. Since we observe higher perceptual with-me-ness in the gaze condition; it suggests that students were following the teacher's gaze, which helped maintain attention to important part of the content. We see in Figure 3 that the overall level of perceptual with-me-ness is low (the students follow the teacher's references about 50% of the time at most). This shows that the students are not mechanically following the visual deictic, but using it as a support for following the teacher. These results are coherent with the results found by Sharma (2015).

Conclusions

Sharing teacher's gaze with students has a lot of potential for augmenting the videos with external information in online education. In environments like MOOCs gaze can be a practical addition to lecture content since the recording of content is not real time and eye tracking technology is becoming more accessible. Our results indicate that sharing gaze information improves learning gain compared to no visual aid and maintains students' attention for longer periods of time compared to the pen pointer condition. Therefore it could be a useful addition to visually rich and complex lecture content. Future analysis will investigate the relationship between linguistic complexity and student gaze given the visual aid representation. We will also investigate the amount of attention shift in the cloud specific slides and its effect on students' gaze on posttest items.

References

- Alkan, S., & Cagiltay, K. (2007). Studying computer game learning experience through eye tracking. *British Journal of Educational Technology*, 38(3), 538-542.
- Alloppenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, 38(4), 419-439.
- Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106(3), 1465-1477.
- Cherubini, M., Nüssli, M. A., & Dillenbourg, P. (2008, March). Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. In Proceedings of the 2008 symposium on Eye tracking research & applications (pp. 173-180). ACM.
- Chetwood, A. S., Kwok, K. W., Sun, L. W., Mylonas, G. P., Clark, J., Darzi, A., & Yang, G. Z. (2012). Collaborative eye tracking: a potential training tool in laparoscopic surgery. *Surgical endoscopy*, 26(7), 2003-2009.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. Perspectives on socially shared cognition, 13(1991), 127-149.
- Finn, J. D. (1989). Withdrawing from school. Review of educational research, 59(2).
- Gergle, D., & Clark, A. T. (2011, March). See what i'm saying?: using Dyadic Mobile Eye tracking to study collaborative reference. In Proceedings of the ACM 2011 conference on Computer supported cooperative work (pp. 435-444). ACM.
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological science*, 11(4), 274-279.

- Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, 57(4), 596-615.
- Holmes, R. M., Pellegrini, A. D., & Schmidt, S. L. (2006). The effects of different recess timing regimens on preschoolers' classroom attention. *Early Child Development and Care*, 176(7).
- Jermann, P., & Nüssli, M. A. (2012, February). Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 1125-1134). ACM.
- Jermann, P. (2014). The EPFL MOOC Factory. Experience track contribution in EMOOCs 2014, Lausanne.
- Juola, J. F., Bouwhuis, D. G., Cooper, E. E., & Warner, C. B. (1991). Control of attention around the fovea. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1).
- Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014, April). Showing face in video instruction: effects on information retention, visual attention, and affect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2095-2102). ACM.
- Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and instruction*, 20(2), 167-171.
- Prasov, Z., & Chai, J. Y. (2008, January). What's in a gaze?: the role of eye-gaze in reference resolution in multimodal conversational interfaces. In *Proceedings of the 13th international conference on Intelligent user interfaces* (pp. 20-29). ACM.
- Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration? *Journal of Experimental Psychology: Human Perception and Performance*, 12(3), 361.
- Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination common ground and the coupling of eye movements during dialogue. *Psychological science*, 18(5), 407-413.
- Sharma, K., Jermann, P., & Dillenbourg, P. (2015). Displaying Teacher's Gaze in a MOOC: Effects on Students' Video Navigation Patterns. In *Proceedings of the 10th European Conference on Technology Enhanced Learning*.
- Sharma, K. (2015). Gaze Analysis methods for Learning Analytics.
- Sharma, K., Jermann, P., & Dillenbourg, P. (2014). "With-me-ness": A gaze-measure for students' attention in MOOCs. In *International conference of the learning sciences*.
- Van Gog, T., Paas, F., van Merriënboer, J. J., & Witte, P. (2005). Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, 11(4), 237.
- Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior*, 25(3), 785-791.
- Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human perception and performance*, 16(1).

Designing the Idea Manager to Integrate STEM Content and Practices During a Technology-Based Inquiry Investigation

Erika D. Tate, bluknowledge llc, erika@bluknowledge.com

Mingyu Feng, Center for Technology in Learning, SRI International, mingyu.feng@sri.com

Kevin W. McElhaney, Center for Technology in Learning, SRI International, kevin.mcelhaney@sri.com

Abstract: This design study investigates the role of the Idea Manager, a formative technology-based learning tool, in student learning that integrates STEM practices and disciplinary content. We coordinated curriculum and assessment frameworks to design instructional features of a STEMGenetics unit that engages middle school students in a set of interrelated STEM practices: use a model to construct an explanation of the trait expression mechanism in rice plants. Our analysis of 99 student team responses to three embedded assessments generated formative evidence that delineated three tiers of explainers (non-mechanistic, partial mechanistic, and mechanistic). These students differed in their knowledge and abilities to perform these interrelated STEM practices. We discuss implications of the findings for how (a) teachers can differentiate instruction for students of all levels and across a range of STEM disciplines and interrelated sets of practices and (b) technology tools can be designed to support teachers.

Keywords: science practices, technology, design, student learning, life science

Introduction

A growing emphasis on STEM (science, technology, engineering, and mathematics) requires researchers and practitioners to design learning activities that promote an integrated understanding of STEM topics, applications, and approaches (NRC, 2012). In the past, student learning in STEM has been limited to (a) remembering a set of isolated scientific and mathematical facts often perceived as irrelevant and (b) applying this nominal understanding to prescribed calculations and procedures with limited opportunity for reflection and revision of thinking. New standards for STEM teaching, learning, and assessment call for students to engage in a variety of STEM practices (e.g., use a model, such as a scientific visualization) in order to understand content comprised of disciplinary core ideas (e.g., heredity) and crosscutting concepts (e.g., cause and effect) that bridge STEM fields. These standards also require students to draw upon this more robust understanding of STEM to design and test solutions to existing or potential real-world challenges (NRC, 2012; NGSS, 2013).

These shifts in policy and practice have created a great demand for the learning sciences community to transform STEM teaching and learning. Researchers have problematized what it means to teach and learn STEM practices, asserting that these practices are valuable dimensions of disciplinary work that are learnable, interrelated, and benefit from ongoing evaluation and critique. This requires researchers and practitioners to reframe STEM teaching and learning and generate evidence for how teachers, curricula, and technology can shape STEM participation across different learning settings (e.g., Erduran & Dagher, 2014; Ford, 2015; Stroupe, 2015). This need presents opportunities to bridge curriculum and assessment frameworks that inform the principled design of STEM learning environments intended to engage students in STEM practices and deepen their STEM disciplinary content knowledge (Debarger et al., 2014).

In this study, we focus on a Grade 7 STEMGenetics curriculum unit developed at Michigan State University and SRI International that builds on students' understandings of key genetics concepts. The development of the unit was informed by knowledge integration (KI) and evidence-centered design (ECD) perspectives (Debarger et al., 2014). We investigate the role of Idea Manager (McElhaney, Matuk, Miller, & Linn, 2012), a formative technology-based learning tool embedded in the unit, to answer the following research questions: (1) How does the Idea Manager scaffold middle school students to integrate STEM practices and disciplinary content during a technology-based inquiry investigation? (2) How does the formative evidence generated with the Idea Manager differentiate students' ability to engage in interrelated STEM practices: use a model and construct an explanation?

This study contributes to the growing body of research that seeks to (a) identify the knowledge and abilities needed to perform STEM practices and (b) develop a more nuanced understanding of the types of design features that can support learning of STEM content and practices (NRC, 2012; McElhaney, 2012; Debarger, 2014; Stroupe, 2015; Erduran & Dagher, 2014). The remainder of this paper (1) describes our coordinated theoretical approach, (2) details the methodology, and (3) presents findings about how the Idea Manager scaffolds student

learning and provides formative evidence that distinguishes proficiencies in STEM practices and content knowledge. We conclude with a set of implications for STEM classroom learning and future research.

Coordinated theoretical approach to designing for integrated STEM learning

We draw upon knowledge integration (KI) and evidence-centered design (ECD) to (a) articulate integrated learning goals and outcomes, (b) define learning engagement and evidence, and (c) organize learning task features in the STEMGenetics learning environment. KI characterizes learning in everyday life and designed learning environments as the opportunity to (a) add ideas about STEM topics, practices, and disciplines to learners' existing repertoire of ideas and (b) revise the connection learners make between existing and new ideas into an increasingly integrated understanding of science. The KI design framework guides the design of coherent STEM instruction (Kali, Linn, & Roseman, 2008; Linn & Eylon, 2006; Lee et al., 2010). The *metaprinciples* (make science accessible, make thinking visible, help students learn from each other, and promote autonomous and lifelong learning) communicate the nature and quality of learning experiences that promote knowledge integration. The *instructional patterns* (e.g., develop a model, engage in argument with evidence) help designers coordinate learning activities that (a) engage students in the knowledge integration *process* (elicit ideas - add ideas - distinguish ideas - reflect on/explain ideas) and (b) scaffold students' participation in STEM practices (Bransford, Brown, & Cocking, 2000; Linn, Davis, & Bell, 2004).

While KI emphasizes deepening students' understanding in a discipline through practice, evidence-centered design (Mislevy & Haertel, 2006) can heighten instructional designers' awareness of how assessments provide evidence for learning as it occurs during instruction. ECD provides a principled framework to analyze the target learning domain and structure the articulation of (1) integrated learning goals that specify the knowledge and abilities of interest, (2) evidence that will be produced in the form of student work products, and (3) design features of learning environments and tasks that elicit this evidence from students (Mislevy & Haertel, 2006; DeBarger, 2014). The coordination of KI and ECD frameworks results in the principled design of instruction and assessment that guide students toward an integrated understanding of STEM content and practices. The instruction includes deliberate learning activities that explicitly engage students in knowledge integration and offer multiple opportunities for teachers to formatively evaluate student progress. With careful design, the coordinated framework allows these learning and assessment opportunities to distinguish separable proficiencies within the domain so that teachers can focus their subsequent instruction where it is most needed.

Design features that scaffold students to integrate STEM content and practices

This research utilizes the Web-based Inquiry in Science Environment (WISE), an online platform aligned with the KI design framework. WISE features a suite of learning tools to support STEM teaching and learning, such as writing prompts (e.g., critique and feedback), activity templates (e.g., inquiry and role play, brainstorm), interactive simulations (e.g., dynamic scientific models), and explanation generation tools (e.g., Idea Manager) (Slotta & Linn, 2009). The Idea Manager includes the (1) Idea Basket, a persistent space where students add and sort information ideas as they learn and (2) Explanation Builder, an organizing space where students (a) consider all the ideas in their idea basket and (b) distinguish and sort relevant ideas into the desired scientific explanation (McElhaney, 2012). We designed instruction using the Idea Manager to (a) guide student engagement in the knowledge integration process, (b) scaffold student learning of the disciplinary core ideas (e.g., trait expression mechanism) and the STEM practices (e.g., use a model and construct an explanation), and (c) offer opportunities to formatively evaluate progress toward integrated STEM understanding. Table 1 organizes the contributions of the KI and ECD design frameworks across three design dimensions and exemplifies resulting design features of the Idea Manager in the Grade 7 STEMGenetics unit.

Learning context

Our collaborative design team of education researchers, curriculum and assessment designers, biologists, technology developers, and teachers designed the Grade 7 STEMGenetics unit to promote students' understanding of inheritance through STEM practices of using models and constructing explanations (NRC, 2012). The unit introduces a situation where the world will be running short of food and fossil fuels in 2052 and students inquire about how to selectively breed for more nutritious rice in order to propose a genetics-driven solution to end world hunger. As students answer the driving question, "How can you use genetics to feed the world in 2052?" they engage in a variety of WISE learning activities, hands-on laboratories, and whole class discussions facilitated by the teacher. This study was based on the second design iteration of the STEMGenetics Grade 7 unit, which has been informed by evidence from the pilot study (DeBarger, 2014) and teacher feedback during a co-design professional learning session.

Table 1: Contributions of KI and ECD design frameworks across design dimensions for instructions

Design Dimensions	Contributions of KI and ECD Design Frameworks	Resulting Design Features of Instruction using the Idea Manager
Learning Goals and Outcomes	KI: Connections between key disciplinary ideas that promote lifelong learning	Learning Performance: Students can use a model showing the mechanistic processes of trait expression to construct an explanation about why an organism has a particular trait
	ECD: Combination of content knowledge and practices as the target of instruction and embedded assessments	
Learning Engagement and Evidence	KI: Learning as engagement through eliciting, adding, distinguishing, and sorting ideas	Learning Engagement: Students (1) observe model to add ideas about trait expression and (2) construct an explanation to distinguish and sort ideas about trait expression
	ECD: Clear articulation of evidence produced by students to indicate progress toward and attainment of learning goals	Learning Evidence: Students (1) use a model of trait expression to document mechanistic elements of the phenomenon of trait expression and (2) construct a scientific explanation integrating observable and non-observable aspects about why an organism has a particular trait
Learning Task Features and Flow	KI: Application of design patterns and tools to promote knowledge integration within and across tasks	Prompts for students to (1) focus students' use of the model on mechanistic aspects of the phenomenon and document them in the idea basket and (2) use idea basket entries to construct an explanation about the phenomenon
	ECD: Specification of instructional and assessment design features to elicit desired evidence	

Learning goals and outcomes

Based on our analysis of the Framework for K-12 Science Education (NRC, 2012), the design team initially developed twelve learning goals for the unit that centered on the disciplinary core ideas that comprise a deep understanding of inheritance for 7th grade students (e.g., trait expression and sexual reproduction) and two STEM practices: developing and using models and constructing explanations. Findings from our pilot study prompted us to recast the twelve learning goals into five learning performances that explicitly consider the application of the STEM practices within genetics (Debarger, 2014). This study focuses on the learning performance: *Students can use a model showing the mechanistic processes of trait expression to construct an explanation about why an organism has a particular trait.*

Learning engagement and evidence

For each learning performance, we articulated statements that clarify the evidence students would need to indicate progress toward the learning goal. The evidence for the focal learning performance included: *students (1) use a model of trait expression to document mechanistic elements of the phenomenon of trait expression and (2) construct a scientific explanation integrating observable and non-observable aspects about why an organism has a particular trait.* To generate this evidence, students were requested to (a) first observe a dynamic scientific model that visibly explained how the structures, functions, and interactions of alleles, messengers, ribosomes, proteins, and starch yield the high-, medium-, or low-nutrition traits observed in rice plants, (b) add their observations as ideas in their idea basket that described how the model elements represented the trait expression mechanism, and (c) use the Explanation Builder to distinguish and sort these model observations into an organizing space, and use selected observations as evidence to build an explanation of the trait expression mechanism in high-nutrition rice plants.

Specific prompts were designed to guide students' observation of the models; students were asked to add at least one idea for each mechanistic element during or after their model observations (Figure 1a). The ideas added at the model steps provide formative evidence for how well students understood the model. Later in the Explanation Builder, students were instructed to (a) consider all the observations in their idea basket and (b) distinguish which ideas were relevant to their scientific explanation of the trait expression mechanism in high-

nutrition rice plants (Figure 1b). As students constructed their explanations, they sorted these ideas into an evidence-based characterization of trait expression mechanism in rice plants.

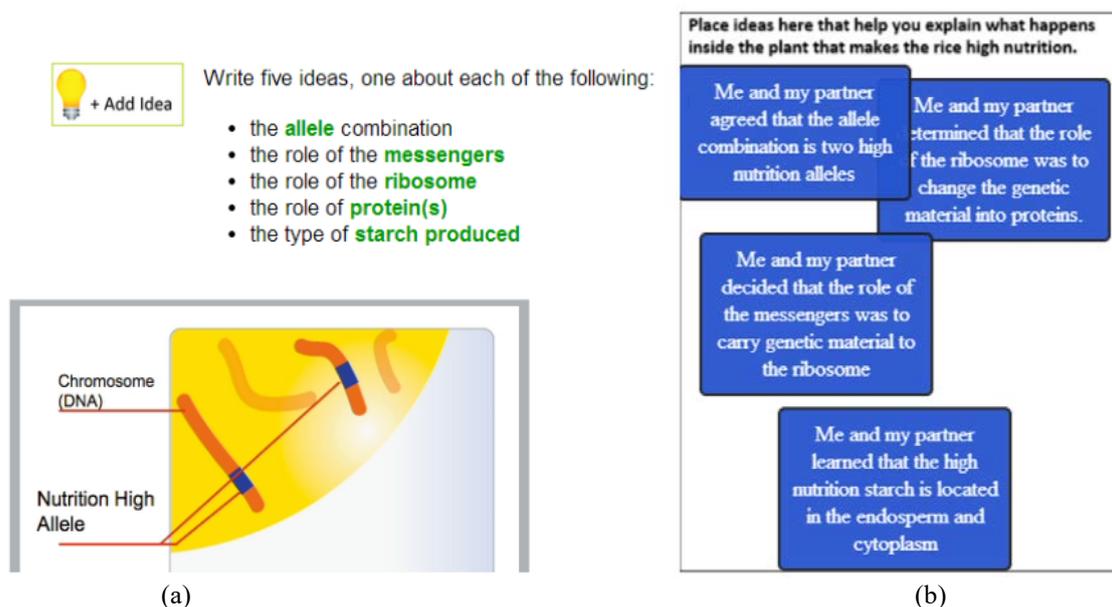


Figure 1. Screenshots of (a) model observation step with specific prompts and (b) examples of students' ideas placed in the organizing space.

Methodology

Research setting and participants

Two 7th grade science teachers who taught 50 and 49 teams of 2-3 students at two middle schools within the same district in the southwestern USA used the STEMGenetics unit in their classrooms during the 2013-14 school year. The district is comprised of 64% African American, 28% Hispanic, and 8% Caucasian students; sixty-one percent of the district's students receive free or reduced price lunch.

Data sources and scoring

The data of this study come from a set of embedded assessments in the unit: (a) the idea basket in the two activities, (b) the organizing space at the end of the two activities, and (c) students' responses to the prompt of the Explanation Builder at the end of activity four, "Explain what happens inside a rice plant that makes it high nutrition."

Idea basket and organizing space

To investigate how the Idea Manager tool scaffolded students' use of a model to construct coherent scientific explanations of the trait expression mechanism, we designed a conceptual inventory (CI) rubric to score ideas in the idea basket and organizing space (Table 2). The CI rubric identified ideas for five non-observable cellular structures that comprise the trait expression mechanism: genetic information, messengers, ribosomes, proteins, and starch. For each of these mechanistic elements, a set of possible normative ideas were listed to guide scoring of each idea in the idea basket and organizing space. If an idea was present and accurate about a mechanistic element, it received a score of 1. If absent or non-normative, it received a score of 0. Two researchers (the first and second authors) established inter-rater reliability of 95%, then coded separate sets of student responses. We then computed composite scores for the idea basket and organizing space for each student group by adding their scores on all ideas together (e.g., *normative model score*: number of normative ideas added during a model observation step).

Scientific explanation

We constructed and applied a 5-score KI rubric to students' explanations of the trait expression mechanism at the end of the activity. The rubric was iteratively revised based on several rounds of team feedback and review of samples of student work with consideration of boundaries between scores and distribution of scores (Table 3).

Two researchers established inter-rater reliability of 90%, then coded separate sets of student group responses. Each team's explanation was assigned a KI score of 1-5.

Table 2: Sample CI rubric for scoring of students' ideas about trait expression mechanism in rice plants

Mechanistic Elements	Possible Normative Idea	Student Examples
(G) Genetic Information	G3. The allele combination is two high nutrition alleles	<i>The allele combination was high nutrition. -- WISE ID 137243</i>
(M) Messenger	M2. The messenger is/has/carries genetic information or instructions from the allele/chromosome	<i>The messengers are genes that are sent out to the ribosomes to make proteins. -- WISE ID 136931</i>
(R) Ribosome	R2. The ribosomes use instructions from the messenger or genetic information to create a protein	<i>The ribosomes make proteins and use the genetic info from the messenger to make up the body in a certain way. -- WISE ID 13589</i>
(P) Protein	P2. The glucose/reactants interacts with or passes through proteins	<i>The role of the proteins is to come together and transform glucose into starch. -- WISE ID 137040</i>
(S) Starch	S1. The rice plant has/produces high-nutrition starch	<i>The starch inside of the Endosperm cell inside has high nutrition. -- WISE ID 136929</i>

Table 3: KI rubric for explanations of the trait expression mechanism in rice plants

	KI Score	Characteristics of Model-based Explanations	Student Examples
Mechanistic Explainers	5 Complex Integration	DESCRIBE KEY INTERACTIONS of NON-OBSERVABLE cellular structures and functions that comprise the trait expression MECHANISM .	<i>The way the rice plant is high nutrition is when the high nutrition allele sends messengers to the ribosomes and the ribosomes makes protein and the protein makes starch and forms in a granule. -- WISE ID 136899</i>
	4 Full Integration	DESCRIBE the one (1) KEY INTERACTION of NON-OBSERVABLE cellular structures and functions that comprise the trait expression MECHANISM	<i>Inside the nucleus, the high nutrition alleles are located in the chromosome. From there, the messengers transport genetic material to the ribosomes. Then the ribosome turns the genetic material into proteins. -- WISE ID 137402</i>
Partial Mechanistic Explainers	3 Partial Integration	DESCRIBE the RELATIONSHIP between the FUNCTION of NON-OBSERVABLE cellular STRUCTURES of the trait expression mechanism --OR-- <u>Partially</u> DESCRIBE the KEY INTERACTIONS of NON-OBSERVABLE cellular structures and functions that comprise the trait expression MECHANISM	<i>The alleles tell the ribosomes to make protein and starch to make the high nutrition cells. -- WISE ID 136847</i> <i>The first thing that will happen is the messenger will help the ribosome produce proteins. Then the protein will produce starch. -- WISE ID 137263</i>
Non-Mechanistic Explainers	2 Isolated, Non-observable	IDENTIFY NON-OBSERVABLE cellular structures of the trait expression mechanism	<i>Inside of a rice plant the allele combination, proteins, messengers, and ribosomes all work together and produce High Nutrition. -- WISE ID 136825</i>
	1 Isolated, Observable	IDENTIFY OBSERVABLE characteristics or traits	<i>The farmer can use IKI to indicate whether or not it's high or low in nutrition. -- WISE ID 137265</i>
	0 Incorrect or Irrelevant	<u>Only</u> incorrectly IDENTIFY, DESCRIBE RELATIONSHIPS, or INTERACTIONS of model elements	<i>In a rice plant there is starch, and starch is a complex carbohydrate and it turns to sugar in the body, which acts as an energy boost. -- WISE ID 136930</i>

Analysis

The KI analysis provided evidence about the range of model-based explanations among students. Given the formative design of the Idea Manager, we compared composite scores for idea basket and organizing space to differentiate how students engaged in the interrelated STEM practices: use a model and construct an explanation to deepen their disciplinary content -- trait expression mechanism in rice plants.

Findings

The two teachers in this study implemented the Grade 7 STEMGenetics unit with fidelity. Students completed the designed learning activities; they added a mean of 18.71 ideas to their idea baskets during the two activities (SD=6.41). No statistically significant differences in total ideas added were detected between higher and lower performing students as measured by their KI score. However, we did notice a significant difference in the quality of the ideas as explained further below.

Knowledge integration of the trait expression mechanism

Across both classrooms, 99 student teams scored a mean of 2.21 for their explanation of trait expression (SD =1.52). We grouped student teams into three significantly different categories: non-mechanistic explainers (KI score 0-2), partial mechanistic explainers (KI score 3), and mechanistic explainers (KI score 4-5), [$t(53)=-14.354$, $p<.001$; $t(16)=-11.785$, $p<.001$].

The 54 student teams categorized as non-mechanistic explainers identified only (a) observable characteristics or features (e.g., rice plant is high nutrition) or (b) non-observable cellular structures that comprise the trait expression mechanism (e.g., alleles or ribosomes). The explanations of the 28 student teams in the partial mechanistic explainers category either (a) describe the structure-function relationship of at least one non-observable cellular structure of the trait expression mechanism (e.g., ribosomes create proteins) or (b) partially describe the key interactions of non-observable cellular structures and functions that comprise the trait expression mechanism (e.g., messenger helps ribosomes create proteins, which produce starch). In the mechanistic explainers category, 17 student teams fully describe at least one key interaction of non-observable cellular structures and functions that comprise the trait expression mechanism (e.g., messenger carries genetic material from the alleles to ribosomes; ribosomes create proteins, which produce starch).

Using a model to construct an explanation of trait expression

Overall, students added a mean of 3.01 (SD=2.27) normative ideas to their idea baskets when they observed a dynamic scientific model and were prompted to add ideas about specific elements of the trait expression mechanism in three steps of the two learning activities. Despite a small difference in the total number of basket ideas, non-mechanistic explainers differed significantly in the mean number of normative ideas (2.33, SD=2.04) added to their idea baskets compared to the partial mechanistic (3.61, SD=1.77, $t=-2.931$, $p=0.005$) and mechanistic (4.18, SD=2.96, $t=-2.393$, $p=0.026$) explainers (Figure 2). The addition of normative model ideas across all coherence groups suggests the modeling steps scaffold students to (a) attend to the structure, function and interactions of the non-observable cellular structures of the trait expression mechanism and (b) add normative ideas about these mechanistic elements to their idea baskets.

Student teams who constructed a more mechanistic scientific explanation (KI scores 3-5) about the trait expression mechanism added a mean of 1.28 to 1.85 more normative model ideas to their idea baskets than the non-mechanistic explainers. This suggests that non-mechanistic explainers may have struggled more than other students to correctly identify the structure, function, and interactions of the non-observable cellular structures represented in the trait mechanism model.

During the Explanation Builder step, student teams selected relevant ideas from their idea baskets to place in the organizing space. Overall, student teams placed a mean of 1.78 (SD=1.78) normative model ideas in their organizing space. Again, the non-mechanistic explainers differed significantly in the number of normative model ideas (1.28, SD=1.70) placed in the organizing space compared to the partial mechanistic (2.32, SD=1.61, $t=-2.730$, $p=0.008$) and mechanistic (2.47, SD=1.87, $t=-2.339$, $p=0.028$) explainers (Figure 2). Students who constructed a more mechanistic scientific explanation about trait expression (KI scores 3-5) placed 1.04-1.09 more normative ideas in the organizing space. These students had more normative ideas about the mechanistic elements available in their idea baskets during the Explanation Builder step where they distinguished and sorted these relevant ideas into scientific explanations of the trait expression mechanism.

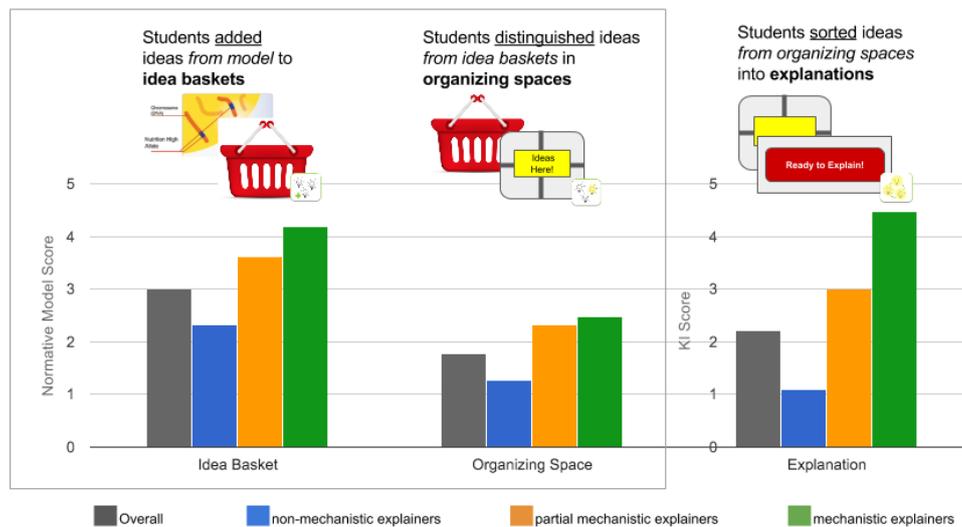


Figure 2. Differences among explainers when engaged in a set of interrelated STEM practices: use a model to construct a scientific explanation.

Despite the significant difference in the KI scores among partial mechanistic and mechanistic explainers, these students did not differ significantly in the number of normative model ideas added to their idea baskets or placed in their organizing space. This finding suggests that these students engaged similarly in the STEM practice, use a model, to identify the mechanistic elements in trait expression. Yet, mechanistic explainers may be better able to (a) understand the relationship between mechanistic elements represented in the trait expression model or (b) sort their distinguished ideas into a coherent mechanistic explanation of trait expression.

Conclusions and implications

Our coordination of the KI and ECD design frameworks afforded the opportunity to design instruction based on the Idea Manager tool for integrated STEM learning of content and practices. Embedded in the Grade 7 STEMGenetics unit, this formative learning tool engaged students in the observation of dynamic scientific models of the trait expression mechanism in rice plants. Students were prompted to document their observations by adding ideas to their idea baskets about the non-observable structures, functions, and interactions that comprise the trait expression mechanism. Later, the organizing space within the Idea Manager tool prompted students to distinguish and sort the ideas in their idea basket, in order to construct a written explanation for how the trait expression mechanism yields the high-nutrition trait in rice plants.

The findings from this study delineated three tiers of students who differed in their abilities to use a model and construct an explanation, and their integrated understanding of the disciplinary content -- trait expression (Figure 2). These documented differences in student learning revealed opportunities for (a) teachers and tools to differentiate scaffolds that engage students in STEM practices and (b) refinement of the Idea Manager's design features to generate more nuanced formative evidence for how students use a model to construct an explanation.

Non-mechanistic explainers demonstrated limited use of the trait expression models. They added an insufficient number of normative model ideas to their idea baskets, indicating a limited understanding of model elements. They articulated isolated ideas about observable traits or non-observable structures in their explanation of the trait expression mechanism, which was likely due to the limited availability of normative model ideas in their idea baskets. These findings suggest that non-mechanistic explainers need differentiated scaffolding to add normative model ideas to their idea baskets so that they are made visible to students when it is time to distinguish and sort relevant ideas into scientific explanations. Also, these students likely need scaffolding beyond the structure-specific prompts during model steps, such as computer-aided or teacher feedback about the contents of their idea baskets.

Both partial mechanistic and mechanistic explainers demonstrated requisite use of the trait expression models, indicating an understanding of model elements. These students differed in their ability to construct an explanation. The mechanistic explainers demonstrated a higher ability to construct an explanation as evidenced by articulation of key interactions of structure and functions that comprise a mechanism, indicating an integrated disciplinary content knowledge of trait expression. Alternately, partial mechanistic explainers articulated non-observable structure-function relationships or partial interactions of structures and functions that comprise the

trait expression mechanism. This is likely due to partial understanding of the key interactions in a mechanism. Differences in how students understand the relationship between mechanistic elements might become more evident when students construct an explanation that characterizes the trait expression mechanism. These findings suggest the need for (a) additional scaffolding in the organizing space to help students organize and sort their ideas into a coherent scientific explanation and (b) refinement of the Idea Manager tasks to elicit evidence of students' abilities to describe the relationships between model elements and the correspondence to the trait expression mechanism. For example, we could redesign the organizing space to use a sequential organizing principle and help students organize their normative model ideas into the trait expression mechanism.

The implications of this study extend beyond the disciplinary content in Grade 7 Genetics and the interrelated STEM practices, use a model and construct an explanation. Formative learning tools, like the Idea Manager and Explanation Builder, can provide intermediary artifacts for teachers and researchers to (a) identify which practices challenge students most and (b) analyze formative evidence to determine how to best to provide scaffolds and adjust instruction. In addition, the Idea Manger could be used to distinguish proficiencies between different sets of interrelated STEM practices, such as analyzing and interpreting data and engaging in argument from evidence. Further design and study of formative learning tools will advance more nuanced understanding of learning design that can support learning of STEM content and practices.

References

- Bell, P. (2004). The educational opportunities of contemporary controversies in science. In M. Linn, E. Davis & P. Bell (Eds.), *Internet environments for science education* (pp. 233-260). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bransford, J., A. Brown, & R. Cocking. (2000). *How People Learn" Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.
- Debarger, A., Tate, E. D., Beauvineau, Y., Feng, M., Shank, P., Heck, T., and Williams, M. (2014). A design inquiry: Bridging assessment and curriculum frameworks to engage students in science practices. In *International Perspectives in the Learning Sciences: Proceedings of the 11th International Conference of the Learning Sciences*. Boulder, Colorado: International Society of the Learning Sciences, Inc.
- Kali, Y., Linn, M. C., & Roseman, J. E. (2008). *Designing Coherent Science Education: Implications for Curriculum, Instruction, and Policy*. New York, NY: Teachers College Press.
- Erduran, S., & Dagher, Z. (2014). *Reconceptualizing the nature of science for science education: Scientific knowledge, practices and other family categories*. Dordrecht, the Netherlands: Springer.
- Ford, M. J. (2015). Educational implications of choosing "practice" to describe science in the Next Generation Science Standards. *Science Education*, 99(6), 1041-1048.
- Lee, H. - S., Linn, M. C., Varma, K., & Liu, O. L. (2010). How do technology-enhanced inquiry science units impact classroom learning? *Journal of Research in Science Teaching*, 47(1), 71-90.
- Linn, M. C. & Eylon, B-S. (2006). Science education: Integrating views of learning and instruction. In P.A. Alexander and P. H. Winne (Eds.), *Handbook of Educational Psychology*, (2nd Ed., pp. 511-544). Mahwah, NJ: Lawrence Erlbaum Associates.
- McElhaney, K., Matuk, C., Miller, D., & Linn, M. (2012). Using the idea manager to promote coherent understanding of inquiry investigations. In *International Perspectives in the Learning Sciences: The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences*. Sydney, NSW, AUSTRALIA: International Society of the Learning Sciences, Inc.
- Mislevy, R. J., & Haertel, G. D. (2006). Implications for evidence-centered design for educational testing. *Educational Measurement: Issues and Practices*, 25, 6-20.
- National Research Council (NRC). (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.
- NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: The National Academies Press.
- Slotta, J. D. & Linn, M. C. (2009). *WISE Science: Web-Based Inquiry in the Classroom*. New York: Teachers College Press.
- Stroupe, D. (2015). Describing "Science Practice" in Learning Settings. *Science Education*, 99(6), 1033-1040.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant Number DRL-1119055. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Situated Learning, Situated Knowledge: Situating Racialization, Colonialism, and Patriarchy Within Communities of Practice

Joe Curnow, University of Toronto, joe.curnow@mail.utoronto.ca

Abstract: This paper brings feminist theories of situated knowledge into theorizations of situated learning, arguing that racialization, colonialism, and patriarchy shape members' experiences of communities of practice. I argue that situated learning must accommodate standpoint epistemologies in order to theorize learning from the margins. Using an environmental activist group, I demonstrate how modes of full participation can be racialized, colonial, and gendered in ways that inhibit women and people of colour from being recognized as full participants. Video data from activist actions shows that dominant practices are rooted in standpoints and that the social locations of participants shape their learning and willingness to participate in racialized, colonial, and gendered modes of participation.

Introduction

As we focus our conversations around transforming learning at this International Conference of the Learning Sciences, we are presented with an opportunity to expand situated learning theory (Lave and Wenger, 1991) so that it is better able to account for transformations of learners from non-dominant racialized, gendered, and colonized standpoints. For decades, Black feminists (Collins, 1986), third world feminists (Mohanty, 1988; Sandoval, 1991), and standpoint feminists (Haraway, 1988; Harding, 1992; Hartsock, 1983) have argued that the social location from which we experience the world shapes what we know and how we know it. Within situated learning theory, though, there has been little attention to power dynamics (Contu and Willmott, 2003) and social relations of racialization, class, gender, colonialism, dis/ability, or others. Situated learning argues learning is a process of becoming, where people move from positions of novice peripherality into positions of mastery through their engagement with the skills, dominant frames, and social processes of the community of practice (Lave and Wenger, 1991). This view privileges the ongoing construction of community and how the immediate context shapes how participants perform skills, talk about their work, and work together.

However, within situated learning theory and the subsequent development of communities of practice theory (Wenger, 1998), the focus on local context leads to an erasure of the trans-local context. Communities of practice appear to exist outside broader social relations, free to enact whatever norms they want, unbound (or bound) by patriarchy, colonialism, and racism depending on the practices of the local community. In situated learning, it is assumed that communities of practice are not bound by social location or the standpoints of members in positions of mastery. This limits the theory's ability to analyze learning in diverse groups, where members of a community may not share worldviews or modes of engagement based on their social locations. Since this body of theory underpins much of sociocultural theories of learning, we must address this gap.

In the sections that follow, I revisit situated learning and critique the social constructivist elements of communities of practice that suggests communities create their own modes of practice and perform them outside of or ambivalent to dominant social relations. Drawing on feminist standpoint theorists, I argue that we cannot extract communities of practice from their historical context, and suggest an anti-racist, feminist, dialectical materialist approach to situated learning that centres, rather than erases, the standpoint of people producing and circulating knowledge in a community of practice. I then use a case study of learning in the environmental movement to illustrate how standpoint matters in the ways that practices are privileged, how people participate, and how mastery is achieved and not achieved in a community of practice. This data demonstrates that moving into full participation is not available to everyone equally, and people's social locations – constituted beyond the local community of practice as well as within it – dramatically impact their ability to become full participants.

Situated learning and communities of practice

Lave and Wenger (1991)'s theories of situated learning and legitimate peripheral participation explain learning as becoming, where through participation, members learn the skills and social performances of the community. Communities of practice are understood as groups of people engaged in ongoing collective work in a joint enterprise using shared repertoires (Wenger 1998). The theory articulates the learning process as legitimate peripheral participation, where new members move from the periphery into full participation gradually, through deeper immersion in the community. Rather than envisioning learning as mimicry or acquisition, situated learning claims people learn through absorption and integration into the community, where all members participate in generative, ongoing negotiation of the community itself, though in differentiated ways.

Situated learning and communities of practice theories have been critiqued for their inattention to power dynamics (Contu and Willmott, 2003) and their consensual framing of communities (Hughes, Jewson, and Unwin, 2011). Most of these critiques have centred on the nature of the employee relationship in workplace settings, where worker exploitation by capital is a core function of the relationship, defining and limiting the learning relationship to one of reproduction (Contu and Willmott, 2003; Hughes, Jewson, and Unwin, 2011). Far less attention has been paid to social relations of racialization, colonialism, and gender. While gendered participation has been addressed in a small portion of the literature (Hodges, 1998; Paechter, 2003; Salminen-Karlsson, 2006), nothing has extensively addressed the ways racialization and colonialism shape experiences of legitimate peripheral participation in a community of practice. Notably, the pieces that have engaged women's experiences in a community of practice centre White women's experiences or undifferentiated women's experiences, a move that universalizes and flattens women's experiences into a unified whole that feminists of colour have argued against. The articles on women in communities of practice show not all practices are attainable (Hodges, 1998; Salminen-Karlsson, 2006). They identify the naturalized standpoint of mastery in the community of practice and name practices as masculine and in service to the ways men navigate the world.

In the next section, I introduce standpoint theory as a way of addressing situated knowledges within situated learning. Understanding and accounting for the different ways that people of colour and women know and participate enables us to theorize learning dynamics more effectively. Rather than assume all members will believe and adopt the same narratives that circulate in a community of practice or be able to enact forms of participation in the same ways, we need a situated learning that situates the knowledge of the members of a community of practice within their historical social relations.

Situated knowledge and standpoint epistemologies

Situated learning focuses attention on the immediate community and centres how particular groups produce and sustain practices and philosophies of the local community of practice. Situated knowledge and standpoint epistemologies, in contrast, locate people's ways of knowing within their social location. They argue, for example, that Black women understand and experience the world differently than White women and Black men based on how relations of patriarchy and racialization shape their lived experiences. This view of knowledge, while rooted in the immediate community experiences of gender and racialization, shift our attention to trans-local relations of power (Smith, 1987). So while it may seem that these two bodies of theory are speaking to each other, they actually speak past each other, with situated learning deeply attentive to the specifics of the local while situated knowledge attends to the ways the trans-local is enacted in the local.

Situated knowledge and standpoint theories have been sites of extensive feminist scholarship. Without engaging deeply in the nuances of the field, it is important to identify some of the major contributions situated knowledge theories offer situated learning. The first contribution is the basic premise that social location shapes what we know and how we know it (Collins, 1986; Harding, 1992; Haraway, 1988). Early feminist theorizations argued against the so-called objectivity and neutrality of dominant thought in sciences and sociology, and demonstrated the ways that the types of knowledge that those fields produced were rooted in the standpoint of White men. As a result, the questions asked and their answers interpreted the world through the experiences of men, ignoring questions that feminist scholars prioritized. Standpoint theorists argued that feminist research should be built from the everyday/night experiences of women (Smith, 1987), and include attention to reproductive and affective labour (Hartsock, 1982). Other feminist standpoint theorists argued that feminist viewpoints were superior because those standpoints saw the world with fewer distortions than men in their relative positions of power, suggesting that subjugated knowledges (Sandoval, 1991) were preferable because they allowed a more comprehensive view of social relations. This approach stresses that non-dominant standpoints not only perceive the world differently, but also that those standpoints open up more expansive understandings of social relations.

Patricia Hill Collins' work (1986) is foundational to this line of argumentation, and her contributions of Black feminist thought bridge situated learning in productive ways. Collins argues that people from outside White male dominance can learn to act in masculine and White ways, but that those people can also see and question dominant paradigms. She claims that people outside dominant social relations have perspectives and insights to offer, based on where they are situated within relations of racialization, gender, and other relations of oppression. She states that people of colour may choose to remain "outsiders within", opting out of certain communities and practices, and that learning to navigate dominant social relations is itself an accomplishment.

Black feminist thought and feminist standpoint theories are necessary components of situated learning. Without acknowledging that all forms of participation are rooted in certain peoples' standpoints, we cannot theorize learning, and we certainly cannot accommodate oppositional consciousness that seeks to challenge dominant ways of knowing and being. Acknowledging how communities of practice are situated within social

relations and enact them at the same time allows us to interrogate how dominant social relations can become normalized in a community of practice and how that reflects the standpoints of participants in positions of dominance. This articulates a more nuanced critique of learning that moves beyond mere reproduction, allows for non-dominant modes of participation, and centres critical praxis within communities of practice.

Situated knowledge and learning in and beyond communities of practice

Lave and Wenger's discussion of full participation in a community of practice fails to accommodate social relations of race, colonialism, and gender in several ways. The first is their suggestion that communities establish their own practices and guidelines. While this is true in the sense that members jointly enact and produce their co-negotiated modes of participation, it separates the practices and the communities from the larger context. Communities of practice are situated within historical relations of racism, patriarchy, and colonialism, as much as we may try to resist. Suggesting that communities of practice are able to create their own practices outside of those dominant social relations de-historicizes them, making it seem as though the racism people face in a particular community of practice is rooted in the racism of that community alone, rather than being rooted in socio-historical and trans-local relations of colonialism and racialization. While the particularities of racialization are navigated in specific ways within a community, they do not act in isolation.

Another way that Lave and Wenger's descriptions of communities of practice fail to integrate racialization, colonialism, and patriarchy into their analysis is through the neutralization of dominant forms of participation by members in positions of mastery. Lave and Wenger propose a value-neutral approach to learning that suggests that whatever the norms of the community are is what will need to be learned within that community in order to become a full participant. This erases socio-historical difference between members in communities of practice and the ways social relations shape and limit members' ability to act in certain ways. This approach effectively removes the standpoints embedded in ways of knowing and being in a community, making it seem as though the full practices are not tied to certain standpoints, that all forms of practice are universally available and unbound by members' social locations. Lave and Wenger's approach also encounters problems in their articulation of "community", which has been critiqued extensively (Hughes, Jewson, and Unwin, 2011). It presupposes that communities are homogenous, free from power differentials, and open to everyone. When combined with the neutralization of standpoint, which makes it seem as though performing Whiteness, for example, is equally available to members of colour within a community, and the de-contextualization and de-historicization of social relations, this view of community erases the experiences of marginalization that many members of a community of practice experience based on their social location. This may have little to do with their learning, abilities to perform skills, or understanding of a community, but rather may have everything to do with their social location. It neglects the subjugated knowledges that non-dominant members bring communities of practice and erases different epistemologies, ontologies, and cosmologies of members from outside dominant social relations. In this way the notion of "community" masks a process of colonization, patriarchy, and racialization that erases forms of knowing and being.

This tension between situated knowledge and the theorization of full participation within communities of practice is an epistemological and ontological problem for situated learning. Theorizations of situated learning do not take seriously that how we know and what we know are rooted in how we are situated within social relations. This is an ontological problem in that it does not address the dialectical relationship between the local experiences of a community of practice with the trans-local relations. While racism, sexism, and colonialism may manifest in nuanced ways within a community of practice, those relations extend beyond the boundaries of the community, while being enacted and reproduced in the local space.

In the sections that follow I demonstrate how practices in an activist community of practice illustrate the ways modes of participations are rooted in particular ways of knowing and being – in this case, a White masculine mode of participation in a multi-racial, multi-gender environmental campaign. I trace how White male participants easily enacted the modes of participation and were rapidly recognized as masters while women and people of colour were kept from full participation despite their learning and mastery of skills within the community. I demonstrate how the modes of participation that were valued in the group in order to achieve full participation were rooted in White, masculine ways of being and knowing, while other modes of participation based in Black, Indigenous, and/or women's social locations were not valued and prevented members from marginalized social locations' participation from being recognized as full members.

Methods

Self location

Indigenous and feminist researchers (Absolon and Willett, 2005; Collins, 1986; Wilson, 2008) emphasize locating the self to clarify relational accountability (Wilson, 2008) and allow readers and research participants to fully understand why we engage particular questions. They challenge researchers to explicate how research fits into a broader theory of change that benefits communities we are accountable to. As a White settler, my approach to this research stems from White anti-racist allyship work while being accountable to communities of colour. As a cis woman, I enter into questions of marginalization and silencing in group settings through my experiences, in schooling and activist spaces, of having my voice silenced, ignored, and talked over. I feel a responsibility to the young women I work with to carve out alternative spaces and to help them to recognize the ways that their contributions are not equally valued based on their gender. I recognize that as a White cis woman, my racial and cis privilege shape my experiences as a woman and that the way I position myself and am positioned by others in groups is very much shaped by the multiple locations of privilege.

Data collection

This paper is based on an ongoing participatory action research project that examines how student activists learn about race, colonialism, and patriarchy through their involvement in environmental activist campaigns. Video was collected with the University of Toronto (UofT) fossil fuel divestment campaign. The group of young people met weekly. Meetings had from 11-42 participants, representing all years in school, (1,2,3,4,5, Law, MA, PhD). Men and women students attended in roughly even numbers and there were no openly trans or other gendered students. Racial and ethnic make-up shifted over the course of the campaign, but the group remained predominantly White, even as Indigenous, Black, Latino, South Asian, and East Asian students became increasingly involved, in terms of numbers and leadership in the group. Video was collected at 29 meetings, 6 actions and 2 focus groups over the course of the campaign, resulting in over 6,500 minutes of video. Videos last from 60-240 minutes, depending on the meeting length. Interviews and stimulated recall interviews (SRIs) were conducted with 8 focal participants. Interviews were semi-structured, lasting from 30-90 minutes. SRIs were held 1-2 days after a meeting and participants reviewed video, discussing group dynamics and their own thoughts during the meeting. SRIs lasted from 40-120 minutes and were held intermittently over the year, based on participants' interest and availability. Videos were captured from one to four angles and stacked so all streams are visible and watched and coded simultaneously.

Analysis

After video was collected, it was content-logged and pre-coded using preliminary codes based on the research question (including race, gender, and colonialism). The first substantial analytic pass of coding was conducted by five women participants from the group, one White, one South Asian, two East Asian, one Indigenous, and one Black man. We watched segments of videos from across the year together and coded "interesting" segments, asking the broad question of how race matters in our group, discussing every instance someone raised and making extensive notes. After conducting "interesting" coding on three segments from the beginning, middle, and end of the year, we reviewed consolidated them into codes that were most present in the video we reviewed and in our experiences of the group. Codes included default space, including instances where White settler participants assumed the room was White or settler by using "we" and universalizing their experience or by mobilizing White settler ways of acting and knowing in the group. Codes also included taking up space, tokenizing, dismissing, and labeling strategies as distractions. We assessed who was in a position of mastery by analyzing who made decisions for the group, who controlled what the whole group prioritized including types of work and mode of participation, who other members included in decisions or deferred to, and who was publicly recognized. Based on the coding, relevant segments of video were transcribed. A draft of this analysis was shared with the participants who co-developed the codes and their feedback was integrated.

Findings

The full practices of the divestment campaign at UofT included three main areas. The first was participating in meetings, where decisions and strategy were decided. The second set of practices was interacting with the official institutional actors of the university, including the office of the President, the Governing Council, and the President's Ad Hoc Committee. The third set of practices was publicly representing the campaign, most often by presenting to a group or presenting on a panel. Since the members who positioned themselves as masters were overwhelmingly White men, they effectively dictated what the important tasks were and how they

ought to be completed. We can trace how women and people of colour's performances were not recognized as legitimate by men in positions of power, which shows how the practices themselves were rooted in the standpoint of White men and were thus difficult for people from other social locations to master.

Participating in meetings

The divestment campaign met once a week for one and a half to two hours. Meetings were open to all members and were facilitated by rotating chairs. One of the main ways people established themselves as core members of the community was by participating vocally in the meetings. This ensured one's voice would be present and one could be involved in the strategy setting and decision making of the group. White men – particularly those in positions of full participation, though not only them – participated in meetings at very high levels. Men spoke far more often and for longer duration, and in doing so, they positioned themselves as experts.

Graham provides a good example of how participating in the space enabled him to move into a position of full participation very quickly. Graham had never been involved in activism before, but joined the group during his second month of law school at UofT. By the first all-group outreach meeting, he had already become a campaign lead and presented to new potential members, speaking assertively about the campaign and the importance of the “inside game strategy” where we focused on working through institutional protocols in order to win divestment. Graham's participation was marked by frequent speaking, responding authoritatively to questions, and weighing in often on the decisions of the group, typically masculine modes of engagement. Other members of the group often deferred to him, named him explicitly as an “expert”, and cited his experience as important and helpful, despite the fact that he had less experience and formal expertise than other members.

In contrast, we can trace Ariel and Amil's participation in the group. Though they had been involved in the campaign and environmental organizing on campus longer and in more comprehensive ways, their modes of participation, rooted in their racialized and gendered performances, kept them from being recognized as full participants. Ariel, a White woman in fourth year, had been involved in the group longer than all but two participants. She was a co-president of the group, yet her authority was continually questioned, with her male colleagues regularly talking over her, making decisions without her, and attributing her contributions to other men. One example of the ways she had participated in the key practices of the community was her interaction in group meetings. This was one of the most important and visible ways that group members demonstrated their authority and positioned themselves as experts. In meetings, women spoke far less frequently and for shorter durations each turn, and were engaged in exclusive talk between a small group of insiders half as often. When women were engaged in exclusive talk, Ariel was almost always the woman included. She often inserted herself into conversations and participated vocally in a very different way from most of the women and people of colour in the room. In these exclusive talk turns, though, she was almost always outnumbered 2:1 by men and her participation was often made up of continuer statements, to acknowledge other people's contributions and encourage them to keep going, rather than her own substantive contributions. This was true in her speaking over all; her mode of participation tended to be highly gendered. She asked questions rather than speaking in declaratives (as the men in positions of authority did), she hedged her statements with qualifiers and apologies, and she often ended her statements at a higher pitch, even when she was not asking a question. Her feminine performance of the practices often meant that the men in positions of mastery did not treat her as an authority, and engaged other men in decisions without her. At the end of the year in a focus group, multiple women said they had not known she was a co-president. Her mode of participation was considered peripheral, and she was kept out of many of the important high stakes activities, having to fight for her inclusion in many instances.

While Ariel's participation was kept peripheral because she performed the core practice but did not do it in a masculine way, Amil's experience showed how his social location precluded doing them. As a fourth year Black man and immigrant to Canada, he did not perform certain practices because he knew his performance of would be judged as inadequate for his Black racialized mode of engagement. Though he had been involved in environmental organizing for three years on campus, he almost never spoke and was asked for his opinion only once during the first year. His non-participation in group meetings meant that most of the men in positions of mastery cast him as peripheral, despite his engagement in the tasks of the community of practice. Amil described his decisions to not speak in a stimulated recall interview saying:

I'm generally more reserved in meetings... being in White spaces... in general... have that effect, personally, on me. It's something I've been dealing with since I've been at this university, in terms of putting out your opinions and having them scrutinized, because in instances it's been very vicious.... [White people say] “you see things this way because it affects you” and it's like you can't make an objective argument because you are somehow self-

interested in this... and it just shuts me down right away [...] every time I say something in a White space, in my mind it's like: how are these people perceiving what I say?

Amil made explicit the racial norms of the community that he had to navigate, including the Whiteness of the space. He talked about his pervasive sense of otherness and his “hyper awareness” of how White people perceived what he said (or did not say) and acknowledged that he did not think and speak in the same way those in positions of authority within the group did. Rather than push back against them and experience the blowback he anticipated, he stayed silent in meetings and talked to people one on one or in small groups outside the meetings. His strategy made sense for his social location, but had the impact of keeping him from being recognized by the men in positions of mastery in the group, despite his skills, his experience, and his expansive understanding of climate change and impacts in frontline communities.

Both Ariel and Amil were not treated as masters in the community of practice, by other members or by the men in positions of mastery, because the ways they participated were shaped by their experiences of patriarchy and racism, both in this specific community of practice and beyond it. As a result, they did not get acknowledged for their contributions to the community, while a brand new White man did quickly and easily and was able to drive the strategy decisions for the group.

Notably, women of colour almost never spoke in the group, regardless of how long they had been involved or how much work they put in behind the scenes. Their opinions were rarely sought and decisions were regularly made without any participation from women of colour, even on discussions of race, gender, and intersectionality. Their non-participation seemed to be considered a legitimate form of participation, which raises concerns about whose standpoints are required for decision-making, whose are not, and how that impacts framing, strategies, and approaches to political action.

Framing, strategy and theory of change

Strategy is the next area where we can see a core practice being demonstrated. At UofT, strategy was often negotiated at the weekly meetings, where the speaking dynamics noted above played a significant role in establishing whose ideas were included in short and medium term strategies. Group strategy was a process of frame alignment, where in order to be brought into the full forms of participation, one had to buy into the dominant logic of divestment. The people who started the campaign initially had charted an institutional course that consisted of going through the UofT's official process for divestment, including preparing a brief, gathering signatures, and supporting the brief in consultation with the President's Ad Hoc Committee for evaluating the petition. This required alignment with a theory of change that centred logical argumentation and respecting the institutional process, rather than protest, confrontation, or more direct resistance.

Graham demonstrated his alignment with the philosophies of the group right away, and he took up leadership of the inside game strategy within his first month of involvement. Here he advocated for meeting with the President's office staff and working through official channels. He, and the other White men from the inside game group, stressed “not rocking the boat”, and in a strategy meeting he said, “We don't want to give the Governing Council any reason to dismiss us... they are on Governing Council because the game works for them, because the system works for those people. They're all wealthy and a lot of WASPs in there, and that sort of thing. So we're left with, how do we play that game?” Again and again, the White men in positions of mastery brought the group back to a theory of change based on compliance with institutional protocol and rooted in non-confrontational action, including providing logical arguments to the administration when they engaged us.

In contrast, Amil balked at this strategy, as did other members of colour, arguing that the institutional pathways do not work for everyone, and rarely work for people of colour, poor people, and other marginalized groups. In one moment of contestation in a meeting this exchange occurred:

- Sam: If we can get 200 people sit in the Governing Council room at the meeting wearing, like, suits with divestment X's pinned to them and divestment signs – I think that'd be powerful.
- Amil: Do we have to wear suits?
- Sam: Uh, generally Governing Council meetings are like – you dress up.
- Amil: Is there like a dress code?
- Graham: No, it's not like ‘men must wear blazers’ –
- Sam: It would be frowned upon if you just showed up in a t-shirt and jeans.

- Amil: But if a bunch of students go in there... let's not play that weird respectability politics –
- Sam: I think we *should* play –
- Amil: – They're not going to kick us out –
- Graham: No, they won't.
- Sam: Look, they won't kick us out, but they *will* respect us more if we play into their respectability politics.

The exchange made explicit some of the philosophical underpinnings of the group, where women and folks of colour pushed back against the logic that the system would work for us and that we should try to conform to the dominant relations that govern it. Amil's disagreement made plain the different ways that different participants navigate the world and how it underpins the strategic choices of the group. Even though Amil was active in the inside game breakout group, he never fully bought into the idea that the governing council would bend to persuasive arguments. His life experience suggested that the system would not always work for him. Because of his view on this, rooted in standpoint, he was kept at the margins of the inside game group. He was not invited to speak to the President's staff or the Ad Hoc Committee (only White people from the group interacted with them). However his analysis of the situation does not suggest he is not competent, has not mastered the skills, or has not learned—in fact his ability to navigate the system differently demonstrates his learning and ongoing negotiation of White supremacist colonialism. Keara, a Cree woman in first year, also identified the inside game philosophy as central to the group, and said that she could never support it, because she knew that institutions like UofT did not work in Indigenous students' interest. Students of colour within the group believed that Governing Council would not interact with them as they did with the dominant White men, and did not always trust the men from our group to broker in our collective interest, but worried that the Whiteness of the approach would serve to tamp down the racial justice components of the campaign that they advocated for.

Public presentations

Our group also prioritized outreach to groups through public presentations. In order to be a full member of the community, one had to be able to coordinate and deliver presentations to groups on and off campus in order to persuade them to endorse the brief. Being able to perform this public task was an important marker of full participation. It demonstrated one's grasp of the issues, framing, and strategy, as well as one's ability to speak persuasively and build coalitions with other power brokers on campus.

Joanna serves as an example of this process. As a 3rd year White woman, she had been involved in other environmental work on campus, but joined the divestment campaign halfway through the year. She talked about herself as a “bold woman” and about her ability to “lean in”, joining in the activities that only the dominant men tended to participate in as a way of proving herself. On her second meeting, she volunteered to chair the meeting and she continually volunteered to take on public speaking tasks that tended to include only the people in positions of full participation. In a women's focus group she noted that it was not enough, saying:

Our breakout group has consistently been, like, going to groups to seek out endorsements. I feel like – pretty consistently – um, we've had a buddy system of two people going each time. And every time there has been at least one male – like, intentionally – it feels like intentionally. And yeah, I mean, I went to the medical society, originally I think I was supposed to go with Sam, and then Sam couldn't make it, so then they were like, oh Graham, he'll go with you. It's good because they are knowledgeable... but women are knowledgeable too. I don't think there's a single group that we presented to this entire year that there hasn't been one of the like, three or four main White males that went with someone else – whether that was a man or a woman. I wasn't encouraged to go alone, I wasn't even told it was an option would be for me to go and present alone or to go with another woman.

She recognized the gendered control of recognition of full participation and that she was kept from it. Joanna, more than any of the women or people of colour in the group, played by the rules of the community or practice. She presented publicly, spoke often and assertively in meetings, but still was not recognized.

Conclusions

In this community of practice, race, colonialism, and patriarchy shaped whose/which modes of participation were deemed legitimate and whose/which were not. Dominant social relation played out through the full

practices of the community and entrenched White men in positions of mastery, while people of colour and women's modes of participation remained peripheral. White men in the group experienced the modes of participation as neutral and familiar, while other members worked at a disadvantage in adopting the forms of participation deemed appropriate or the logic that underpinned them. Regardless of engagement and experience, people of colour and women did not become recognized as full participants in the community of practice.

As learning scientists, we cannot afford to ignore the ways race, colonialism, and gender shape communities of practice through their histories and continuities. When people of colour and women are kept at the periphery – either through so-called “failure” to perform Whiteness and masculinity or through acts of resistance, it is not adequate to suggest they cannot/have not reached full participation, or that their position of peripherality is legitimate. None of these explanations capture what is happening as people in dominant positions normalize their everyday ways of interaction and have their reproduction of dominant social relations conflated with learning, mastery, and full participation. They do not account for how people at the margins deftly navigate the community to achieve small concessions, to gain recognition, and to withdraw, resist, or embrace their outsider status. Through the terminology of the theory we would articulate their position as legitimate peripherality, but there is nothing legitimate about being marginalized based on one's race and gender. When we ignore the standpoints of people in positions of mastery and the full practices they seek to reproduce, we suggest to members from other social locations that through their participation and engagement they might achieve full participation. But when racism, colonialism, and patriarchy make it impossible or disproportionately difficult, this may not be the case. We cannot assume that everyone has the same potential to act and understand in the same ways. In doing so, we pathologize oppressed peoples and their inability to become full participants and legitimize the racist and sexist modes of participation that keep women and people of colour on the periphery of communities. Standpoint epistemologies represent one intervention that begins to bring attention to social relations, history, and difference. Acknowledging standpoint is a necessary step toward a more comprehensive transformation of situated learning to account for race, colonialism, and patriarchy in and beyond communities of practice.

References

- Absolon, K., & Willett, C. (2005). Putting ourselves forward: Location in Aboriginal research. *Research as resistance: Critical, Indigenous, and anti-oppressive approaches*, 97-126.
- Collins, P. H. (1986). Learning from the outsider within: The sociological significance of Black feminist thought. *Social problems*, 33(6), S14-S32.
- Contu, A. & H. Willmott. (2003). Re-embedding Situatedness: The Importance of Power Relations in Learning Theory. *Organization Science*, 14(3), 283-296.
- Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist studies*, 575-599.
- Harding, S. (1992). Rethinking standpoint epistemology: What is "strong objectivity"? *Centennial Review*.
- Hartsock, N. C. (1983). *The feminist standpoint: Developing the ground for a specifically feminist historical materialism*. 283-310. Springer.
- Hodges, D.C. (1998). Participation as Dis-Identification with/in a Community of Practice. *Mind, Culture, and Activity*, 5(4), 272-290.
- Hughes, J., Jewson, N., & Unwin, L. (Eds.). (2013). *Communities of practice: Critical perspectives*. Routledge.
- Lave, J. & E. Wenger. (1991). *Situated Learning: Legitimate Peripheral Participation*. NY, NY: Cambridge.
- Mohanty, C.T. (1988). Under Western Eyes: Feminist scholarship and colonial discourses. *Feminist review*.
- Paechter, C. (2003). Learning masculinities and femininities: Power/knowledge and legitimate peripheral participation. *Women's Studies International Forum*, 26, 541– 552.
- Salminen-Karlsson, M. (2006). Situating gender in situated learning, *Scandinavian Journal of Management*, 22.
- Sandoval, C. (1991). US third world feminism: The theory and method of oppositional consciousness in the postmodern world. *Genders*, (10), 1-24.
- Smith, D. E. (1987). *The everyday world as problematic: A feminist sociology*. Toronto: U of T Press.
- Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge: Cambridge Press.
- Wilson, S. (2008). *Research is Ceremony: Indigenous Research Methods*. Winnipeg: Fernwood.

Acknowledgments

Many thanks to Jody Chan, Amil Davis Tresanne Fernandes, Keara Lightning, and Jade Wong for their participation in developing this analysis. Thanks to Thelma Akyea, Andrew Kohan, and Genevieve Ritchie for their conversations throughout the process. This work was funded by the Vanier Canada Graduate Scholarship.

Opportunities to Learn Through Design: Mapping Design Experiences to Teacher Learning

Emily Horton, Jahneille Cunningham, Louis M. Gomez, and Kimberley Gomez,
eshorton@gmail.com, jah.cunningham5@gmail.com, louismgomez@gmail.com, kpg1321@gmail.com
University of California, Los Angeles, CA

Katherine Rodela, Washington State University Vancouver, katherine.c.rodela@gmail.com

Abstract: This paper examines the ways in which specific design experiences lead to certain types of learning. We engaged thirteen community college instructors in iterative design through Plan-Do-Study-Act (PDSA) cycles to design and develop new developmental mathematics lessons. In this study, as in most design-partnerships, teachers touch only some aspects of the design process. In the work reported here, we make explicit the varying ways in which instructors took part in PDSA cycles and examine the types of knowledge generated by their participation. We use Edelson (2002) as a lens to categorize learning into three types: domain, design framework, and design methodology. Results indicate that instructors do not need to be involved in every aspect of design to learn. Our findings highlight the role time plays on engagement and learning in design. Implications for design efforts, to the extent they are focused on learning, are discussed.

Introduction

Increasingly, education practitioners, policy makers, and researchers recognize that teaching quality is key to student achievement (Darling-Hammond & Richardson, 2009). The link between high-quality professional development and student outcomes is sound (Borko, 2004; Desimone, 2009); however, only recently have researchers become concerned with deciphering the underlying, and often messy, links between professional development conditions, what and how teachers learn, and transformation of classroom practices (Borko, 2004; Clarke & Hollingsworth, 2002). This study seeks to contribute to this growing body of literature, probing how teacher engagement in design can foster meaningful learning and promote changes in classroom practice.

A number of studies point to design as a site for teacher learning (Gomez et al, 2015; Koehler & Mishra, 2005; Voogt, 2015). Design is the systematic development of an educational innovation (e.g: curricular material, technology) to support some aspect of student learning (Edelson, 2002; Joseph, 2004). Edelson (2002) conjectures that if one participates in design processes, s/he will have many opportunities to learn. Most commonly, however, teachers, less often the engines driving the design effort (Penuel, Fishman, Yamaguchi & Gallagher, 2007), participate in only some elements of design. Thus, to better understand what teachers can learn as they engage in design, we must understand how specific design experiences lead to certain types of professional learning. Although the current literature is sparse, we posit that it would be fruitful to probe opportunities to learn in specific aspects of the design experience.

We report results from a two-year study, where we take up this question. We engaged 13 community college instructors in Plan-Do-Study-Act (PDSA) cycles (Langley et al., 2009), our design methodology guiding the iterative design of developmental mathematics lessons. We make explicit the varying ways in which instructors, who are critical to the design effort but do not drive the design process, took part in PDSA cycles, and examine the types of knowledge generated by their participation.

Community college developmental mathematics classrooms are important sites for learning by design. Efforts to increase the quality of teaching are gravely needed. Yet this problem receives relatively little attention (Boylan, 2002; Stigler, Givvin & Thompson, 2010). Each year, over thirteen million students enroll in community colleges across the U.S. For 59% of these students, the dream of graduation is quickly shattered when they are placed into developmental, or remedial, mathematics courses (Bailey, Jeong, & Cho, 2010). With a 30% success rate (Levin & Calcagno, 2007), developmental mathematics have been called the “graveyard of dreams and aspirations” (Merseth, 2011). Students are often doomed to retake courses, resulting in prolonged enrollment, increased debt, and in many cases, eventual dropout (Stigler, Givvin & Thompson, 2010). A core assumption of this paper is that professional development could figure largely in the reform of developmental mathematics. As such, we seek to examine ways in which developmental mathematics instructors learn through engagement in design as a means to improve instructional practices, and in turn, student outcomes.

Design as professional development

Studies that explore the potential of design for professional learning suggest that the elements of design key to learning include: situating learning in practice (Joseph, 2004); active inquiry into the problem (Koehler & Mishra, 2005; Kolodner et al., 2003), sustained and organized engagement (Collins, 1992; Koehler & Mishra, 2005), and collaboration (Voogt, 2015), as they align with key characteristics of effective professional development (Borko, 2004; Desimone, 2009; Little, 1990). Active and situated learning opportunities allow teachers to integrate new knowledge with existing knowledge (Davis & Krajcik, 2005), resulting in meaningful and authentic learning (Greeno, 1998). Using cyclical attempts to improve the intervention, designers learn most in their moments of failure. When one aspect of the design does not work, designers must reason through the ways different design elements work together, considering how change to one area of design may impact another (Collins, 1992). This decision-making process, requiring exploration of nuanced relationships between the tool, students, and local context (Koehler & Mishra, 2005; Krajcik et al., 1998), provides designers with important opportunities to learn (Edelson, 2002). Disciplined inquiry, coupled with the collaborative nature of design, provides a venue for instructors to encounter distributed knowledge as they take part in purposeful discourse with collaborators from diverse backgrounds. Such collaboration can be transformative as it allows designers to gain increased awareness of their own practices and beliefs (Koehler & Mishra, 2005).

Design is sustained over time and organized, another important characteristic that supports learning. The use of cyclical testing to develop optimal design solutions (Collins, 1992; Koehler & Mishra, 2005; Kolodner et al., 2003) naturally extends participation in design over time. There is a growing consensus that professional development that is sustained over longer periods of time presents increased opportunities for teachers to assimilate new knowledge into practice (Desimone, 2009). Studies have shown that when not given enough time in professional learning, teachers often do not retain what they learn and evidence little change in classroom practice (Coburn, 2004). Other scholars suggest that extended duration is not enough, and that time must be organized into structured activities, with purpose, to produce effective change (Garet, Porter, Desimone, Birman & Yoon, 2001). Adherence to a specific design methodology provides this organization, guiding those participating in design through design activities (Collins, 1992).

In this study, we ask, what is the relationship between kinds of design contact and specific opportunities to learn? The guiding hypothesis of this paper is that amount of time instructors spend in design may have consequences for the quality of their engagement in design activities and professional learning generated. We follow Edelson (2002) to explore three types of learning: 1) domain learning, which refers to increased knowledge about the design setting, such as increased understanding of the language and literacy needs of developmental mathematics students; 2) design framework learning, or an instructor's increased understanding of the design ideas involved in the design solution (in this case, mathematics, language and literacy tools, problem situation, and pedagogy in the new lesson); and 3) design methodology learning, which indicates an instructor's increased understanding of the design procedures, or in this study, PDSA cycles.

Methods

This study focuses on the collaborative design of 12 new developmental mathematics lessons. Our design goals were to contextualize lessons and reduce language barriers. Contextualization, or the integration of academic and occupational curricula, engages students in real-life, authentic problems resulting in more meaningful learning, making it easier to internalize, understand, transfer, and retain (Herod, 2002).

Participants. We examined data from 13 instructors from 6 community colleges across the U.S., who participated in lesson testing. Instructors volunteered to participate, and received a small honorarium for their work. Instructors engaged in design in varying ways (see Table 1); this allowed us to examine the ways in which specific design activities generate learning. We will detail this involvement in the next section.

PDSA Cycles. We used Plan Do Study and Act (PDSA) cycles to guide the developmental arc of the design work. An Improvement Science tool, PDSAs are characterized by quick iterative learning, fast failure, and rapid refinement (Langley et al., 2009). Instructors touch the PDSA cycles in the following ways. Our cycles began with a few instructors teaching the same lesson. "PLAN" occurs as instructors prepare to enact the lesson. "DO" occurs as instructors conduct a test by enacting the lessons. Within two days of enacting a lesson, instructors participated in either a follow-up semi-structured, open-ended interview (Seidman, 2006) or a survey aimed to gain insight into instructor experience teaching the lesson and recommendations for refinements. "STUDY" includes instructor's participation in these interviews and surveys. We then collected and analyzed data from lesson enactments, summarized results, and made quick revisions to the lesson before a new cycle of instructors tested the same lesson. "ACT" occurs as faculty, along with the design team, work out the plan for the next testing cycle. These cycles continued until all participants taught the lesson. Halfway through the PDSA cycles, we made intermediate refinements to the lesson based on more extensive data analysis. At the end of

testing, six of the thirteen participants attended a two-day design workshop to make long-term changes to the lessons. Participation in this meeting falls under “ACT”¹. The use of PDSA cycles allowed the design team to manage design revisions efficiently, as team members decided when and how design changes should be addressed. PDSAs recognize that straightforward design changes could be implemented immediately while more complex changes were put aside for later revisions. Data resulting from instructor engagement in PDSAs include: 51 instructor interviews, 52 instructor surveys, artifact design changes (documented changes within and across lessons), and ethnographic field notes of the 2-day instructor design meeting.

Table 1: Instructor Engagement in PDSAs

Frank: designed 4 lessons, enacted 9 lessons, 9 interviews, 9 surveys, design meeting	Catherine: enacted 7 lessons, 6 interviews, 5 surveys, ½ design meeting	Kelly: enacted 2 lessons, 2 interviews, 2 surveys
Nate: designed 8 lessons, enacted 8 lessons, 4 interviews, 4 surveys, design meeting	Maria: enacted 3 lessons, 3 interviews, 3 surveys, design meeting	Ted: enacted 2 lessons, 2 interviews, 2 surveys
Kyle: enacted 9 lessons, 9 interviews, 9 surveys, design meeting	Natalie: enacted 3 lessons, 3 interviews, 3 surveys	Kristen: enacted 2 lessons, 1 interviews, 2 surveys
Henry: enacted 9 lessons, 9 interviews, 9 surveys, design meeting	Nancy: enacted 3 lessons, 2 interviews, 1 surveys	Quincy: enacted 3 lessons, 2 surveys
		Dana: enacted 1 lesson, 1 survey, 1 interview

Data Analysis. We analyzed design workshop ethnographic field notes, transcribed interviews, interview notes, survey responses, and artifact design changes. We coded instances of professional growth using Clarke and Hollingsworth’s (2002) Interconnected Model of Professional Growth (IMPG). In accordance with the model, we coded for changes in: 1) knowledge, belief, or attitudes; 2) classroom practice; 3) salient outcomes; and 4) use of new materials. This allowed us to identify instances of change and determine where learning occurred. In the second cycle of coding, Edelson’s (2002) framework guided our coding for three types of learning: domain learning, design framework, and design methodology. This allowed us to understand the types of instructor learning generated by design participation.

Table 2: Duration and Engagement

High Design Time Group (HDT)	
HDT Participants	High Engagement Scores
Nate 106	Catherine 5
Frank 79.5	Frank 4.69
Kyle 39.5	Henry 4.67
Henry 39.5	Low Engagement Scores
Maria 26.5	Nate 4.42
Catherine 23	Maria 4.21
	Kyle 3.56
Low Design Time Group (LDT)	
LDT Participants	High Engagement Scores
Natalie 10.5	Natalie 4
Nancy 8.5	Nancy 5
Ted 7	Ted 5
Kelly 7	Kelly 5
Quincy 4	Low Engagement Scores
Kristen 4	Quincy 2.83
Dana 3.5	Kristen 2.25
	Dana 3

Table 3: Duration and Learning

High Design Time Group (HDT)	
HDT Participants	More Learning Scores
Nate 106	Nate 28
Frank 79.5	Frank 20
Kyle 39.5	Less Learning Score
Henry 39.5	Henry / Maria 6
Maria 26.5	Catherine / Kyle 4
Catherine 23	
Low Design Time Group (LDT)	
LDT Participants	More Learning Scores
Natalie 10.5	Ted 3
Nancy 8.5	Natalie / Quincy 1
Ted 7	Kristen 1
Kelly 7	Less Learning Scores
Quincy 4	Nancy / Kelly/ Dana 0
Kristen 4	
Dana 3.5	

Given the variation in instructor involvement in PDSA cycles, we developed a scoring system to stratify participants into groups based on their number of hours engaged in design activities, quality of engagement, and learning generated by participation in design. First, we calculated participants' total hours spent in design work (including lesson enactment) and performed a median split to separate the top and bottom 50% in regards to total time spent on design (duration ranges from 3.5-106 hours). Within each 'design time' category, we performed a sub-split to compare instructors on the basis of quality engagement (high or low) (see Table 2). The authors gave each participant an engagement score using a Likert scale (low engagement=1 and high engagement=5) based on his or her willingness to give feedback on classroom experiences and provide design revisions in each design opportunity (engagement scores range from 1-5). The authors scored engagement individually (interrater reliability is 84%), and inconsistent scores were averaged. We then created a second system using the initial 'design time' median split to perform a sub-split on the basis of learning (high or low) (see Table 3). Learning scores take into account the number of instances of teacher change and the kinds of learning generated (domain, design framework, and design methodology), and range from 0-28. The goal was to examine the following relationships: 1) the relationship between design time and level of engagement, and 2) the relationship between design time and learning. Given that research supports the learning potential for sustained, meaningful professional development (Garet, Porter, Desimone, Birman & Yoon, 2001), we predicted that we would find a positive relationship between design time and engagement and instructor learning.

Findings

We used qualitative methods to uncover the ways in which engaging in key elements of design might lead to learning. Our findings are limited by the small sample size that rendered inferential testing less optimal for examining the relationship between design time and engagement and learning. Our intention was not to establish correlation or cause, but rather to initially characterize the broad outlines of phenomena that might connect design experience to professional growth.

The results of this study suggest the relationship between design time, design engagement, and teacher learning are more complex than predicted. That is, an increase in design time was not always associated with an increase in engagement or learning. Our findings in the *High Design Time (HDT)* group do suggest, however, that a relationship exists between duration of engagement and types of learning when instructors engage in at least 20 hours of design work, consistent with Desimone (2009). Our findings support Edelson's (2002) perspective; ten out of thirteen instructors showed evidence of professional learning in at least one of three categories of learning: design domain, design framework, and design methodology. In what follows, we describe in more detail the relationship between duration and engagement and duration and instructor learning.

Duration and engagement

The results of our analysis found the relationship between amount of time spent on design and level of engagement to be unclear. Some instructors received engagement scores (ranging from 0-5) comparable to their design time (ranging from 3.5-106 hours). For example, all three instructors who participated in less than five hours of design work showed a low level of engagement (≤ 3). However, in most cases, the scores were unpredictable. For example, three *Low Design Time (LDT)* instructors earned the maximum engagement score of 5, outscoring most of their *HDT* counterparts. The results of the level of engagement surprised us for particular instructors even within design time groups. Catherine, who only spent 23 hours on design work, was the most engaged instructor in the *HDT* group with a score of 5. Nate, who participated in lesson design and spent the most time in design activities (106 hours), received only the fourth highest engagement score (4.42).

Duration and learning

Our findings indicate that increased duration of engagement in design activities is linked to increased and varied types of learning. Learning scores range from 0-28. The *HDT* instructors (duration ranging from 23-106 hours), and four of seven *LDT* instructors (duration ranging from 3.5-10.5 hours) showed evidence of domain learning. All six of the *HDT* instructors, but only one of the seven *LDT* instructors showed evidence of design framework learning. Only one *HDT* instructor of the thirteen total participants showed increased understanding of PDSA cycles, our design methodology. In all cases, the *High Design Time* (learning scores range from 4-28) instructors learned more than their *Low Design Time* counterparts (learning scores range from 0-3).

High Design Time (HDT) group

Our findings in the *HDT* group suggest that a relationship between time spent in design and learning may exist, but another variable may be at play. The *HDT* instructors were the only instructors in the study to participate in

“ACT” activities, which engaged instructors in making design decisions with researchers to refine the lessons. In the *HDT* group a relationship between time spent in design and learning clearly exists, as we predicted. Instructors in this group participated in 23-106 hours of design work, and instructor learning scores range from 4-28. Instructors Frank and Nate, who co-designed the initial lesson drafts with researchers, participated in > 75 hours of design work and evidenced the highest learning scores (≥ 20). The other four instructors in the *HDT* group engaged in only 20-40 hours of design work and evidenced lower learning scores (≤ 6). While it seems that increased duration leads to increased learning, there is some unexpected variation within the “low-learning” group. Kyle had the highest number of design hours (39.5 hours) within the “low-learning group” yet had the lowest learning score (4), while Maria participated in only 26.5 hours of design, but generated a learning score of 6. Interestingly, Maria, like Kyle, had low a score of engagement, suggesting that engagement may not relate to instructor learning. It is important to note that while learning scores vary based on the number of occurrences, all instructors in the *HDT* group showed instances of learning about the domain and design framework, while only one evidenced learning about the design methodology. In what follows, we provide examples of the specific types of learning generated in the *High Design Time* group.

Learning about the Design Domain. All six *HDT* instructors exhibited increased domain knowledge. For example, throughout nine lesson enactments, Henry learned about the interactions between the local setting and the Comprehension and Synthesis² (CaS) Chart (see Figure 1), a language and literacy tool critical to our design. He showed a change in his beliefs about the usefulness of the CaS chart for his students, which resulted in a shift in his instructional practices, and in turn, student outcomes. Following the first lesson, Henry said: “...I like it, think it’s useful...fits my feeling of how we should approach information mathematically. It will help students figure out what they’re doing before they put a number on it.” (Henry Interview, 1st enactment). However, his students did not necessarily agree: “There is some distrust over the system whether this [CaS chart] is going to be beneficial to students. Students are unsure about how to complete the CaS chart. They especially struggle with Column C” (Henry Interview, 1st enactment).

Column A	Column B	Column C
What are the main issue(s) in this problem situation?	What is the key quantitative information you need to solve the issues in the problem situation?	Describe in writing how the information in Column B will help you address the issue(s) in Column A later in the lesson?

Figure 1. The CaS Chart.

Part A: What is the dilemma or are the main issue(s) in this problem situation?	Part B: What quantitative information is important to address the issue(s)?
Part C: Describe how you might use the information in Part B to approach the issues in Part A. (without using numbers)	

Figure 2. Kyle’s Adapted CaS Chart.

However, by the ninth lesson enactment, Henry gained insight into the usefulness of the CaS chart for his students, resulting in a shift in instructional practice: “I talk about the CaS as a tool for helping them understand reading...separate thinking into small parts, and start to organize a strategy for calculating...Thinking through a strategy before calculating is something I’ve added to the discussion about the CaS chart.” (Henry Interview, 9th enactment). Henry evidenced the impact of this change on student learning, reflecting on his enactment of the CaS Chart with a new class:

Students gave decent reviews of the Cas Chart...reporting verbally that it was worth the time to talk things out and sort information. One student specifically said that this [CaS Chart] matched the way she likes to think...Both classes...recognized the value of Column C... They were thinking about how they might approach the problem before they dive into it. We had...positive vibes from the class as they were discussing what they've found in that third column.” (Henry Interview, 9th enactment)

It is important to note that while this data evidences what domain learning looks like, other examples of design domain learning do not evidence the impact of instructor learning on student outcomes, as this example does.

Learning about the Design Framework. All six *HDT* instructors learned about the design framework. In this example, Kyle learned about the core ideas behind creating language and literacy supports. After his first enactment of the CaS chart, Kyle reflects: “Having the two columns filled in with examples was helpful as acquiring the tool and material at the same time isn’t good. I don’t think the CaS chart was useful this time

because of the scaffolding. It might become useful in the future when they do it on their own.” (Kyle Interview, 1st enactment). Following the second enactment, Kyle re-evaluates his belief that scaffolding is useful: “You should get rid of the scaffolding. It’s still not useful. You should introduce the CaS chart in a short lesson on it’s own.” (Kyle Interview, 2nd enactment). In the first two lesson enactments, Kyle builds understanding of how the design of CaS chart (i.e: scaffolding) impacts students. Following this cycle of testing, Kyle reflects with instructors and researchers about the purpose and formatting of the CaS chart, “Column C seems superfluous by the time Column A and B are completed. The directions instruct students to complete column A then B then C, but it is more cyclical. Students should know that it is an iterative process.” (Kyle, Instructor Design Meeting) Kyle developed a new version of the CaS chart, adapting it to fit his students’ needs (see Figure 2). He has continued to use this adapted version of the CaS chart in his classes, and is presenting his adapted version of the CaS chart at a practitioner’s conference this year.

Learning about Design Methodology. Only Maria, a *HDT* instructor, showed evidence of increased understanding of PDSA cycles. While PDSAs guided and documented design activities, instructors did not use them directly. Maria became familiar with PDSAs through her participation in design activities, including interviews, surveys, and participation in the in-person meeting. As a result of this familiarity with PDSAs, Maria is presenting PDSA as a tool for curriculum development to colleagues, and has sought additional consultation with the research team to gain a better understanding of this design methodology. Although it is unlikely that Maria would have gained familiarity with PDSAs without engagement in this work, without Maria’s subsequent presentation as a external prompt, increased duration would not likely result in design methodology learning.

Low Design Time (LDT) group

The relationship between duration and learning is unclear in the *LDT* group. Instructors in this group participated in 3.5-10.5 hours of design work, with learning scores ranging from 0-3. Two of the four instructors with the most design time (7-10 hours) were also in the “high learning” group, but Nancy, who spent the second highest amount of time in design (8.5 hours), and Kelly (7 hours) do not evidence learning at all. In contrast, Quincy, who only engaged in four hours of design, evidenced one instance of learning. In the *LDT* group, four of seven instructors evidenced learning; three increased design domain knowledge, but only one, Ted, evidenced learning about both design domain and design framework. In this example, Ted builds his understanding about how the underlying design ideas of the Double-Entry Journal (DEJ), a language and literacy tool, interact with the time allotment for his class as he engages in a post-enactment survey and interview. This is an important example because it provides evidence that engaging in “STUDY” activities may lead to learning.

Before the interview, Ted completed a survey providing feedback on the lesson, in which he wrote: “You may want to consider using the DEJ with a shorter lesson.” (Ted Survey, 1st enactment). During the interview, Ted discusses two different ways to save time while still including the DEJ:

It’s just a lot of reading. It might be helpful for students to see a copy of the DEJ upon completion of the reading so that they would not have to go back and reread the introductory instructions...the task [DEJ] is out of context. I think it would have more meaning if they had it in the context of the lesson...Then question six...You could use the DEJ there. In the [left column of the DEJ] say ‘Yes’ or ‘No’ and ‘Why’, in the [right column of the DEJ] have students use statistics to back up [the left column]. You could take most questions and turn it into a DEJ. (Ted Interview, 1st enactment)

Through DEJ enactment and reflection in the interview, essential components of PDSAs, instructor Ted learned how to reach a design solution that would address students’ needs, rather than simply eliminate the DEJ. The two ideas he presented in the interview, displaying the chart for students after the reading and embedding the chart in existing mathematics questions are evidence of Ted’s learning about the DEJ’s importance.

Ted spent the median number of hours (7) engaged in design work in the *LDT* group, but received the highest learning score (3). Why did Ted stand out amongst his colleagues, as other instructors with similar design time and engagement scores did not exhibit any learning at all? Nancy and Kelly, who spent 8.5 and 7 hours on design, respectively, received learning scores of zero, but like Ted, received the highest engagement score possible. This suggests that either engagement is unrelated to learning for these instructors, or that Nancy and Kelly learned in ways that we could not capture using the Edelson (2002) model.

Discussion

Very commonly, instructors involved in design-partnerships only engage in some elements of design. We have offered some granular, though very preliminary, evidence of how engagement in specific elements of design

contributes to certain types of learning. This work helps us understand the role of time in design experiences. The relationship between learning and time spent in design activities seems clear for instructors who participated in more than 20 hours of design work, but this relationship is less apparent for the *LDT* group. Further, the relationship between duration and engagement remains unclear. It is possible that some instructors learned information that was not captured in our current analysis. The relationship between engagement and learning might also have impacted our findings. That is, instructors who were not as engaged during design meetings or phone interviews may not have vocalized their learning. Future research should examine more explicitly the relationship between engagement and learning.

In our analysis of instructor learning an unexpected category of learning emerged: pedagogical design capacity (PDC), an instructor's ability to recognize and employ resources to adapt existing, or develop new, instructional materials (Brown & Edelson, 2003). In this work, five of the thirteen instructors (and, in the *HDT* group, five out of six) adapted lessons to better support their students' needs (Barab & Luehmann, 2003). Adaptations included changes to the language and literacy tools embedded in the lessons (as evidenced in Kyle's example above), creating additional mathematics questions, developing and integrating examples in areas where students struggle, and developing new problem contexts. As it became clear that the time instructors spent in design work played a role in kinds of learning, we began to see that PDC is consequential, resulting from the confluence of increased understanding of the design domain and design framework. Of the five instructors who evidenced increased PDC, none evidenced learning about design methodology; thus, we do not believe that this is a critical category of learning in the development of PDC. Although PDC is indicative of certain types of learning, we argue that it is, in itself, an important type of learning generated from design, as adaptation of curricular materials to align with local needs is critical for effective implementation (Barab & Luehmann, 2003). It is important to note that all instructors who evidenced PDC engaged in "ACT" design activities, while the others (with the exception of 1 (Maria)) did not. While this may be a significant factor in developing towards PDC, more work must be done to better understand this relationship. While this study evidences increased PDC in instructors, it does not shed light on the alignment of adaptations to designers' intentions; future work is necessary to understand how to support instructors in adapting materials to meet their needs while maintaining integrity of the design solution (Davis, Beyer, Forbes, & Stevens, 2011).

Our findings highlight the opportunities for learning provided by design. Importantly, these results suggest that participation in design work can take on different forms; teachers do not necessarily need to be involved in every aspect of the design process to learn from the experience. Instructors in this work differed in duration and forms in which they were involved with PDSA cycles. In general, our findings suggest that instructors who spent more time in design experienced more learning. Yet, four of the seven instructors who spent less than 10.5 hours in design showed evidence of learning, suggesting that even a short period of design activity can present a learning opportunity. However, more work, which systematically assigns instructors to specific PDSA activities, keeping duration the same, must be undertaken to better understand the relationship between PDSA activities and learning. For example, the 6 instructors who participated in "ACT" activities were also in the *High Design Time* group. All instructors participated in "PLAN", "DO", and "STUDY" activities, however, increased participation in these activities also increased overall duration spent in design activities. With the exception of Ted, who evidenced immediate learning as a result of participation in a "STUDY" activity, our data is insufficient to parse how engagement in specific PDSA activities generate learning.

This work has important implications for future design efforts, especially in models concerned with professional learning, that engage teachers in some, and not all, elements of design. We argue that it is critical, as we did with PDSA cycles, to keep track of instructor duration, the ways in which instructors touch design, and quality of engagement. The use of PDSA cycles allowed us to trace the evolution of both the lessons and participants. With each iteration, informed by the instructors, the PDSA cycles captured design problems and potential solutions. Thus, we were able to simultaneously gain insight into the lessons themselves, as well as the people who were enacting them. We believe that embedding these practices in design work provide a rich way of talking about the kinds of learning generated from design.

Endnotes

- (1) All 13 instructors participated in "PLAN", "STUDY", and "DO" activities, but only the six instructors at the design meeting participated in "ACT" activities. We did not collect data on "PLAN" activities.
- (2) The CaS chart is a tool the researchers developed to support student comprehension in mathematics word problems.

References

- Bailey T., Jeong D. W., Cho S. W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review*, 29(2), 255-270.

- Barab, S. A., & Luehmann, A. L. (2003). Building sustainable science curriculum: Acknowledging and accommodating local adaptation. *Science Education*, 87(4), 454-467.
- Borko H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3-15.
- Boylan H. R. (2002). *What works: Research-based best practices in developmental education*. Boone, NC: Continuous Quality Improvement Network, National Center for Developmental Education, Appalachian State University.
- Brown, M., & Edelson, D. (2003). Teaching as design: Can we better understand the ways in which teachers use materials so we can better design materials to support their changes in practice. *Design Brief. Evanston, IL: Center for Learning Technologies in Urban Schools*.
- Clarke D., Hollingsworth H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947-967.
- Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77(3), 211-244.
- Collins, A. (1992). *Toward a design science of education* (pp. 15-22). Springer Berlin Heidelberg.
- Darling-Hammond, L. & Richardson, N. (2009). Research review/teacher learning: What matters. *Educational leadership*, 66(5), 46-53.
- Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational researcher*, 34(3), 3-14.
- Desimone L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Edelson D. C.(2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences*, 11(1), 105-121.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American educational research journal*, 38(4), 915-945.
- Gomez, K., Gomez, L. M., Rodela, K. C., Horton, E. S., Cunningham, J., & Ambrocio, R. (2015). Embedding Language Support in Developmental Mathematics Lessons Exploring the Value of Design as Professional Development for Community College Mathematics Instructors. *Journal of Teacher Education*, 0022487115602127.
- Greeno J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5-26.
- Herod, L. (2002). Adult learning from theory to practice. Retrieved March, 2, 2009.
- Joseph, D. (2004). The practice of design-based research: Uncovering the interplay between design, research, and the real-world context. *Educational Psychologist*, 39(4), 235-242.
- Koehler M. J., Mishra P. (2005). Teachers learning technology by design. *Journal of Computing in Teacher Education*, 21(3), 94-102.
- Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design (tm) into practice. *The journal of the learning sciences*, 12(4), 495-547.
- Langley G. J., Moen R., Nolan K., Nolan T., Nomran C., Provost L. (2009). Using the model for improvement. *The improvement guide: A practical approach to enhancing organizational performance*.
- Levin H., Calcagno J. C. (2007). *Remediation in the community college: An evaluator's perspective*. New York, NY: Community College Research Center, Teachers College, Columbia University.
- Little J.(1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *The Teachers College Record*, 91(4), 509-536.
- Merseth K. K. (2011). Update: Report on innovations in developmental mathematics—Moving mathematical graveyards. *Journal of Developmental Education*, 34(3), 32-33.
- Penuel W. R., Fishman B. J., Yamaguchi R., Gallagher L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44(4), 921-958.
- Seidman I.(2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. New York, NY: Teachers College Press.
- Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator*, 1(3), 4-16.
- Voogt J., Laferrière T., Breuleux A., Itow R. C., Hickey D. T., McKenney S. (2015). Collaborative design as a form of professional development. *Instructional Science*, 43(2), 259-282.

No One Ever Steps in the Same Discussion Twice: The Relationship Between Identities and Meaning

Murat Öztok, Lancaster University, m.oztok@lancaster.ac.uk
Maarit Arvaja, University of Jyväskylä, maarit.arvaja@jyu.fi

Abstract: The concept of identification is a relational construct; that is, identities are not static but rather negotiated based on available material and symbolic resources. However, we know relatively little about how identities play a dual role when students collaborate. The aim of this paper is to explore this process through multiple case studies: we aim to explore how identities are enacted and used in making personal sense and understand the content knowledge, while at the same time we are interested in how this process can take a form of renewing process in the sense that the identities enacted are themselves changed, transformed or re-negotiated. Our results show that due to its dual role, identities mediate collaborative learning not only because knowledge is constructed in relation to identities but because online selves are articulated and constructed in relation to knowledge construction.

Keywords: identification, identity, I-position, online learning

Introduction

Learning is simultaneously an individual and social process (Brown, Collins, & Duguid, 1989; Cole, 1996). It is the material, symbolic, and intellectual reconstruction of self; a process of discovering and articulating oneself in relation to others. In other words, it is a process of knowing the self through mediation between self and others. What mediates between the self and the others – between the individual and social – is referred to as identification. Thus, the process of education is the process of constructing and shaping individuals' identification (Holland, Lachicotte Jr, Skinner, & Cain, 1998).

The concept of identification (or identities-in-practice) is a relational construct; that is, identities are not static but rather negotiated based on available material and symbolic resources. It is through these negotiations that individuals identify themselves or are identified with various cultural categories within particular situations (Buckingham, 2008; Hall, 1996). In this sense, identification has both individual and social aspects by which individuals perceive, categorize, situate, and understand themselves and those around them. Such an understanding is particularly important since it “reconceptualizes learning from an in-the-head phenomenon to a matter of engagement, participation, and membership in a community” (Nasir & Cooks, 2009, p. 49). Indeed, much learning sciences research has long demonstrated that sense-making has profound impact on collaborative learning practices (Holland et al., 1998; Stahl, 2006). Despite its significance for group work, however, we know relatively little about how identities play a dual role when students collaborate. In particular, we do not exactly know how identities are utilized for sense-making, and in return how the sense-making process leads to re-negotiations of identities. The aim of this paper is to explore this process.

Background and rationale

This research is nestled at the intersection of three related theoretical frameworks: sociocultural learning theories, Dialogical Self theory, and the concept of identification. At the center of this intersection is the idea that each individual has a unique perspective through which they make sense of the world around them and create their own narrative (Cole, 1996; Ochs & Capps, 1996; Wortham, 2001). These narratives, then, constitute “boundary structures” for learning since they influence the way people understand and act in the world. Similar accounts can be found in more recent studies. For example, Rajala and Sannino (2015) use the notion of “personal sense” which allows conceptualizing task interpretation within the wider perspective of the students' life and interests. According to the research reported in their work, it is important that task resonates with the reality of the student's own life; otherwise, it is likely that students interpret the task as uninteresting or irrelevant, which can create discrepancy between sense and meaning. Their study concludes that for a subject, meanings exist only in relation to personal sense. Such a relationship between meaning and sense also provides possibilities for the re-negotiation of one's identities and “being, thinking and doing” (Akkerman, Admiraal, & Simons, 2012). Indeed, our previous works have already illustrated that drawing on personal experience (i.e. professional background or personal interests) to make sense of the subject-matter provides opportunities for individuals to advance personal and collective understanding (Oztok, 2013, 2014; Arvaja, 2015).

Dialogical Self theory (Hermans, 2003) is associated with Bakhtin's (1984) concept of voice, especially multivoicedness. It provides a tool for understanding how different perspectives manifested in identities (or I-positions) are embedded in the person's self. An "I-position" of a person, according to this theory, "is a particular voice that has been internalised in one's self-presentation" (Akkerman et al., 2012, p. 230). Consequently, the self is diverse in the sense of *multiple* I-positions that can be used in expressing oneself (e.g., I-position or identity of a professional, student, mother or hard worker). The concept of multiplicity can help with understanding peoples' varying positions and identities. For example, in addition to a voice or identification of a student, students' also have more personal voices, inner voices, containing personal and intimate experiences, such as an assertive voice or passive voice that also shapes their sense making as a student (Wortham, 2001). Consequently, when people are talking or thinking, they often integrate, contrast, and move between different I-positions (Hermans, 2003).

The concept of identities-in-practice implies an interrelationship with the broader collective or social group (Buckingham, 2008; Holland et al., 1998). In particular, identity is something that is unique to each individual due to unique personal biography (Buckingham, 2008; Linell, 2009) while at the same time, it refers to a collective sense due to a sense of belonging (Hall, 1996; Wenger, 1998). Research has illustrated that individuals often speak the words of the groups or society to which they belong; therefore, the social world has an important role in the construction of self: it mediates the voice of traditions, generalized others, institutions, groups, communities, colleagues, relatives and friends through the dialogical participants (Hermans & Kempen, 1993; Linell, 2009). In this way the voices of others become woven into what one says and as part of one's thinking, reasoning and acting, as part of one's different identities or I-positions. According to Wortham (2001) "speaking with a certain voice means using words that index some social position(s) because these words are characteristically used by members of certain group(s)" (p. 38). Therefore, the self is not a pure intra-psychological process but a relational process that includes the social environment (Akkerman & Meijer, 2011).

The concept of identities-in-practice also connotes discontinuity in its nature; that is, different identities, identifications, or I-positions changes according to the people and type of situations one encounters (Akkerman & Meijer, 2011; Hermans, 2003). This is what Goffman (1959, 1983) calls impression management: the ways in which people represent themselves and engage with others is a socially situated process. The concept of impression management suggests that particular situations evoke particular parts of the self and that identities enacted in the situation depend upon the presence of others in the context. Therefore, the content of what is said does not only reflect person's attitude towards the object at hand but also person's attitude towards preceding and succeeding actions and identities of others (Akkerman & Meijer, 2011). Even though, the presence of 'real' others influence identities evoked in the situation, dialogical approach also stresses the importance of generalized others, "virtual" others, or different third parties evoked by a situation, and their influence on the identities enacted (Linell, 2009). In a Bakhtinian sense, the *other* is pervasive even though the person is alone (Hermans, 2003). Indeed, our previous work has demonstrated how contextual, symbolical and material aspects have an influence on the identities represented and enacted (Oztok, 2014). For example, a text used as a learning material can be regarded as a third party, a voice, that evokes enacting different identities relating to one's professional, personal, national or political self or identification (Arvaja, 2015). These identities, in turn, are used for making sense of the text or related discussion (Oztok, 2013) or the texts are used in making sense of one's identities (Arvaja, 2015). Texts may invite different aspects of the self into dialogue within the self (internal dialogue) or with others (external dialogue).

Our ultimate aim is to study a two-way process: we aim to explore how identities are enacted and used in making personal sense and understand the content knowledge, while at the same time we are interested in how this process can take a form of renewing process in the sense that the identities enacted are themselves changed, transformed or re-negotiated. In this two-way process not only knowledge is co-constructed through different identities but also identities are negotiated, and therefore, a change may also occur at the level of one's perspective of the world and of the self.

Current research

We demonstrate our conceptual approach through multiple case studies (Creswell, 2006). Two purposefully selected cases (named hereafter Case 1 and Case 2) comprise data from two different online courses that took place in different universities, and provide an in-depth exploration of a certain phenomenon in a given context. Each case focuses on the different aspects of the phenomenon at question while keeping the concept of identity at the center of analysis; Case 1 probes how identities are enacted in making personal sense whereas the Case 2 exemplifies how this process enables identities to be renegotiated. Next we describe the contexts, subjects, data, and analytical approaches used in the studies.

Case 1 is based on a fully-online graduate level education course offered at a large Canadian research university. Typically, these graduate courses have students from diverse historical and cultural backgrounds, from

different geographical locations, and of various ages and professions. The course comprised twelve modules, each corresponding to one week. Students were asked to introduce themselves (create their profile pages) and meet with their peers (read and comment on others' profile pages) in the first week and submit their final paper in the last week. In each module, one or two students acted as moderators: they facilitated discussion throughout the week, kept discussions on track, and offered a summary of the week's issues; they provided opportunities for sustained discourse, increased interaction, and rich discussions. The online discussion occurred asynchronously; the environment does allow synchronous communication through instant messaging, but such activity was not mandatory (nor was it a major communication tool) in this course. 14 students enrolled in the course and worked together as a single group throughout.

In order to illustrate the variety of identification traits individuals manifest, participants' profile pages (personal pages in which students create their online existence by introducing themselves with their own words and a picture or avatar) were analyzed. An online persona is created for each participant to materialize the salient identification traits in their profile pages. Considerable attention is paid to choosing individuals who use a variety of identities and selected four individuals who maximize the exploration of the phenomenon. Then, the research team analyzed the notes in these threads semantically (Fairclough, 2001) with three different lenses: (1) the use of identification, (2) the process of knowledge construction, and (3) the relationship between the two. The use of identification is analyzed simply with probing "who says what" in language-in-use. The analysis of language-in-use reveals how identification traits are manifest in ways of saying, doing, and being: "to understand anything fully, you need to know who is saying it and what the person saying it is trying to do" (Gee, 2011, p. 2). Since the language-in-use is linked with the role that identification traits play in mediating experiences among individuals, "who says what" is critical for understanding the otherwise hidden intersections between identification, situated meaning-making, and knowledge construction in online learning environments.

Case 2 is based on a three-month online science philosophy course for health science students and professionals at a university in Finland (see Arvaja, 2015). The course consisted of six learning tasks, all of which dealt with historical approaches in the philosophy of science. Each task was a reasoning task where the students were first supposed to read a given text (or texts) dealing with a particular approach within the philosophy of science. In reasoning about the task, the students were asked to use their prior experiences or conceptions about their own field of science or work as resources in interpreting the texts. Based on these tasks, each of the students was first supposed to write an individual reasoning text. In the next phase, the students posted their individual writings onto a shared web-based (asynchronous) discussion forum, and their task was first to read each other's writings and finally to have a shared discussion based on these.

For the purposes of this study one student, Aino, was selected for the analysis to exemplify a change in her I-positioning. The analysis leaned on dialogical approach to narrative self-construction (Wortham, 2001). According to this view the self is constructed through relationships with others and emerges through constant interactional positioning with respect to others in daily life. Hence, according to Wortham (2001), the self (and different I-positions within) is narratively constructed through positioning different voices in the social world in relation to each other, and by positioning oneself with respect to these voices. Data consisted of Aino's individual writings and asynchronous discussion postings. From Aino's discourse, two layers of positioning, that is, voicing and evaluating (ventriloquation), were analyzed. The process of voicing (i.e. characterizing oneself and others) draws on positions and ideologies from the larger social world, as the others described come to speak like recognizable types of people with their related characteristics, viewpoints or ideologies (Wortham, 2001). In ventriloquation (Bakhtin, 1984) one evaluates the other voices by differentiating or identifying and/or by distancing or standing closer (e.g. taking a critical or supportive stance) with these voices.

Findings

Case 1: Identifications enacted for making personal sense and co-constructing knowledge

The thread being analyzed here is a slice from an online asynchronous discussion among a cohort of students. While each student in this cohort is included in the analysis, the excerpt below focuses on how four purposefully-selected individuals (as explained above) utilize their identities when they collaborate.

Three students articulated their perspectives before Judith joined the discussion. The third note is worth quoting at large as it sparked an exchange of ideas around the issue of cultural diversity. Enacting her professional identity, a student wrote:

... I read [the weekly reading] differently. Here is why. I have worked with students from different cultures, students who are first generation Canadians whose parents have migrated

here; students whose parents are asylum seekers; students with a range of learning difficulties. I am convinced that teachers and educators have negative assumptions about these students – as if they know what's needed for them. I am not sure if diversity can ever lend itself to equality in classrooms because teachers don't know what they are dealing with. Do you think students expect that teachers will understand their cultural differences and requirements?

The rhetorical question at the end of this note became a focal point from which others departed by articulating their perspectives and experiences. Judith was the first to react; she acknowledged that teachers' beliefs about cultural differences are important:

I agree, [anonymous student 1], that teachers have assumptions about students. ... In my experience, it is very difficult to change other teachers' beliefs about cultural differences. It is because the term 'cultural diversity' is often misused (especially by stakeholders) – as though it is more important that teachers, educators, school principals, the director of education, etc. say that they have well thought out “cultural diversity” ... than they actually understand it. ... This is the reason why teachers have misconceptions about their students' cultural needs. In my experience, teachers are just worried about ticking the boxes off in official reports when it comes to cultural diversity.

Similar to the student in the previous note, Judith enacted her professional identity. However, while she agreed with the issues identified in the previous note, she also provided an alternative perspective. As a teacher, Judith believed that cultural diversity means more than addressing teachers' negative assumptions. As she continued articulating her understanding, Judith started enacting her maternal identity, explaining that an authentic learning context requires active dialogue between parents and teachers:

... I think that it is not only the responsibility of the teacher but also the parent to help establish an equal learning environment for all students. As stated in previous posts by others and you, as parents we want to make sure nobody is being left out, we want to make sure we are being inclusive and doing our best to help teachers to better accommodate our kids' needs. How does a teacher provide authenticity just by herself? How does a parent expect teachers to do everything?

Enacting both her professional and maternal identity, Judith identified a source of disagreement based on her experience. It is through this type of identification that she was able to provide a counterargument; that is, the tension between diversity and equality is not only about teachers' attitudes but also requires parents' active involvement. Manu responded to this message, also enacting her maternal identity along with her teacher identity. She built on Judith's perspective by further elaborating her experience:

I totally agree with you both – though you have different points on teachers. I appreciate the usefulness of taxonomies in general, but think human nature is too messy to be classified. ... [F]or managers and principles inclusivity is about numbers, but teachers have nothing to do with that. Diversity is not about numbers ... As a teacher, when I think of diversity what comes to my mind is students who not only have different learning needs but also [students] who come from diverse social backgrounds. Learning diversity encompasses diverse learners with different academic needs, such as students with disabilities and English language learners – such as my kids. But then, I see a big mismatch between the articles and my kids' schooling. I wonder if the authors of these articles have any kids or ever taught at schools. Judith is right in a way, how can a teacher do it all?

Manu acknowledged both sides' perspectives on diversity and the capacity of teachers to recognize and appreciate diversity in the classroom. By so doing, she attempted to clarify differences between Judith and Anonymous Student 1, and tried to link the points of disagreement between the two. She then incorporated her understanding of diversity based on her experience as a teacher and a mother. Manu continued:

Maybe a different approach would be to clarify to what extent learning differs by calling it culture. Although we can call on a number of stock words – nationality, race, gender, ethnic group, social-class, sexual orientation, etc, etc, etc – how they impact on learning is not

straightforward. One thing is for certain; students differ in one way or another, but is it enough to make claims on learning? Whether it be race, class, gender or language this thing we know as culture helps give students identity. That's all. Let's agree on that.

Tackling the relationship between culture and learning, Manu suggested a new lens for understanding the disagreement and started to develop her own hypothesis in order to unite strands of consensus. Then, she continued:

But [the weekly readings] argue that it has an impact on performance, learning styles and learning rates, learning experience and expectations, attitudes and achievements. Isn't it downright wrong? How could you categorize people so easily based on the ideals of culture? This is an open-ended question for you all; can you simply categorize people in your daily life just like that? Let me tell you; [the weekly readings] assume culture [to be] monolithic. Like the principles and managers you mentioned above, and the ones that I've been working with so far, I believe [the authors] try to ensure they 'deal with' the diversity. They just idealize it; it is far from real-life situations. Simple is that...

Manu tested her own hypothesis by providing rhetorical answers to her own questions based on her experience as a teacher. She suggested that the weekly readings, perhaps, offer an idealized understanding of diversity and thus do not reflect real-life situations.

Two other students replied and agreed with Manu, enacting their professional identities. Ken was the third replying back to Manu. He enacted his ethnic and professional identity, and picked up on Manu's new proposal of the lack of congruence between idealizations and real-life situations in learning and teaching. He tried to reconcile differences among them by suggesting that as a teacher, he believes readings are “just idealized scenarios” and that there are “unavoidable power tensions between cultural groups”. Ken continued enacting his professional identity:

I'd agree, culture is difficult to quantify, in addition, students differ so much within their respective cultures so it is not unified. The whole aspect of the impact of culture on teaching and learning, how we accept, accommodate and celebrate student diversity is a fascinating element of our day-to-day job as teachers. This is what we all agree so far.

Ken's cohesive view of disparate ideas led others to build on agreed facts, transitioning from debating to knowledge construction. Chun-Li was the second one to reply. She enacted her ethnic identity and further discussed “the idealized scenarios” by providing examples from her learning experience:

I did my MA in UK and I felt more Chinese then[sic] ever. ... But it doesn't mean that I was quiet or shy. Idealized scenarios? Yes! But then you are also right Ken that all of my teachers, lecturers, instructors, professors – what ever you call them – accommodated differences. But how do they accommodate? I think we have to understand what we mean by difference. Difference or diversity is not about where we were born or what kind of skin color we have. Diversity or difference is not about geographical location. Where I was born, where I studied, and where I am right now are completely different locations. So, where do I fall into?

Chun-Li built on Ken's summary and exemplified the current understanding based on her experience, testing the proposed synthesis. She continued:

... again, how do teachers accommodate these differences? Maybe [reading 1] offers an answer for dealing with different cultural groups: an ‘inclusive’ approach, which not only incorporates cultural perspectives from minority groups but also challenges the dominant model. I think this explains what I faced when I was in UK. I found that the lecturers were good at allowing individuals to express themselves. In my experience this allowed inclusivity because cultural practices are often shaped by individuals and their own dynamic. I look forward to future discussion.

Chun-Li was able to draw from Ken's summary, and bring together her experience and the readings to construct knowledge. According to Chun-Li, “if diversity is thought of as a matter of individuality, then the issue of the

inclusion or exclusion can be better understood". Ken enacted his student identity in his response and noted the importance of a learning community:

As classmates we want to make sure nobody is being left out from the discussion, we want to make sure we are being inclusive in all our discussions and activities and doing our best. Therefore effort also needs to be made on the students part, on our part. Perhaps, we can consider trying what [reading 1] suggests and help each other, especially those who are excluded. In sum, I think the key is being aware of any exclusiveness and making the effort to establish a community.

This particular note from Ken received great attention from his peers (indeed, this is the most *replied-to* note throughout the course according to the automated-log data) and constituted a point of agreement for the whole class.

Summarizing the weekly discussion and affirming Ken, Manu synthesized that "educators should teach their students ways to foster diversity in all its forms (ethnic, sexual, gender, learning styles, etc.) and create a sense of community to create inclusive educational contexts". Judith built on this and summarized that "most of us have the best of intentions as teachers and parents, but as all of you put it so well, life... happens!".

Case 2: Identities re-negotiated through the process of making personal sense

This case demonstrates another aspect on the role of identification and identities in learning. It demonstrates a situation when I-positioning itself is re-negotiated as a result of engaging in online discussions and reading of course material. Examples analyzed here are drawn from one student, Aino, who is studying part-time in the online course while also working full-time as a physiotherapist. Next example demonstrates how encountering a different voice or perspective in contrast to Aino's own leads to internal dialogue (Linell, 2009):

Aino: "In this week I've been reading texts from web and on paper, and frankly speaking I feel that my head is somewhat overloaded. One doesn't really know anymore what to think of what issue, and now one is questioning one's own work and science and research and whatever it was and I cannot make any sense of this, there are simply too many ideas. So I decided to look once more at this 'what is science' issue, on the basis of Niiniluoto's article, because it bears most relevance to me personally. I have always considered myself a type very much oriented to science and especially to natural science, and being somehow schematic and mathematical. For this reason it feels somehow overwhelming to question everything now. Admittedly at the same time really interesting, too. What's hard for me is that one can keep elaborating the idea endlessly and never reach a solution."

While voices are drawn from the complex social world, they get engaged in a dialogue that involves multiple perspectives and often conflicting positions (Wortham, 2001). A double-voiced discourse often involves "a conflict". This is what happens in the example above. Aino's inner tension and confusion she is facing when reading the philosophical course material is explicitly expressed in her discourse: "I feel that my head is somewhat overloaded. One doesn't really know anymore what to think of what issue... it feels overwhelming to question everything now". This results in "questioning one's own work, science and research". Aino voices and characterizes herself as "a type oriented to natural science", and it seems that from that position she is facing a challenge when being introduced with fundamental questions of the essence of science. Encountering a different view challenges Aino's current way of seeing and understanding. It seems that her orientation to natural science (according to her constant characterizing and identification) represents a dominant voice adopted from authorities in her working and study environment. According to Linell (2009), an authoritarian voice is often like cultural assumptions that the individual does not question and once the ideas of this voice are internalized it often becomes a kind of self-discipline. However, it seems that when Aino is introduced with texts that offer different perspectives or alternative conceptions (i.e., different ontological and epistemological voices of science) her beliefs pertaining to her "natural science oriented" position becomes questioned. This becomes more explicit as the course proceeds. Aino starts questioning the dominant authoritative voice behind her thinking in her discourse with other students:

Aino: Yes, indeed, this is precisely the way I see it in physiotherapy and for my own work. The problem just lies specifically in that, for example, at work people have too high regard for the views of natural science. One has to measure mobility and muscular strength etc. and compare

the results and assess effectiveness in that way. [...] Another issue I face at work is compilation of statistics. If I spend time at the ward discussing with a patient, talking about goals and motivation, listening to the person and evaluating her emotional state, without performing actual physiotherapy i.e. muscular or mobility or functional exercises, so can I then record the visit as a physiotherapeutic visit? As I didn't actually perform any therapy, but as much time was spent and after the discussion the patient is likely to be more motivated to engage in rehabilitation and more cooperative when we start actual training. [...] There's only the problem that I can't really mark on my daily nursing record sheet just that 'discussed about therapy'. Then one will cheat and take say a stretching or a quick inspection in the end. That's how it is; the emphasis is too much on natural science. :)"

While in the beginning of the course Aino has voiced herself as "considered myself a type very much oriented to science and especially to natural science, and being somehow schematic and mathematical" repeatedly, in this example from the end of the course we can see that she distances herself from that I-position. In her discourse she places this natural-science oriented voice more to an external voice than as a voice of her own e.g., "at work people have too high regard for the views of natural science". In this process of positioning (Wortham, 2001) Aino distances and differentiates her voice from the authoritative voice in the workplace and the values and ideologies it represents. Aino argues against the dominance of the natural sciences and describe work practices that reflect and support this dominance. Her redefined I-position as a physiotherapist is in conflict with the prevailing practices such as compilation of statistics. Aino's internal dialogue reveals a struggle between her professional/personal voice and the authoritative voice in her work community. In other words, Aino is engaged in a double-voiced discourse (Wortham, 2001) between her own I-position which makes personal sense (Rajala & Sannino, 2015) and the authoritative voice she recognizes in her work and related science practices. Her redefined professional I-position as a physiotherapist does not fit with the practices in the workplace and ideologies behind these practices. Through her redefined professional position Aino interprets her experiences in a new light through a new frame of reference which acknowledges the patient as a whole: not only the body; anatomy and physiology but also the mind; feelings, motives, and values in inter-personal interactions. A change in Aino's I-position is also explicitly stated in her last writing:

Aino: "At least for me this course has taught a quite different way of thinking for doing research and broadened my approach to science in general. It seems that I started from Sharply positivistic notions and ended up in a fairly broad and open view on the importance of qualitative research and human sciences, for example. It's good to stop and reflect on things and their meanings every now and then. At work one is often measuring just for the fun of it and it bears no significance, after all, to the patient let alone for science. Actually it may have been the most important lesson for me in this course; to consider what really significant science is. It is by no means about angle degrees and gauges but consideration of causal relationships more broadly and consideration of humans and interaction. Although it sometimes feels that thinking was really tangled, in the end one must say that this has been a good process."

Aino's statement clearly demonstrates how the things discussed, read and written in the course had an impact on her changed I-position. Aino implicitly states that what I think now in relation to science is different of what I thought before, therefore indicating also a re-negotiation of a scientific I-position. Scientific paradigms can be seen as collective voices which function as social positions in the self (Hermans, 2003). Such positions or voices are expressions of historically situated selves that are constantly involved in dialogical relationships with other voices. At the same time they are constantly subjected to differences in power as in this case where the dominant natural science position in the workplace is in contrast with more holistic position (minor position). It seems that exposure to the diversity and even opposed voices through the different approaches and perspectives in science challenges and changes Aino's scientific I-position. This change is further applied in seeing and thinking as a physiotherapist i.e., change in Aino's professional I-position.

Discussion and conclusion

This paper aimed to demonstrate the dual role identities play in (collaborative) learning. The underlying argument was that individuals construct, position, and understand themselves and those around them in relation to their identifications or I-positionings (Hall, 1996; Wortham, 2001). When we foregrounded identities-in-practice for analyzing and interpreting collaborative learning, the dual role of identities became clear: (1) identities are enacted and used in sense-making and understanding the content knowledge, and (2) this sense-making process can take

a form of renewing process since the identities enacted are themselves redefined and changed. Identities are of paramount importance for collaboration when learning is regarded as both social and individual process; as a matter of engagement and participation in a community. Our results show that due to its dual role, identities mediate collaborative learning not only because knowledge is constructed in relation to identities but because online selves are articulated and constructed in relation to knowledge construction.

References

- Akkerman, S. F., Admiraal, W., & Simons, R. J. (2012). Unity and diversity in a collaborative research project. *Culture & Psychology, 18*(2), 227–252. <http://doi.org/10.1177/1354067X11434835>
- Akkerman, S. F., & Meijer, P. C. (2011). A dialogical approach to conceptualizing teacher identity. *Teaching and Teacher Education, 27*(2), 308–319. <http://doi.org/10.1016/j.tate.2010.08.013>
- Arvaja, M. (2015). Experiences in sense making: Health science students' I-positioning in an online philosophy of science course. *Journal of the Learning Sciences, 24*(1), 137–175.
- Bakhtin, M. M. M. (1984). *Problems of Dostoevsky's poetics*. (C. Emerson, Trans.). University of Minnesota Press.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.
- Buckingham, D. (2008). Introducing Identity. In D. Buckingham (Ed.), *Youth, Identity, and Digital Media* (pp. 1–22). Cambridge, MA: MIT Press Journals.
- Cole, M. (1996). *Cultural Psychology: A once and Future Discipline*. Cambridge, MA: Harvard University Press.
- Creswell, J. W. (2006). *Qualitative Inquiry and Research Design: Choosing among Five Approaches* (2nd ed.). Thousand Oaks, CA: SAGE Publications.
- Fairclough, N. (2001). *Language and Power* (2nd ed.). London, UK: Longman.
- Gee, J. P. (2011). *An Introduction to Discourse Analysis: Theory and Method* (3rd ed.). New York, NY: Routledge.
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York: Anchor.
- Goffman, E. (1983). The Interaction Order: American Sociological Association, 1982 Presidential Address. *American Sociological Review, 48*(1), 1–17. <http://doi.org/10.2307/2095141>
- Hall, S. (1996). Who Needs “Identity”? In P. du Gay & S. Hall (Eds.), *Questions of Cultural Identity* (pp. 1–17). Thousand Oaks, CA: SAGE Publications.
- Hermans, H. J. M. (2003). The Construction and Reconstruction of a Dialogical Self. *Journal of Constructivist Psychology, 16*(2), 89–130. <http://doi.org/10.1080/10720530390117902>
- Hermans, H. J. M., & Kempen, H. J. G. (1993). *The dialogical self: Meaning as movement*. San Diego, CA, US: Academic Press.
- Holland, D., Lachicotte Jr, W., Skinner, D., & Cain, C. (1998). *Identity and Agency in Cultural Worlds*. Cambridge, MA: Harvard University Press.
- Linell, P. (2009). *Rethinking language, mind, and world dialogically: Interactional and contextual theories of human sense-making*. Charlotte, NC: Information Age Publishing.
- Merriam, S. B. (2009). *Qualitative Research: A Guide to Design and Implementation* (3rd ed.). San Francisco, CA: Jossey-Bass.
- Nasir, N. S., & Cooks, J. (2009). Becoming a Hurdler: How Learning Settings Afford Identities. *Anthropology & Education Quarterly, 40*(1), 41–61. <http://doi.org/10.1111/j.1548-1492.2009.01027.x>
- Ochs, E., & Capps, L. (1996). Narrating the Self. *Annual Review of Anthropology, 25*, 19–43.
- Oztok, M. (2013). Tacit knowledge in online learning: community, identity, and social capital. *Technology, Pedagogy and Education, 22*(1), 21–36. <http://doi.org/10.1080/1475939X.2012.720414>
- Oztok, M. (2014). Social presence and social absence: socio-cultural production of self in online learning environments. Presented at the The 15th Annual Meeting of the Association of Internet Researchers, Daegu, Korea.
- Rajala, A., & Sannino, A. (2015). Students' deviations from a learning task: An activity-theoretical analysis. *International Journal of Educational Research, 70*, 31–46. <http://doi.org/10.1016/j.ijer.2014.11.003>
- Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA: The MIT Press.
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, UK: Cambridge University Press.
- Wortham, S. E. F. (2001). *Narratives in action: A strategy for research and analysis*. New York, NY: Teachers College Press. New York, NY: Teachers College Press.

“Hearts Pump and Hearts Beat”: Engineering Estimation as a Form of Model-Based Reasoning

Aditi Kothiyal, Indian Institute of Technology Bombay, aditi.kothiyal@iitb.ac.in
Sahana Murthy, Indian Institute of Technology Bombay, sahanamurthy@iitb.ac.in
Sanjay Chandrasekharan, Homi Bhabha Centre for Science Education, sanjay@hbcse.tifr.res.in

Abstract: Professional engineers are often required to make estimates of physical quantities and processes, as well as user requirements and system performance. However, engineering undergraduates, as part of their curriculum, do not learn how to make such estimates. In order to design a technology-enhanced course in engineering estimation, we conducted a study of experienced engineers working on estimation problems. A three-phase process for engineering estimation, and the cognitive mechanisms underlying them, emerged from the analysis of this data. We highlight the roles of mental simulation and external representation in the estimation process. These results will be used to design learning environments for engineering estimation.

Keywords: engineering estimation, expert study, technology-enhanced learning environment

Introduction

Engineers build things: they build bridges across oceans and tiny memory devices that can store 30 full length feature films in less than 1 cm² of physical space. But before they build anything, they analyze whether they *can* build it, whether the designs are feasible or not (Linder, 1999; Mahajan, 2014). For instance, consider the following problem, “A toy company has designed a remote-controlled toy car for kids driven by a 5W hand-cranked generator. They expect that a child will be able to play with it for 5 hours before needing to crank again. Is this reasonable?” Professional engineers must routinely answer questions such as this, which involve estimating an unknown quantity and using it to make a judgement. Formally engineering estimation has been defined as “An analysis to determine all quantities to some level of specificity” (Linder, 1999). Such estimates are difficult to make and even the definition of a good estimate is unclear, as it varies from situation to situation and may include a range of values rather than a finite number of options. However such estimates are necessary in engineering practice when complete accuracy is unnecessary or adequate information is not available. Yet, estimation has not been included in engineering curricula, nor teaching-learning strategies developed specifically for it. Consequently, even graduating engineering students cannot make estimates of physical quantities such as force and energy (Linder, 1999). It is thus important for researchers and educators to provide appropriate learning environments so that engineering students can explicitly learn how to do estimation.

In order to design a learning environment for estimation, either for classroom or a computer-based learning environment, we need to articulate the learning goals, the process and the skills and sub-skills required in this estimation process (Jonassen et al, 1999). This will enable us to include the appropriate learning strategies and necessary scaffolds. While Mahajan (2014) has developed a set of tools to perform approximations and estimation in science and engineering, and Linder (1999) has identified a set of effective actions for solving engineering estimation problems, these methods and actions are not in sufficient for learning design. Paritosh (2007) proposes a model for back-of-the-envelope reasoning, but the set of problems that this applies to are not engineering-specific. One way to identify the learning goals and the required component skills of engineering estimation is to study the estimation process of experienced engineers to understand the cognitive mechanisms they use. To the best of our knowledge, a study identifying the characteristics of the engineering estimation process has not been done. In this paper, we report a study of the estimation process of two experienced engineers as they work on an estimation problem, with the broad goals of understanding the process of engineering estimation and the cognitive mechanisms underlying it.

Related work on expertise

There exist several research studies documenting the characteristics of experts in various domains and the multiple approaches followed to study expertise (Ericsson et al, 2006). For instance, there is sufficient evidence documenting the differences between experts and novices in the solving of physics, chemistry and mathematics problems (Ericsson et al, 2006). While there have been no studies with experienced engineers solving engineering estimation problems, Wankat & Oreovicz (2015) have compiled the expert-novices differences identified in various problem-solving domains and summarized them along the dimensions of problem representation, solution

strategies, monitoring etc. There is ample evidence, however, documenting the characteristics of experts and experienced practitioners in the process of engineering design (for example, Dym et al, 2005; Atman et al, 2007; Cross & Cross, 1998), which often requires estimation as a first step.

In engineering estimation, Linder (1999) compared the performance of experienced engineering practitioners and students on two engineering estimation tasks and observed significant differences. In particular, practitioners were able to provide values of the right order of magnitude for quantities in their domain of expertise and values off by a few orders of magnitudes for quantities outside their domain of expertise. Students, on the other hand, in both cases provided estimates off by several orders of magnitudes. Linder analyzed student solutions and conjectured that these differences were due to the reasons that 1) students do not have a sound understanding of fundamental engineering concepts, much lesser in fact than was expected and 2) students do not relate the estimates they make to their physical significance, do not have reference values for the quantities they are estimating and have difficulties working with units. However, Linder did not offer a model for engineering estimation as done by expert practitioners who obtain order of magnitude estimates.

Theoretical basis

In order to identify the probable cognitive mechanisms that could be at work in engineering estimation, we looked at related work in engineering thinking, design and problem-solving. We identified the critical role of visual thinking in engineering as documented by case studies of engineers (Ferguson, 1994). An example of this is given by Nelson (2012) when he says, “Engineers are visual or non-verbal thinkers in general. Not only do we represent physics in our minds, we are also able rotate static objects to understand them better.” A recent study found that practicing engineers describe visualizing and improving by manipulating materials, mental rehearsal of the physical space, sketching, and doing thought experiments as engineering habits of mind (Lucas, 2014). Unpacking visual thinking, we note that it has the following components – imagination of statics, mental simulation of dynamics, and external representations of procedures, systems or objects, design or analysis.

The role of mental simulation in science and engineering has been extensively studied. Using think-aloud protocols of experts solving problems outside their domains, Clement (2009) argued that “imagistic simulation” (mental simulation) played a role in the thought experiments used by experts and that these simulations generated new knowledge. Similarly, Nersessian (1999) studied the artefacts produced by scientists in the process of developing new concepts and argued that mental simulation is the mechanism by which model-based reasoning produces conceptual change. Hegarty (2004) has a review of research which provides evidence for the use of mental simulation in mechanistic reasoning. The role of mental simulation for uncertainty resolution in engineering design has been discussed in Ball & Christensen (2009) and for creating knowledge and technologies in Nersessian (2009). Research also suggests that gestures result from these mental simulations (Hostetter & Alibali, 2008). Since engineering estimation begins with understanding how a system works (Linder, 1999), we conjecture that experts must be performing mental simulations when they are identifying the working of the unfamiliar problem system in engineering estimation.

Recently the role of representations in engineering has received a lot of attention (Johri et al, 2013). In Moore et al (2013) the authors found that students use multiple representations and translate across representations during a complex modeling task. In Aurigemma et al (2013), the authors found that building external representations serves more purposes than offloading cognitive load in the engineering design process. The theory of distributed cognition suggests that cognition emerges from the interaction between internal and external (environmental) resources because external representations allow processing that is not possible in the mind (Kirsh, 2010). External representation allow actions such as rearrangement, reformulation and sharing as described in Kirsh (2010). In the domain of problem solving, Zhang (1997) showed that external representations are more than merely memory aids and/or stimuli to the internal mind. He argued that the form of the representation determines what information is perceived and how the problem is solved. In mathematical problem solving research, Hegarty & Kozhevnikov (1999) found that using schematic spatial representations rather than pictorial representations improves problem solving performance. Martin and Schwartz (2009) found that experts take the time to create external representations before starting because it improves their overall performance on a medical diagnosis task. We conjecture that experts will use external representations extensively in their estimation process and we will identify from the data how and where they use them.

Methods

The broad research questions guiding this study were,

1. What is the process by which experienced engineers perform estimation?
2. What are the roles of mental simulation and external representation in performing estimation?

Data collection

Problems

The problems given to the experts are below and were chosen after pilot studies, based on their potential to elicit a wide range of problem solving behaviors from the experts. The first problem required estimation based on the structure of an object, while the remaining two required estimation based on function. The problems progressed from simple to complex, and from requiring little to more domain knowledge. Each problem had two versions which were conceptually similar but worded differently as it was conjectured that this would elicit different estimation behaviors from each expert. For example, the expert 2 version of problem 2 is formulated as an evaluation question whereas the expert 1 version is a numerical estimation problem. Therefore it is plausible that rather than estimate the power of the human heart and then compare it to the power required by a wine opener, experts would consider the human heart and wine opener as a system and evaluate whether the heart could drive wine opener. Thus we may observe different problem solving behaviors from both experts.

Table 1: Problems given to the experts

Expert 1	Expert 2
Suppose I told you that the pit spacing on an ordinary CD is 2 μm , would you agree with me? Why/why not?	How far apart are the pits on a CD?
What is the output power of the human heart?	Could a human heart run a wine opener?
The hand cranked radio is for use far from supplies of domestic electricity or batteries. For decent sound performance (say a single 5 W speaker) how heavy would you expect the radio to be?	Consider radios used far from supplies of domestic electricity or batteries. They have to be cranked by hand for them to work. How heavy would such a radio have to be to be heard within a tent at a campsite?

Procedure

Two experienced engineers specializing in electrical engineering were chosen for the study. These experts are faculty members at a premier technology university in India, and have several years of industry experience as well. They have active research programs in their respective areas of research. The study was done separately with each expert, and conducted in a location of the experts' choosing. Each expert was given a sheet with a problem written on it. They were told to write as many details while solving the problem. They were free to use any books or other materials they wanted to consult in solving the problem, including looking up supporting information needed to solve the problem on the Internet on their personal laptop/computer. In permitting the experts to look up the Internet we were trying to simulate a natural work environment as it exists in the engineering workplace where engineers often consulted the Internet for domain related facts and knowledge.

In order to record every action that the expert took towards estimation, the entire session was recorded using two video cameras. The first was focused on the task area (i.e. the sheet of paper and surrounding area on the desk) to capture their sequence of writing and small hand gestures. The second was focused on their face in order to capture facial expressions and large body movements. Their interactions with the computer were captured using the screen capture software CamStudio (<http://camstudio.org/>). The researcher recorded regular unstructured observations while the expert solved the problems, looking out for critical events which would require elaboration in the follow-up interview. The experts were free to solve in their natural mode, silently or talking aloud as they felt comfortable. The researcher didn't interrupt except to offer a new problem sheet. We did not require experts to think aloud as doing this effectively without placing a cognitive load on the solver requires extensive practice which was not possible with the experts. So we interviewed experts immediately after they had completed all problems using a semi-structured interview protocol. In all we had 45 minutes of video with expert 1 and 2 hours and 20 minutes with expert 2.

Data analysis

We followed the method of cognitive ethnography which is based on traditional ethnography but is concerned with identifying how members of a cultural group make meanings (Williams, 2006; Hutchins & Nomura, 2011) by interpreting observed behaviors. The analysis began by creating detailed transcripts using ELAN (<https://tla.mpi.nl/tools/tla-tools/elan>) of the two videos plus the screen capture and the follow-up interviews for each expert. Threading these transcripts together, a single researcher wrote, for each problem, a detailed description of the problem solving process as it happened sequentially in time. From this description, the authors of this study collaboratively abstracted out the stages or phases of the problem solving. Finally we focused on specific episodes which were interesting in the larger context of the study because they allowed the expert to move forward in the

task and identified what were the mental and physical activities that contributed to forward progress during this episode. Specifically, we analyzed the roles that gestures, talk, writing and computer search played in these episodes. The analytical framework evolved as we did multiple passes through the data.

Results

While we have analyzed all the problems solved by each expert, in the interest of space we restrict ourselves in this section to the description of the second problem as solved by each expert.

Workflow of expert 1

Expert 1 (E1) is an academic with three years' experience in academia and eleven years' experience in industry. She spoke out loud intermittently while solving the problems. After reading the problem, E1 almost immediately searched "Flow rate of blood" on the Internet. She scrolled through the results, highlighted two links related to blood velocity, but did not click either. She picked up the pen to write, dropped it and then picked it up again and began writing. Initially she started writing "Pressure = $F \times$ ", but after a while she struck through that and wrote "Power = $\frac{mgh}{t}$ ", then after a pause added another equality " $= \frac{F \times h}{t}$ ". After a long pause, she wrote " $= \frac{P \times A \times h}{t}$ ". After another pause, she searched for "Blood pressure" on the Internet and clicked on the first search result that popped up called "Normal Blood pressure" and read it. She stated that there were two readings given for blood pressure whose meaning she didn't know so she chose a value between the two which is 100 mm Hg and wrote down that value.

Next, she wrote down " $F = P \times A$ ", said "r is the distance - the head that it pushes the blood around" and added " $\times h$ " next to the equation " $F = P \times A$ ". She identified that she knows "pressure" by underlining the value she had written down, "100 mm Hg" and then stated that "A" (area) is probably the cross-section of the two blood vessels. This she estimated to be "2 cm²". Next she said that "h" was hard to determine "because the diameters of the pipes keep changing". She added that it's a closed loop system and went silent for a while. For part of that duration her pen was hovering over the equation " $= \frac{P \times A \times h}{t}$ ". After this she added "Okay flow rate. That's what I need to know" and searched for "Flow rate of the blood from the heart" on the Internet. She clicked on the Wikipedia page titled "Blood Flow" and read the section "Velocity". She noted down the cross-sectional area of the aorta and the blood velocity and calculated flow rate and then power as "flow rate \times pressure". She decided to convert all values into MKS units and for that searched "100mm Hg Pascal" on the Internet. She noted down the value, 13332 N/m² and completed the calculation arriving at the result of 6 Watts. Her problem solving process (as drawn by her during the follow-up interview) is shown in Figure 1.

This problem was followed up with the other version of the same problem "Could a human heart run a wine opener?" E1 began by saying that she was going to consider the work done as the work against friction between the cork and the bottle neck. So she needed to determine this force of friction since work done is "force \times displacement" and she estimated distance to be 2cm. For a while she was silent and then said "Work done is just power \times time". After a brief pause she added that "...given the right contraption it could take forever and still open the cork." She wrote this down and ended. Her solution approach (as drawn by her) is shown in Figure 2.

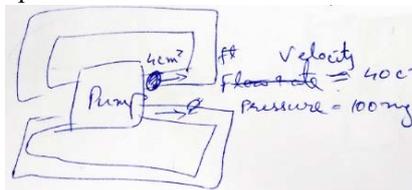


Figure 1. E1's diagram for the heart problem.

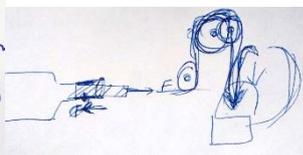


Figure 2. E1's diagram for the follow-up problem.

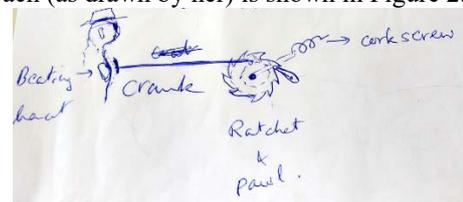


Figure 3. E2's diagram for the heart problem.

Workflow of expert 2

Expert 2 (E2) is an academic with seven years' experience in academia and three years' experience in industry. He worked silently and only spoke to report that he had finished a problem. E2 spent some time reading the problem and after this, while he stared straight ahead silently, the index finger on his left hand moved to and fro a few times. Then he searched for "ratchet" on the Internet, briefly scrolling through results and changing the search term to "to and fro" before returning to the search results for "ratchet". He clicked on the Wikipedia page

for “ratchet” and read the “theory of operation”. While reading, he made a small turning movement with his right hand. Then he wrote down two assumptions, “Assumption 1: It is a beating heart. Assumption 2: It is inside a human body.”

After this E2 spent some time reading the computer screen and then he drew a part of the diagram shown in Figure 3. Next, he air drew what seemed to be a straight line between the man and the ratchet. He formed a “C” with his right hand and rotated it about his wrist. He again drew straight lines and circles in the air. After a while of looking away, he searched for “to and fro motion to rotational motion” and read the first link titled “reciprocating motion”. As he read the screen, he intermittently looked at paper and looked away. Next he searched for “crank machine” and read the Wikipedia link for “Crank (mechanism)”. Then in Figure 3, he drew the straight line from the rectangle to ratchet and labeled it “crank” and completed the rest of the drawing. He drew the flow chart below this diagram to depict his solution approach “Beating heart → Crank (turns gear on ratchet) → Ratchet & pawl → cork-screw. This concluded his solution.

Engineering estimation as a form of model-based reasoning

In this section, we present the answer to our first research question. Based on our analysis of the entire corpus of data, we identified the phases in engineering estimation shown in Figure 4.

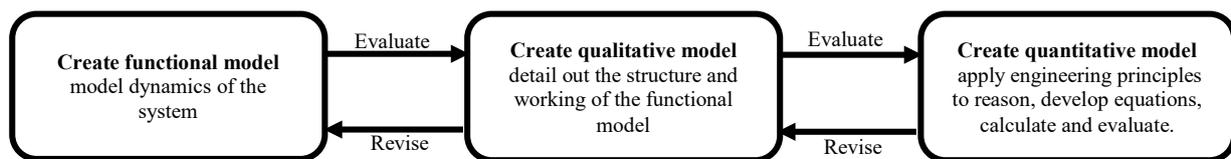


Figure 4. Phases of the estimation process.

Create a functional model

When faced with an unknown system, experts first focused on the dynamics of the system and modeled the dynamics in terms of the dynamics of a known system. In the case of E1 as she reported later “...the first thing that came to my mind was the heart is a pump.” In E2’s case as he reported later he “...thought what does the human heart do that can help me? Which is that it beats. So there’s a rhythmic motion.” In both cases, the experts began by modeling the function of the heart. Its dynamics gave experts a way to identify an object or system with similar dynamics. For E1, the similar object which immediately emerged was the pump and for E2 the heart was reduced to an object that executes repetitive motion. These converted systems that experts began working with were their functional models and their structure was similar to the given system.

Create a qualitative model

Next, experts developed a model of the structure of the system and how its various components work together. We call this the qualitative model. The functional model was constantly evaluated in the mind until it met the requirements of the solution. It was broken down into components to see which component could be modified to get the solution. Thus the final solution was always kept in mind. Since E1 had modeled the heart as a pump, she needed to determine the power of this pump. She identified that the power of a pump is determined by the flow rate and head. Thus her task changed to determining the flow rate and head of this heart “pump”. She was aided in this restructuring by her knowledge and familiarity with pumps due to her recent experience with them. However, when she looked for “flow rate of blood” on the Internet, she thought of “...how much work this pump is accomplishing in this system ... so what does the system look like?” She started to think of the lengths of the veins and arteries, their diameters and the pressure. At this stage E1 fleshed out the details of her model of the heart “pump”. This was her qualitative model.

In E2’s case, because he had modeled the heart as something which moves rhythmically, his task was to find a way to “run” the wine opener (or cork screw as he assumed) using that motion. Thus his problem reduced to converting the rhythmic motion of the heart to the rotational motion of a corkscrew; he changed the nature of the task to coming up with a mechanism for accomplishing the above. Then E2 had to identify the components of a mechanism to convert the beating motion of the heart into the rotation of the corkscrew. He recognized that the heart goes to and fro but he did not want the corkscrew to go both clockwise and anticlockwise, only in one direction. Here he recalled having recently read about the ratchet and that it had something to do with one-way rotation. So he looked it up and decided that it was suitable to the task of turning the corkscrew in one direction. He indicated his partial solution by drawing the heart and the ratchet & pawl. At this point he realized that before the corkscrew could be turned, the linear motion of the heart would need to be converted to rotational motion. As he didn’t know what mechanism could accomplish this, he looked it up on the Internet, learned of the crank and

inserted it into drawing of the mechanism that he had already drawn (Figure 3). Thus E2 also re-examined and restructured his model resulting in his qualitative model.

Create a quantitative model

It is at this third and final stage that experts applied engineering principles and developed (if necessary) a quantitative model or equation corresponding to their qualitative model to calculate the estimate. E1 wrote out the general equation for power and restructured that to arrive at the equation for the power of the heart “pump” in terms of the blood pressure of the heart, namely “Power = $\frac{P \times A \times h}{t}$,” where P is the blood pressure that she looked up. In this new structure, she still did not know “h” in the equation. She evaluated the qualitative model to determine what “h” was; as she reported later, “...basically to what extent is the heart pumping the fluid. So I was trying to think, ok then what are the various diameters of the various arteries and how long are they and all that. But then I was thinking whatever the energy with which it pushes the blood out is expended by the time the blood comes back to the heart.” By re-examining the equation she realized that $\frac{A \times h}{t}$ was actually flow rate and that the distance through which the heart pumps the blood, *h* doesn’t matter. Thus she restructured the equation again and arrived at an equation in which all the quantities’ values could be looked up. She then looked up the standard values on the internet, namely blood pressure and velocity of blood in the veins and completed the estimate.

E2 did not develop a quantitative model of the system or calculate power of the heart to compare it with the power of the corkscrew. Recall that we had expected this to happen due to the wording of the problem given to him. During the follow-up interview, when he was asked to evaluate whether the heart had enough power to turn a corkscrew, he qualitatively reasoned that it probably didn’t. He added that by including two gears – a small one and a large one – which would together turn the corkscrew “very slowly”, he would be able to able the wine bottle, though “it would take forever”. Thus by restructuring his qualitative model he was able to evaluate this alternative scenario and develop another solution.

This process suggests that engineering estimation is an instance of model-based reasoning in which a functional model was iteratively evaluated and fleshed out multiple times, culminating (in the case of a numerical estimation problem) in the calculation and evaluation of the estimate and, if necessary, revision of the model (Figure 4). We also observe that experience or familiarity with certain systems played a critical role in estimation as experts began the process by considering systems from their experience as functional models. In the following, we elaborate on the underlying cognitive mechanisms which support this estimation process.

Cognitive mechanisms underlying engineering estimation

In this section, we answer our second research question and elaborate the roles of mental simulation and external representations in engineering estimation.

Mental simulation

The data shows that when experts read a problem they mentally simulated the dynamics of the problem system, entirely or in part. Some system the expert knew about was used to ‘instantiate’ the simulated dynamics (e.g. heart is a pump). Experts simulated the end point (e.g. the wine opener/corkscrew) or the entire system (e.g. working of the heart) in sufficient detail to evaluate whether their instantiated functional model achieved the desired result. Evidence for this comes from both experts. E1 thought of the heart as a pump and that “it has 4 pipes coming out of it - 2 of which are pumping out and 2 of which are pumping in. So essentially if it has 2 pipes ...I mean, if it’s pushing water out...” In E2’s case, he thought of the heart as “It’s something that is executing repetitive motion.” This simulation helped them to develop their initial functional model of the situation. Further evidence for this mental simulation comes from experts’ gestures. When E1 described the heart, she gestured dynamically with her hands to indicate the flow of water in the pipes, while E2 had been moving his index finger to and fro in the initial phase of the problem solving when he was developing his model. As known from literature (Hostetter & Alibali, 2008), gestures are evidence of mental simulation.

These mental simulations did not stop with functional model building; as experts fleshed out the qualitative model, the functional model was simulated and constantly compared to the desired end-point to ensure that it was still valid. Experts were willing to modify their models if they did not give the desired result. For instance, in the case of E1 while solving problem 3, she initially developed a functional model of a crank and mass attached to it. However, working with the model and re-simulating it, helped her realize that the radio works on electricity and “turning the crank means you are running a generator”, so mass meant the mass of the magnet. Thus constant evaluation of the model led to a breakthrough in problem understanding.

External representations

Diagrams

We found that E1 did not draw diagrams while solving problems. However, she had imagined very clear models for problems 1 and 2 while solving them as was evident from the diagrams she drew when asked during the follow-up interview. We conjecture that because of these clear imagined models she was able to easily restructure and solve these problems. From her very rudimentary diagrams and her verbal reports for problem 3, it appears that she did not have very clear imagined models of this problem, which could have been the reason for her difficulty with this problem, especially because the system had many more components than the previous two problems.

E1 drew diagrams while solving problems 1 and 2. For problem 1, since the question was to estimate pit spacing, by drawing a diagram of his model of the CD he was able to create the other representation required for solving the problem, namely the equations for pit spacing. For problem 2, the final diagram that he drew was his solution. After he drew the first and third parts of his solution, the diagram helped him identify that the solution was incomplete and he needed something in the middle to convert the linear motion to the rotational motion. For problem 3, he did not draw a diagram but a flow chart describing his approach to the problem.

Equations

E1 and E2 both used equations extensively, which is not surprising in engineering. Equations were the way to assign numerical values to physical quantities. However, equations served other purposes besides this in the estimation process. For instance, in problem 2, E1 used equations as an external representation that can be rearranged and reformulated (Kirsh, 2010) and arranged them into a form that was conducive to further action. Equations helped her in mapping the details of her model with the given problem system. Originally she thought that to calculate power she would need to know flow rate and head, but later realized that head was not a valid parameter in this context. Thus working from the basic equation of power ($\frac{mgh}{t}$) she was able to rearrange and reformulate it to a form in which everything was known to her ($\frac{P \times A \times h}{t}$). In the follow-up to problem 2 and in problem 3, she used equations as persistent objects to think with (Kirsh, 2010), as equations helped her in splitting the problem into factors, and in identifying a clear path to the solution.

E2 used equations as persistent objects to think with and for restructuring the problem when he was trying to arrive at an estimate for the weight of the magnet in problem 3. He wrote down a set of equations and then tried to assign approximate values to the physical quantities involved in them in order to determine the volume of the magnet and hence weight. This restructuring of the problem from weight to volume was aided by the equations, which again helped in factorizing the problem and identifying a path to the solution. E2 very often transitioned between text (written by himself and on the computer screen), equations and diagrams in the solving of problems; an instance of this was seen in problem 1 when his pen went back and forth in the air between the diagram and the equation before he wrote; this indicates that he was making a connection between the equation and diagram or using information from one in the other. While this may seem to be obvious to do in engineering, it has been shown that students begin with the equation rather than the model (Wankat & Oreovicz, 2015). While experts used equations to converge their estimation process, which started off with the simulation of dynamics, students may start with the equation, which may not help generate the simulated dynamics or model of the system. Experts used equations to evaluate the simulated model; students may use equations as the only model.

Conclusions and implications

In this study, we have identified a three phase iterative model-based reasoning process for engineering estimation which may be performed in different ways depends on the problem and the solver. We have also identified the roles of mental simulation and external representations in each phase of the estimation process. The process begins with experts simulating the dynamics of the given system and identifying a system with analogous dynamics as a model. The specifics of the initial model may change, especially during the second and third phases, in which it is used to identify, refine and evaluate the structure, working and equations governing the problem system, by constant comparison to the expected behavior from the problem system. These results show us that engineering principles help detail and converge mental simulation and model-based reasoning, and are not themselves generators of solutions. This is different from the classical case of model-based reasoning in science (Nersessian, 1999) in which models are used to infer general principles; in estimation the detailed structure of the models, along with qualitative reasoning and engineering principles, are used to make estimates.

It is interesting to note that while E1 thought hearts pump and E2 thought hearts beat, both arrived at the conclusion that it would take forever to open the bottle “using” the heart, but that it can be done. Their starting dynamics and instantiated models were different, yet their final solutions were conceptually and functionally

similar. The latter two phases contributed to this convergence, and how this happened would be examined in future work. Further, we will repeat this study with two novices and identify the differences with expert performance. The learning environment for estimation will then be guided by learning science principles aimed to reduce the identified differences between experts and novices. If, as we suspect, students do not begin with mental simulation and functional modeling but rather with equations, we propose that the learning environment should require students to work with computer simulations of the given problem systems and model them before writing equations. The exact details will depend on the details of the expert-novice differences.

References

- Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *J. of Engg. Edu.*, 96(4), 359-379.
- Aurigemma, J., Chandrasekharan, S., & Nersessian, N. J. (2013). Turning Experiments into Objects: The Cognitive Processes Involved in the Design of a Lab-on-a-Chip Device. *Journal of Engineering Education* 102(1), 117–140.
- Ball, L. J., & Christensen, B. T. (2009). Analogical reasoning and mental simulation in design: two strategies linked to uncertainty resolution. *Design Studies*, 30(2), 169–186.
- Clement, J. J. (2009). The role of imagistic simulation in scientific thought experiments. *Topics in Cognitive Science*, 1(4), 686–710.
- Cross, N., & Cross, A. C. (1998). Expertise in engineering design. *Research in Engg. Design*, 10(3), 141-149.
- Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering Design Thinking, Teaching, and Learning. *Journal of Engineering Education*, 94(1), 103–120.
- Ericsson, A. K., Charness, N., Feltovitch, P. J., & Hoffman, R. R. (2006). *The Cambridge Handbook of Expertise and Expert Performance*. Psychology.
- Ferguson, E. S., (1994). Engineering and the Mind's Eye. MIT press.
- Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, 8(6), 4–9.
- Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91(4), 684–689.
- Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, 15(3), 495–514.
- Hutchins, E., & Nomura, S. (2011). Collaborative construction of multimodal utterances. In C. Goodwin, J. Streek, & C. Lebaron (Eds.), *Multimodality and human activity: research on human behavior, action, and communication* (pp. 29–43). Cambridge, UK: Cambridge University Press.
- Johri, A., Roth, W., & Olds, B. M. (2013). The Role of Representations in Engineering Practices : Taking a Turn towards Inscriptions. *Journal of Engineering Education*, 102(1), 2–19.
- Jonassen, D. H., Tessmer, M., & Hannum, W. H. (1999). What is Task analysis? In *Task Analysis Methods for Instructional Design* (1st Ed.). Mahwah, NJ: Lawrence Erlbaum.
- Kirsh, D. (2010). Thinking with external representations. *AI and Society*, 25, 441–454.
- Linder, B. M. (1999). *Understanding estimation and its relation to engineering education* (Doctoral dissertation, Massachusetts Institute of Technology).
- Lucas, B., & Hanson, J. (2014). Thinking like an engineer : using engineering habits of mind to redesign engineering education for global competitiveness, In *SEFI 2014: 42nd Annual Conference*.
- Mahajan, S. (2014). *The Art of Insight in Science and Engineering: Mastering Complexity*. The MIT Press.
- Martin, L., & Schwartz, D. L. (2009). Prospective Adaptation in the Use of External Representations. *Cognition and Instruction*, 27(4), 370–400.
- Moore, T. J., Miller, R. L., Lesh, R. A., Stohlmann, M. S., & Rae, Y. (2013). Modeling in Engineering : The Role of Representational Fluency in Students' Conceptual Understanding. *J. Engg. Edu*, 102(1), 141–178.
- Nelson, E. A. (2012). What makes an Engineer an Engineer? In *Structures Congress 2012* (pp. 1152–1159).
- Nersessian, N. J. (1999). Model-based reasoning in conceptual change. In *Model-based reasoning in scientific discovery* (pp. 5-22). Springer US.
- Nersessian, N. J. (2009). How Do Engineering Scientists Think? Model-Based Simulation in Biomedical Engineering Research Laboratories. *Topics in Cognitive Science*, 1, 730–757.
- Paritosh, P. K. (2007). *Back of the Envelope Reasoning for Robust Quantitative Problem Solving*. (Doctoral dissertation, Northwestern University).
- Wankat, P. C., & Oreovicz, F. S. (2015). *Teaching engineering*. Purdue University Press.
- Williams, R. F. (2006, June). Using cognitive ethnography to study instruction. In *Proc. of the 7th International conference on learning sciences* (pp. 838-844). International Society of the Learning Sciences
- Zhang, J. (1997). The Nature of External Representations in Problem Solving. *Cognitive Science*, 21(2), 179–217.

“Doing Double Dutch”: Becoming Attuned to Rhythms of Pathways In and Through Community Spaces

Joyce M. Duckles and Brandon Stroud
joyce.duckles@warner.rochester.edu, bstroud@u.rochester.edu
University of Rochester

George Moses and Ryan Van Alstyne
gmoses@neadrochester.org, rvanalstyne@neadrochester.org
North East Area Development

Abstract: As activists and co-researchers from the university and the Cedarwood community, we have engaged in a long-term participatory ethnography to uncover and inform the transformation of an urban community. Our research focuses on understanding relationships and pathways within and across “hubs” emerging in the neighborhood that lead to authentic and meaningful change. This paper presents our current analytic attention to the boundaries, pathways and spaces between hubs. Adopting the metaphor of “doing Double Dutch” emerging from team analysis, we develop the themes of “pathways in” as opening doors; “pathways through” as feeling the rhythms; and the learning and dexterity that energizes trajectories beyond the hubs in the neighborhood. We propose the constructs of rhythm, attunement and dexterity as conceptual and analytic tools to inform our understandings of pathways and boundaries as “hidden energies” for community development and their generative potential for learning. We conclude that stagnation is not an option.

Keywords: boundaries; community transformation; pathways; rhythms; attunement; dexterity

Rhythm is the soul of life. The whole universe revolves in rhythm. Everything and every human action revolves in rhythm (Babatunde Olatunji, 2000).

Introduction

Educational researchers and activists interested in community transformation have demonstrated increased attention to the generative potential of boundaries as spaces of learning and change. In particular, many scholars have stressed the learning potential of boundary crossing (Akkerman, 2011; Akkerman & Bakker, 2011). Cultural-historical theorists have explored the changing nature of participation over time and across dynamic communities (Gutierrez & Rogoff, 2003) and horizontal learning across activity systems (Engeström, 2001). Boundary crossing has been proposed as a way to introduce new ways of participating in established practices as well as expanding and transforming these practices (Akkerman, 2011). This analytic lens can lead researchers to examine how people “traverse or otherwise connect one environment with another in their everyday lives” and how opportunities to learn are “organized and accomplished through trajectories connecting multiple places” (Leander, Phillips, & Taylor, 2010, p. 331).

Trajectories also imply movement and energy. Leander et al. (2010) apply the metaphors of houses from the critical geography work of Lefebvre (1991) and rhizomes from Deleuze and Guattari (1987) to emphasize the “continual movements and transformations” and connections across places and within spaces (p. 341). Although a house may appear as a stable physical structure, it is actually a nexus of in and out conduits through which energy flows, just as social spaces are produced through energy and movement. Lefebvre encourages us to increase our sensitivities to movement and the “diverse, multiple rhythms of everyday life” (2004, p. 20). In a similar way, rhizomes are connected yet open systems through which nutrients flow. With underground systems of roots, shoots, and nodes, they mirror the ways surface entities that appear distinct have deep subterranean connections. The rhizome metaphor (Deleuze & Guattari, 1987) illustrates the generative yet often hidden energies and capacities across spaces. These theoretical approaches treat structures and places as sets of networked assemblages that are “composed in unfolding activity” (Leander et al., p. 341).

As activists and co-researchers from the university and the Cedarwood community, for the past five years we have engaged in a participatory ethnography aimed to uncover and inform the transformation of an urban community. Our research spans activities across “hubs” that together we have identified, including the Liberty Market (a corner store), The Liberty School Summer Program and Café, Central City Development (CCY) (an active community development office and organization) and the Greenhouse Collaborative (a new initiative of

community gardens and greenhouses). Our overarching goal is to understand relationships and social pathways within and across these hubs that lead to authentic and meaningful change of individuals and community. This paper presents our current analytic attention to the boundaries and spaces between hubs, to the rhythms, pathways, and the generative potential of boundary crossing. It has led us back to our field notes and interviews from the beginning of this long-term ethnography, as we began our research in the hub of the Liberty Market, and our evolving conceptualizations of engagement and belonging in community. As we revisited transcripts and field notes in team meetings with this new lens, a community researcher proposed a new metaphor to capture the movement we are uncovering within and across spaces:

Have you ever done Double Dutch? That's what this reminds me of...these circles are constantly moving. You Double Dutch on the side until you decide to jump in at one of these pathways...then you jump in, you slide, and there you go. [Team Meeting, 10/2015]

Building on this metaphor, we constructed questions to guide our current data collection and analysis: How do community members pick up on, and become attuned to, the varying rhythms of community places and practices? How does this learning lead to engagement in the neighborhood and trajectories beyond the community?

Our interest in transformation and social justice is consistent with the conference theme of transforming learning and empowering learners. The generative nature of boundaries and spaces also leads us to consider how learning environments and community practices are being mutually constituted to enable learners to acquire the kind of learning that empowers them to design their own social futures. This takes place within and across hubs, along pathways and trajectories, and across the community and university. In this paper, we present findings emerging from this new analytic focus. As we explore these spaces, we are learning about “pathways in” as opening doors; “pathways through” as feeling the rhythms; and the learning and dexterity that energizes trajectories beyond the hubs in the neighborhood. We propose the constructs of rhythm, attunement and dexterity as conceptual and analytic tools to inform our understandings of pathways and boundaries as “hidden energies” for community development and their generative potential for learning.

Theoretical framework

We reach across cultural practice and activity theories of learning (Gutiérrez & Rogoff, 2003; Leander, Phillips, & Taylor, 2010) and a rhizomatic model of interdependence and generativity (Deleuze & Guattari, 1987) to frame and guide our interest in spaces and pathways, rhythms and trajectories. Leander et al. (2010) examine the relationship of learning to space and place and, in particular, the relations between learning and movement or mobility. They build on the sociocultural attention to ways learning is distributed across people and places, as well as Lave and Wenger's communities of practice, but also claim that these theories adopt “container narratives” that perpetuate images of bounded spaces. In contrast, they encourage researchers to attend to “how opportunities to learn are organized and accomplished through trajectories connecting multiple places” (p. 331). Cultural practice approaches conceptualize learning as a process occurring within ongoing activity and recognize the diversity of activity settings and practices. Gutiérrez and Rogoff (2003) address some of the challenges of engaging in dynamic practices and shifting activity systems. As people move across systems, they propose that they develop dexterity in navigating practices and participation appropriate to different activity systems.

Our attention to movement has also made us aware of the threats of stagnation. The radical democratic construct of dissensus (Ziareck, 2001) helps us theorize a different boundary crossing; creating spaces to speak across difference and to respectfully challenge differing perspectives of theoreticians and practitioners. This is most evident within our research team, in our daily interactions, where we resist the temptation to insist upon consensus and allow our mutually respected differences to inform authentic debate to energize our research and practices.

Together, these theories provide a framework for considering pathways in and through community spaces and trajectories beyond. They remind us that spaces and practices are shifting and non-linear and that they are situated in “a culture that's constantly being built” (Research team member, 10/2015). We find this focus on movement and change presenting unique methodological challenges.

Methods

Nespor (2006) describes the work of researchers and ethnographers as looking for patterns. He conceptualizes patterns as “descriptions of processes and networks through which things are moving and changing” (p. 300). He points out that the problem is that we, as researchers, are not good at “thinking movement.” Nespor describes the challenge as navigating a “moving dance-floor (historically shaped and propelled) where dissonant orchestras of social relations battle to define the rhythms” (p. 300). To enter into moving and shifting spaces and practices, he

recommends that we take the advice of Latour (1987) and “follow the movement of things.” We are directing our research and theoretical lenses to spaces where practices are in the process of being mutually constituted, where different rhythms are taking shape.

To address this challenge, we adopt grounded theory approaches to ethnographic inquiry (Charmaz & Mitchell, 2001) within a participatory action research (PAR) framework (McIntyre, 2008). In our work, which is by nature iterative in the ways that theory and analysis inform and energize our practice as activists, we are not able to ignore the ethnographic attention to culture and everyday life. We engage the PAR cycle of data collection-analysis-implementation-collection in iterative cycles to build a local evidence base and inform practice. Residents have been trained in research methods and ethical practices as co-researchers, and together with university team members, work as co-authors and co-implementers to address community-identified issues. As we developed our current questions, we returned to transcripts and field notes from across our data set. Our data is coded using grounded theory approaches: themes are iteratively generated and validated through assiduous member checking processes. Adopting the constant comparative method and theoretical sampling (Charmaz, 2014), as we explored the concepts of rhythm and movement in our research team meetings, the metaphor of Double Dutch emerged. We adopt this metaphor to frame our presentation of findings and analysis.

Findings

Nespor (2006) directs us to “note how both people and things change as they move through networks that mesh at some points and generate tensions at others” (p. 301). As we look at what moves, Nespor (2006) recommends challenging the boundaries and looking at how neighborhoods are defined by the intersections of multiple pathways. As we explore these pathways and spaces, we are learning about pathways in as “opening doors”; pathways through as “feeling the rhythms”; and the learning and dexterity that energizes trajectories beyond the hubs of the neighborhood.

Opening doors - Pathways in

As we developed our model of interdependent hubs, our initial concept maps depicted the hubs and pathways but it had the appearance of a closed system (Appendix A). We had focused on the processes and activities, but, as Nespor (2006) noted, we were challenged in seeing the movement and in particular the movement in and out. In the store and across hubs, we recognized the challenge of engaging community members who have traditionally been “dis-engaged,” directing our attention to “pathways in.” Our data suggested that this begins with meeting people “where they’re at” and not judging others:

I think it’s more like, how do we go about reaching the people without it, without them getting offensive. To you talking to them, and then, meeting them where they’re at. Trying to find out what’s going on, to just not always being judgmental. [Greenhouse Collaborative Meeting, 2/2015]

We realized that as community activists we needed to attend to what “meeting people where they’re at” looks like, and how it can be facilitated. As we were analyzing data, a co-researcher suggested that it begins with opening the doors wider:

So when we were talking about the guys on the street corners and the dealers and stuff, but making pathways available for them as well? You can’t judge them because that’s where they are right now, but instead you have to connect with them and have pathways for them...And some of them won’t. We may not get all of them, but if we get one or two to come along we’re making progress. So, that’s harder work, going from feeling totally disengaged to entering the work that you want to do. How do we pull them along or at least make pathways for them? Cause it can’t always be them that are there and ready to move. Sometimes you’ve got to make the door wider open some way or another. [Team meeting, 9/2015]

Parents and community developers talked often of making pathways for everyone who “wants to come along” and opening the doors in intentional ways. Our data has many examples of an apprenticeship model across hubs with attention to providing first time experiences to youth and adults. Through the Work Experience Program (WEP), adults have come to work in the market, the CCY office, and the Liberty School Café. Youth are also trained at the market and as “junior servant leaders” in the Liberty School. A community member described these interactions as “reshaping what interactions look like in the community.” Processes of building relationships and

ways of interacting echo across the hubs – the consistency is intentional. Rather than accepting silos, the connections between spaces ensure the consistencies as individuals move along the pathways.

This means that “pathways in” can occur through any hub. At a team meeting, a co-researcher pointed to the lines we were drawing between hubs on our concept map:

That’s what I would say about these lines. It’s intentional. It’s organic but it’s intentional. It’s intentional for them to come to any one of these hubs and to be able to organically, to be exposed to each of them.

Several years later, this was articulated through the Double Dutch metaphor. We recognized the circles or hubs as constantly moving. And then,

You Double Dutch on the side, until you decide to jump in at one of these pathways. Whether it’s the market, the school, university, or garden, you know, you just Double Dutch. Then you jump in, and you slide, and there you go.

Feeling the rhythms - Pathways through and between hubs

Opening the doors is the beginning. A co-researcher described how participants enter on pathways through and between the hubs. It begins with listening and using ideas shared by residents:

So that they know, hey, I am relevant, I am important, I am a part of it. Then through that, it allows them to spread the word or the message to others. But when you don’t make them feel welcome, you don’t make them feel like their voices are heard, they’re going to continue down the same roads that they’ve been. It’s just the bottom line. [Team meeting, 10/2015]

Feeling the rhythms means feeling part of something bigger. A team member described this as people “becoming conscious or aware and feeling, you know I’m a part, because the conversations that we now have, they participate in, where before they didn’t” [Team meeting, 10/2015]. They are now “at the table.” They are part of the “Freedom Movement.”

The Double Dutch metaphor becomes powerful here for illustrating the challenges of feeling the rhythms across hubs. The Double Dutch conversation continued in our meeting, focusing on “getting the rhythm” so that you can jump in. We noted that you get better at “jumping in” over time, that you can’t be “stagnant” because once you’re in you have to be a part of it. CCY’s director asked if we all knew how to “twirl,” noting that everyone can’t jump in, but most people know how to twirl. He extended it further:

So that metaphor is, you just, what you’re doing is you get into rhythm. So as we start to talk about energy it’s attunement. Cause we’re using static in different ways - static being still or static where tuning in stations. So you’re tuning until you get to the right station. So it’s energy. So what you’re trying to do is get on the right energy level...so that’s Double Dutch. You have that pathway to get in, but you have to Double Dutch to get in, and once you get in you have to have the right energy to get in rhythm with everything else. [Team meeting, 10/2015]

The pathways through and between are also co-constructed in intentional ways. In our early renditions of our model, we began to recognize the hubs and the consistency of processes within them. We spoke of “challenging the norms” that have come to be accepted in the community by building rhizomatic connections between spaces that echo similar messages, practices, and expectations. Again, the rhythms of different hubs were noticed:

All components at some point tie in on a different level, whether it be social, economical, health, educational...so there’s like a constant interjection between them. So even though it may look static-y, it may be static, but at any frequency there’s always going to be static until you fine tune it. It’s the fine tuning of all the components together that makes you come to a strong frequency to push out there. [Team meeting, 9/2015]

Pushing “out there” means moving beyond, extending the movement beyond the neighborhood.

Sharing the rhythm and developing dexterity – Pathways beyond

Part of moving forward is sharing the rhythm and “passing it forward”: “You reach one, you teach them, and then you let them reach someone else.” Keeping up the momentum and sharing the rhythm is described as central to the “Freedom Movement.” This was illustrated during a team meeting using our model of rhizomes:

HC: I know irises grow in a rhizome and that’s how they spread. I mean I don’t know if this is an issue, but after a while if you don’t cut them up and replant them the center will die.

LJ: So what does that mean for our data?

DJ: Well, you’re never going to stop, you have to keep changing and developing and if you’ve got entrepreneurial ventures that become their own nodes, they are building on the same principles, same processes that we’ve been talking about, and then they move to another community.

RG: Yeah, and what does that mean for this work?

LJ: So they can become centers of their own.

RG: There you go. [Team meeting, 4/2013]

Another “pathway beyond” takes shape through developing the skills or the “dexterity” to enter different spaces with different rhythms. Stories of community residents who are “making it” are shared often. Abby began as a WEP worker in the store. She worked in the café and received her safe food certification, and she now has a job downtown. One researcher noted that Abby could have kept working her WEP hours, but “now she’s in the mindset of I want something of my own.” Community members, both adults and children, navigate the multiple rhythms of the different hubs in the community, learning how to “jump in” and participate. Similar to the repertoires for participating in practices proposed by Gutiérrez and Rogoff (2003), we can encourage people “to develop dexterity in determining which approach from their repertoire is appropriate under which circumstances” (p. 22) or across varying activity settings. We propose that following participants like Abby can add to our understanding of “how engagement in shared and dynamic practices of different communities contributes to individual learning and development” (Gutiérrez & Rogoff, 2003, p. 21). We also argue that developing this dexterity is part of the learning that can empower people to design their own social futures.

Within our research team, we also became aware of how work at boundaries generates learning. Akkerman and Bakker (2011) identified transformation as a recognized mechanism for constituting the learning potential of boundary crossing. In this framework, moving or transgressing from expected roles or modes of interpretation can open up spaces for the construction of new ways of being and thinking. In our roles as university and community researchers, we entered this emergent zone of contact and this joint work at the boundaries. We use the concept of dissensus (Ziarek, 2007) to illuminate the ways that collaborative research can foster generative tensions at the boundaries. A doctoral student on our research team summarized this tension in a memo:

We do create shared narratives. We also embrace the continuous joint work we are crafting at the boundaries – our unique frameworks, experiences, identities and practices shape the dialogic spaces we co-construct; lead us to embrace dissensus and recognize the generative potential it holds for us as a research team. [Memo, 5/2012]

Whether in the store or in our team meetings, friction and dissensus emerging from boundaries ground our “constant move forward.” As we work together toward urban transformation, stagnation is not an option

Conclusion and implications

The theoretical lens framing our current analysis encouraged us to treat structures and places as networked assemblages that are “composed in unfolding activity” (Leander et al., p. 341). The themes of boundaries, movement and rhythms have permeated our data from the very beginning of this long-term ethnography. We adopted the theoretical metaphor of rhizomes (Deleuze & Guattari, 1987) to bring attention to pathways and their transformational potential. However, we had been conceptualizing the interdependent hubs as a closed system of relations. Returning to our data and devoting research team meetings to this analysis created a dialogic space to uncover what is happening along the pathways connecting hubs and at multiple boundaries. Metaphors of Double Dutch and notions of “attunement” emerged from these meetings, providing conceptual tools to add to our understandings of the pathways and spaces, their rhythms and “hidden energies” for community development and

their generative potential for learning. This metaphor extended our attention to pathways in and pathways beyond, reminding us that the systems are open.

As social activists and researchers, we align with the conference goal of uncovering ways that we can empower learners to design their own social futures. In our analysis, this emerged as “pathways beyond.” Building on the concept of repertoires of practice developed by Gutiérrez and Rogoff (2003), we propose that examining the dexterity to feel and navigate the rhythms of different spaces suggests a new way for us to explore how people are learning. Our findings mirror notions of learning as increased participation in valued practices of a community, but also provide insights into how this learning can move with and across spaces, and how it might empower learners to extend across and beyond participation in the hubs in the neighborhood.

Our analysis is also pointing to challenges and barriers, to processes that constrain access or challenge the momentum toward transformation. As we focus on movement and dynamics, we also need to consider cohesion or stasis. As a community member stated, “Stagnation is not an option.” We proposed “pathways in” taking shape as opening doors by creating spaces for voices and we recognized that the intentional redundancy across hubs was providing multiple entry points. However, we know that some community members do not “feel a part of” the changes under construction in Cedarwood. We are also interested in why some may choose to stay out, and the physical and psychological barriers that persist. Exploring historically closed doors can help us challenge simplistic views of access to resources based on physical accessibility, especially as we consider “pathways in” to the shifting rhythms of community spaces.

Consistent with the conference theme, our goal as researchers and practitioners is to uncover development and learning across spaces in community, and in particular the kinds of learning that might empower community members to engage with and design their own futures. As we have uncovered the generative potential of working the boundaries and pathways across hubs in the community, we have also explored ways to make the connections and rhythms more intentional and explicit. Recently, our “visual connections” initiative has emerged to create a visual representation of the pathways across hubs of activity. A series of mural and community arts projects are being created that link these pathways through powerful images, spoken word, and digital media (#FreedomSankofa, Figures 1). Our research continues to guide our practice as we work toward engaging community in imagining pathways and creating spaces that foster connections across their histories, within their community, and toward their futures.



Figure 1. #FreedomSankofa

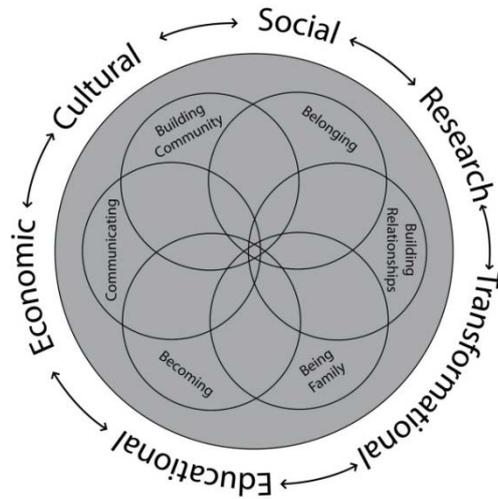
References

- Akkerman, S.F. (2011). Learning at boundaries. *International Journal of Educational Research*, 50, 21-25. doi: 10.1016/j.ijer.2011.04.005
- Akkerman, S.F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, 81(2), 132-169. doi: 10.3102/0034654311404435
- Charmaz, K. (2014) *Constructing grounded theory: A practical guide through qualitative analysis* (2nd ed.). London, UK: Sage
- Charmaz, K., & Mitchell, R. G. (2001). Grounded theory in ethnography. In P. Atkinson, A. Coffey, S. Delmont, J. Lofland, & L. Lofland (Eds.), *Handbook of ethnography* (pp. 160–174). Thousand Oaks, CA: Sage.
- Deleuze, G, & Guattari, F. (1987). *A thousand plateaus: Capitalism and schizophrenia*. Minneapolis, MN: University of Minnesota Press.
- Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 33(1), 107-123.
- Gee, J.P. (2008). A sociocultural perspective on opportunity to learn. In P.A. Moss, D.C. Pullin, J.P. Gee, E.H. Haertel, & L.J. Young, (Eds.), *Assessment, equity, and opportunity to learn* (pp. 76-108). New York, NY: Cambridge University Press.
- Gutiérrez, K., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher*, 32(5), 19-25. doi: 10.3102/0013189X032005019
- Latour (1987). *Science in action*. Cambridge, MA: Harvard University Press.

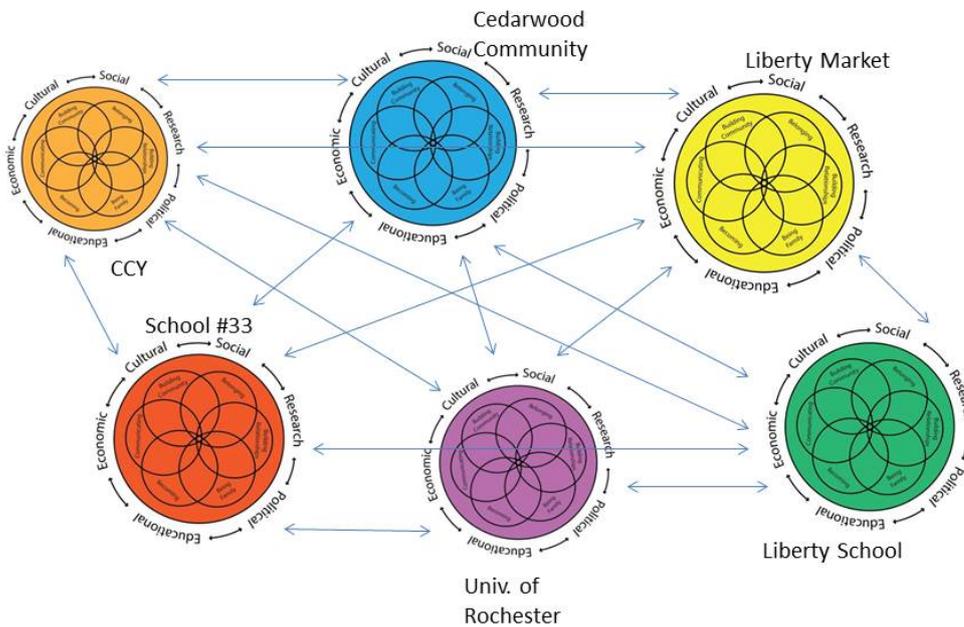
- Leander, K.M., Phillips, N.C., & Taylor, K.H. (2010). The changing social spaces of learning: Mapping new mobilities. *Review of Research in Education*, 34, 329-394. doi: 10.3102/0091732X09358129
- Lefebvre, H. (2004). *Rhythmanalysis: Space, Time and Everyday Life*. New York, NY: Continuum.
- McIntyre, A. (2008). *Participatory action research*. London, UK: Sage. doi: 10.4135/9781483385679
- Nespor, J. (2006). Finding patterns with field notes. In J. Green, C Camilli, P. Elmore, A. Skukauskaite, & E. Grace (Eds.), *Handbook of complementary methods in educational research* (pp. 297-308). Mahwah, NJ: Lawrence Erlbaum.
- Ziarek, E.P. (2001). *An ethics of dissensus: Postmodernity, feminism, and the politics of radical democracy*. Stanford, CA: Stanford University Press.

Appendix A:

Model of Interdependent Processes & Transformational Pathways*



Interdependent Hubs & Transformational Pathways



*As presented at:

Duckles, J.M., Moses, G., Moses, R., Gallegos-Greenwich, J., & Stroud, B. (2015, February). Co-constructing active and collective pathways of family and community engagement: A community-based approach. Data Analysis Session. 36th Annual Ethnography in Education Research Forum: Inequality, Poverty and Education: An ethnographic invitation. University of Pennsylvania, Philadelphia, PA.

Towards a Framework of Pedagogical Paradoxes: A Phenomenographic Study of Teachers Designing Learning Experiences and Environments With ICT in Singapore Classrooms

Wan Ying Tay, Boon Yen Chai, Nur Azarina Khamis, Samuel Tan, and Chew Lee Teo,
hotaywanying@gmail.com, CHAI_Boon_Yen@moe.edu.sg, Nur_Azarina_KHAMIS@moe.gov.sg,
Samuel_TAN@moe.gov.sg, TEO_Cheew_Lee@moe.gov.sg
Ministry of Education, Singapore

Abstract: This paper discusses how teachers reconcile tensions as they design learning experiences and environments with Information & Communication Technologies (ICT) in Singapore classrooms. Through applying a paradoxical lens, we describe conflicting demands and contradictory considerations that teachers face. Adopting a phenomenographical approach, we examine the narratives of 2 teacher participants documented in pre- and post- lesson conferences, classroom enactments and interviews, so as to tease out variations in the ways they experienced and mediated the paradoxical tensions. Based on the findings of these tensions, a pedagogical paradoxes framework is presented, illustrating the interplay between the different dimensions of the paradoxes. By surfacing these paradoxes, we hope to draw awareness to the challenges and reconciliation in the decision-making and rationalisation process that teachers as learning designers may experience. A heightened awareness can enable teachers to consciously negotiate the paradoxes in designing different kinds of learning experiences better tailored for their students.

Introduction

Teachers are increasingly encouraged to leverage on the affordances of information and communication technologies (ICT) for designing learning experiences and learning environments for their students, with the aim of enhancing their learning (e.g., Beetham & Sharpe, 2013; Laurillard, 2012; Luckin, 2010).

In the case of Singapore, there are also keen efforts to encourage such practices. For example, in a programme aimed at spreading good instructional practices and supporting schools in developing and sustaining student-centric pedagogies with technology, Educational Technology Division, Ministry of Education (Singapore), partnered schools in deepening classroom learning by involving teachers from participating schools in designing learning experiences and environments. Besides participating in networked learning communities and working with teachers from other schools to develop their design, teachers also worked closely with a team of officers from ETD (ETOs) and a Research Partner (RP) from an Institute of Higher Learning (IHL). In the process, they participated in pre- and post- lesson discussions and engaged in reflection.

Review of the support process as well as analysis of transcripts of the discussions and teacher interviews suggests that in the course of designing learning experiences and learning environments for their students, teachers often came face to face with paradoxical tensions. While our 2 teacher participants might not be immediately aware of these tensions during their design, enactment and reflection process, they would, as our findings suggest, consciously or unconsciously, attempt to manage and reconcile these tensions.

Understanding what these paradoxical tensions may be and how best to reconcile and cope with them is important as they could influence classroom practice and consequently, student learning (Boling & Beatty, 2012). As Lewis (2000) aptly points out, by intentionally bringing these tensions to the fore and engaging in discussion about them, could help individuals better understand and negotiate them.

We first provide a review of the paradoxical tensions literature, discussing the nature of paradoxes as a constitutive feature. Guided by our findings, we address how a pedagogical paradoxes framework can help support teachers in reconciling these paradoxes as they approach designing learning experiences and environments for their students.

Literature Review

Understanding pedagogical paradoxes

In this paper, a paradoxical view is applied to help identify pedagogical paradoxes that are inherent in the learning design. It also helps us examine and describe the dynamic interrelationship between the each of the paradoxical

dimensions (Papachroni, Heracleous, & Paroutis, 2014). We focus primarily on pedagogical paradoxes which involve examining teaching and classroom practices, from the teachers' point of view.

In a framework proposed by Lewis (2000), he identifies the root causes of paradox, and how individuals react to them as they try to deal with and reconcile these paradoxes. This framework emphasises that paradoxes, while conflicting in its nature, should not be perceived as opposing, polarised elements; rather, they are "contradictory, yet interrelated" (p. 3). Thus when confronted with paradoxical tensions, individuals need not make a clear and distinct choice between two poles of the paradox (Eisenhardt, 2000; Westenholtz, 1993). By viewing paradoxes as dynamically interrelated or even complementary elements, enables us to examine how teachers would try to reconcile or cope with them.

While paradoxes have mostly been described and discussed in organizational management literature, they have been deliberated upon by educators and scholars, one of them being Dewey (cited in Prawat, 1999). Dewey referred to them as epistemological duality (rather than paradox). He observed that there could be more than one duality, and that each duality may be nested upon another. In other words, it is not unusual to find that a teacher may have to approach various paradoxes all at once. Furthermore, we are also very much informed by literature exploring dilemmas and challenges that teachers face (e.g., Ball, 1993; Luehmann, 2008). Thus, above or beneath each single paradoxical dimension, we may find concentric circles or layers of other dimensions. Paradoxical perspective has also been adopted to examine instructional practices (e.g., Tay & Ng, 2014; Warschauer, 2007).

The central thesis of this paper is that, in teaching, the nature of paradoxes in classroom practice is often not made explicit and a teacher may or may not be aware of them, even as they make decisions based on these paradoxes. By surfacing these paradoxical tensions, it helps to bring greater awareness the challenges and considerations that a learning designer may encounter. The greater awareness enable teachers to consciously think through and mediate these tensions in deciding on the kind of learning interactions and learning activities or tasks that are appropriate for the learning experience. This might enable the teacher better negotiate the challenges and decisions in evolving different kinds of learning experiences better suited to the needs of the students.

Guided by our review of the literature, we focused on two main research aims: (1) to describe the different ways in which teachers experience and perceive tensions as they approach learning design; (2) to describe the different ways in which they attempt to reconcile these tensions.

Adopting a phenomenographic approach

So as to describe the variations in the ways teachers experience and perceive tensions as they approach learning design and the ways in which they attempt to cope with and reconcile these tensions, we adopted phenomenographic approach as the guiding methodology for our study.

Phenomenography is the study of the different ways in which people experience, conceptualise and understand a phenomenon (Marton & Booth, 1997, Marton, 1981). It aims to draw a relation between the subject and the phenomenon. Notably, in a phenomenographic study, the investigation is not directed at a phenomenon as such, but at the variation in the ways individuals conceptualise, experience and comprehend the phenomenon. For the purpose of our study, which is to describe and examine the variation in which teacher participants experience paradoxical tensions while designing and enacting learning experiences and environments for their students, phenomenography serves well as a guiding methodology. It had allowed us to tease out the variations as teachers reflect into their teaching-learning experiences, diagnose the tensions they experienced in their learning design process.

Our study was conducted over a span of a year. During the year, teachers planned, designed and implemented their lesson design, with the support from the ETD officers (ETOs) and a research partner (RP) from an IHL. The data collected in this study include pre- and post- lesson conferences between teachers, ETOs and RP, lesson observations captured on video and audio recordings as well as semi-structured interviews conducted with the teachers. These interviews, which were about an hour long each, were conducted at the end of the study. Based on the preliminary findings from the pre- and post- lesson conferences and lesson observations, we mapped out the paradoxical tensions that emerged from the analysis of the data sets. During the interviews, we also requested for our two teachers to map out how they felt the lesson design was intended, i.e., whether they felt that it was tending towards (1) meeting long-term objectives, such as developing 21st century competencies, subject-based literacy and inquiry skills in their students, or short-term objectives, such as delivery of the curriculum and preparing their students for assessment; (2) engaging students in personal sense-making activities or collaborative, collective sense-making activities; (3) guiding/facilitating students in their learning (teacher control), or giving students the agency/autonomy in their learning; (4) having flexibility in the lesson design, or following as close as possible to what was being designed (stability).

We digitally recorded and transcribed verbatim the interviews with the teachers. Interview questions were designed in a semi-structured fashion, and the key questions aimed at teasing out their experiences of the tensions that emerged from implementing their lessons. These key questions included: What were the insights/problems you and your students experience when implementing the lessons? Do you think that the way that the lessons had been designed have made any difference to how your students learn? What were the challenges you faced when incorporating ICT into your lesson and how did you overcome them? In keeping with the phenomenographic methodology, analysis of the various data sets was done through constant comparison and iteration, which was carried out until no additional tensions were surfaced. This analytical procedure was conducted by each of the five co-authors of this paper and the analysis was cross-checked against one another.

Description of the participants

Our participants were 2 Geography teachers from a secondary school in Singapore. They participated in this study as the schools in which they are teaching at were involved in the ETD programme (as mentioned in the introduction section). Teacher 1 is a very experienced Senior Teacher and has been teaching Geography for 26 years. As a senior teacher, she participated in the programme as she felt the need to constantly learn so that she could develop her fellow younger colleagues to the next level. While she has not been using ICT pervasively in her lessons, she is very open in learning new things and is adequately proficient in IT. Teacher 2 is also another very experienced Geography teacher, having taught for more than 15 years. She is the school's Head of Department (HoD) for ICT. She saw the benefits of having ICT-enhanced lessons for both teachers and students.

Description of the lessons

The 2 teachers (Teacher 1 and Teacher 2) were involved in teaching the same topic on stratovolcanos in their respective classes. The pre-lesson conferences involved ETOs and RP sharing a lesson that leveraged ICT to facilitate students' learning about the characteristics of stratovolcano through GI approach that was designed and enacted in other secondary schools. Much focus was on how the design allowed students spaces for investigation, data gathering and the teaching - student interaction that facilitated the exploring of students ideas to further the inquiry-based learning. Both teacher 1 and teacher 2 raised questions and considerations about meeting the requirement of the curriculum, while at the same time, providing more flexibility in the lesson design and allowing more student agency and autonomy in monitoring and assessing their own learning. Finally, the teachers evolved a lesson experience that was ready to be enacted in their respective classrooms.

Teacher 1 was getting students to learn about the characteristics of stratovolcanoes of evidence-based inference through the process of geographical inquiry. The lesson required students to look through learning resources consisting of video recordings of various volcanoes as well as authentic data made available through the online application, ArcGIS. Students would then make sense of the data gathered which included the locality, shape, height, gradient as well as eruption video clips of the various volcanoes identified as learning resources and synthesised the data to arrive at a consolidated understanding of the characters that define a stratovolcano. The lesson was also planned as a group learning endeavour with each group of students working in two teams - one team would work on seeking and finding information such as gradient and shape, while the other looked at video online and identified information such as eruption patterns and locality. The whole learning experience was planned as a lesson of a single one-hour duration. However, in the actual enactment, the lesson took longer than planned and spanned a total of two hours (over two one-hour sessions). Teacher 1 shared that the lesson took longer than planned as there were initial unfamiliarity among students on how to work with the tools that were available to identify and extract useful information. The main challenge that the teacher faced was that of guiding her students to look at the correct set of data and focusing on what was relevant to the pre-planned learning outcomes rather than exploring other learning opportunities that emerged.

Teacher 2, a colleague in the same team, who was present during the design and enactment of Teacher 1's lesson, saw the value of the learning approach that the team was attempting. However, she felt too that it was too time-consuming and she also noticed there were also too many perspectives being explored. She then made the decision to refine the process by getting the same student to do both parts of data gathering and instead of working on a spread of data over 6-8 volcanoes, each student only looked at 2 -3 specific volcanoes. These volcanoes were commonly selected and allocated by the teacher. Teacher 2 then took the students through a structured inquiry with her playing a more central role in modelling and guiding students on how to analysis, understand and interpret the data. She guided students on thinking through what the patterns observed meant and how it linked to the concepts and big ideas. This, she opined, allowed her to shorten the learning process and freed up time for her to teach students the epistemic skills associated with the correct way of expressing their ideas and understanding in geographical language.

Findings and discussion

This section reports findings from the analysis of pre- and post- lesson discussion amongst teachers, ETOs and RP, as well as interview transcripts and observations of classroom enactment of the lesson experiences. The findings reveal dimensions of paradoxical tensions that were present as teachers approached their design of learning experiences and learning environments for their students. The findings also suggest the various ways in which teachers could negotiate and mediate the tensions, taking into consideration their competencies, orientation and personal experiences.

Dimensions of tensions experienced by teachers when designing learning experiences and environments with ICT

Based on the analysis of the pre- and post- lesson conferences and lesson enactment, we propose a pedagogical paradoxes framework. The illustration of the pedagogical dimensions is shown in Figure 1 and the description of each of the dimensions is given below.

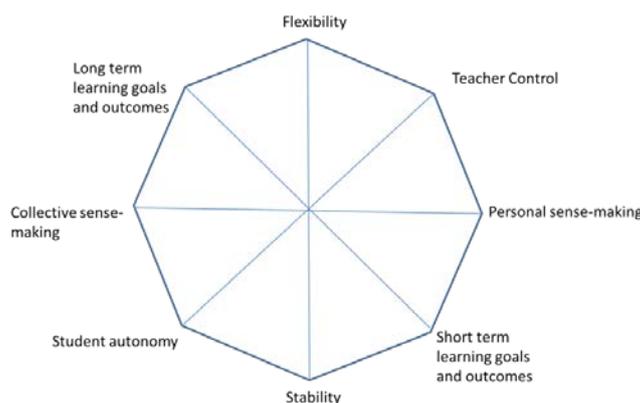


Figure 1. Illustration of the pedagogical paradoxes framework.

Collective sense-making and personal sense-making

This illustrates the teacher's decision in the nature of learning pedagogies to be adopted in a lesson. At one end of the spectrum is the view and design that the process of learning being very much a personal experience, with the individual taking the primary responsibility and making sense of information and knowledge that are generated. i.e., the learning activity is primarily designed as individual activity which involves cognitive learning processes. This dimension aligns well with existing literature, particularly those that explore the nature of learning in classrooms (e.g., Cook, Smagorinsky, Fry, Konopak, & Moore, 2002).

At the other end, learning is seen as a collective responsibility in a lesson, with teachers and students as a learning community taking collective responsibility in advancing the common understanding by working together to make sense of information and knowledge that are generated during the learning interactions. i.e., the learning activity is primarily designed as collaborative activity which involves socio-constructive learning processes.

Teacher control and student agency

This shows the thinking of whether the learning experience is one where the teacher takes centre stage, designing the learning experience to what the teacher is most comfortable with, being the authority and source of information, designing learning tasks/activities as one where teachers need to "teach" so that students "learn", having to make deliberate effort to articulate specific information and knowledge to students to ensure their learning. Students come ready to be instructed.

At the other end is the thinking of letting students profile and ideas lead the learning. Designing the learning experience to leverage on students ability, interest and skills with teacher playing the role of learning orchestrator and facilitator stimulating inquiry, supporting sense-making and scaffolding inductive and deductive learning. Teachers and students are partners in learning. Teachers act as the drivers of the scenario and lead the collective learning activities. The tensions that were surfaced at this dimension align with existing literature, particularly those that investigate how teachers balance teacher control and student autonomy (e.g., Drexler, 2010).

Stability (fixed) and flexibility (adaptive)

On one end of the spectrum is stability. This involves designing learning experiences or environments that include a series of learning tasks and activities that are conducted in a more sequential manner with the content and interactions happening according to prescribed/planned steps. The learning tasks/activities lead to specific pre-planned learning outcomes.

On the other end of the spectrum is flexibility. This involves designing learning experiences or environments where teacher-student, student-student and student-content interactions lead the direction of learning. Teachers (and students to a lesser extent) have the possibility to change the learning scenario, activities and content on the fly as the sense-making process takes place in order to optimise the learning (opportunities) that takes place and maximise the achieved learning outcome.

Short-term learning goals/outcomes and long-term learning goals/outcomes

For short-term learning goals, learning is about covering curriculum. Often, it would involve solving well defined problem, meeting fixed outcome and representation, focusing on factual knowledge and information transmission, and on practicing of or training up a certain skill.

For long-term learning goals, learning is to develop cognitive and meta-cognitive (and/or affective) competencies. Thus, this may involve designing activities which may engage students in solving authentic tasks/ill-structured problem, learning tasks that have multiple outcomes and representations, contextualised authentic application and/or extension of learning. Thus, it focuses on conceptual learning & process skills.

Mapping of the paradoxes

During the interviews with the 2 teachers, they were asked to map the paradoxical tensions they felt best describe their lessons. Teacher 1's mapping of the pedagogical paradoxes of her lesson enactment is illustrated in Figure 2.

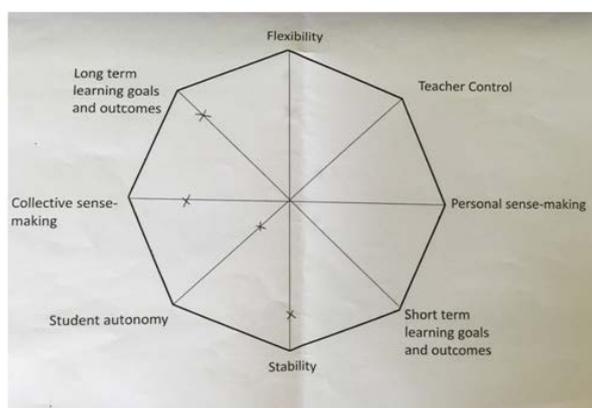


Figure 2. Teacher 1's mapping of pedagogical paradoxes of her lesson.

As could be observed from the mapping in Figure 2, Teacher 1's conception and the ways in which she negotiated the pedagogical paradoxes differed significantly in her two lesson enactments. In the interview, Teacher 1 shared that she felt quite "useless" during the first lesson as she was not doing her job as a teacher but merely facilitating the activities. Though learning was done through group discussions, Teacher 1 felt learning were still very much a personal enterprise as students learnt more through personal sense-making. She also felt the demands of having to constantly adapt to the learning that emerged in different groups of students. The demands, she felt, required a teacher to have a very strong and deep content and pedagogical knowledge.

For Teacher 2, as could be observed from her mapping of the pedagogical paradoxes in Figure 3 below, that she felt she had to negotiate in her lesson, she experienced and negotiated pedagogical paradoxes differently from Teacher 1. In explaining the thinking behind the design and enactment of her lesson, Teacher 2 observed Teacher 1's lesson that students were struggling with group work and were not able to understand how the data they gathered had to work together to help them inferred the required knowledge in relation to the big ideas. Hence, she decided to give each student fewer but more concrete examples so that each student could follow her guidance on looking at the data presented and linked these examples to the concepts they were to learn. Thus, her lesson was focused on providing a structured inquiry with her (as the teacher) explaining key concepts based on the data that was shared with the students.

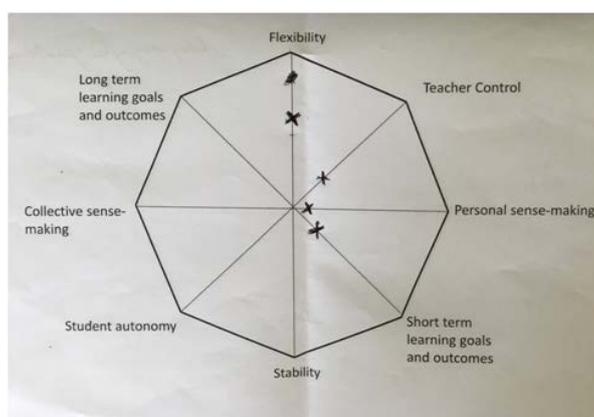


Figure 3 – Teacher 2’s mapping of pedagogical paradoxes of her lesson

Our interviews with the two teachers also suggest their encounters with the four dimensions of the pedagogical paradoxes. These are described with some of the interview excerpts in Table 1:

Table 1: Dimensions of tensions as experienced by teachers

Extract from transcripts	Dimensions of Tension
<p>Teacher 1: “Personal sense-making. Because they’re all given the free will to think of various ways. So actually it’s more about personal sense-making.... Because the software itself is a platform but the way the lesson is conducted has to be group work. It can’t be one to one.</p> <p>Teacher 2: “So the whole structuring of the lesson needs to be carefully thought out, and it’s not like we didn’t think it out in the beginning, the idea was there...different groups come together. But in essence, there are a lot of technical issues also la, so it takes up a lot of time.”</p>	<p>Collective Sense-making and Personal Sense-making</p>
<p>Teacher 1: “... I want to know the evidence of learning from my students. Like knowledge or skill... The knowledge target and the skill target. I must be like clear. Then once I have that, I will plan my lesson based on the evidence of learning. So... erm, what do I want to look at? There are so many things, I mean... depending on the topic.”</p> <p>Teacher 2: “So we have to use this kind of learning judiciously, we can’t, much as I would like to do it in everything, design such exciting lessons all the time. I can’t because of the limited time that we need to ensure that syllabus is completed. In terms of curriculum and syllabus, you must be very clear about the learning outcome. The next thing is you must be very clear about the skills, the skills that the kids should have. So you must know the concepts and content, followed by the skills.”</p>	<p>Long Term Learning Goals and Outcomes, and Short Term Learning Goals and Outcomes</p>
<p>Teacher 1: “Students will be able to come out with the answer very directly but the thing is it depends on the example, the data that we plant in. I realised that the volcanoes for example, there is some ambiguity, because the example that we plant in is not so clear-cut (and) not so straightforward. There is this blur area.”</p>	<p>Flexibility and Stability</p>

<p>Teacher 2: “...it took on its own life when you give kids a voice in expressing what they have discovered, through their own exploration of the various volcanoes they around the world.”</p>	
<p>Teacher 1: “... I feel that my role is just helping them to start the computer and giving them the instructions and after that they are on their own. But later on, following that, the data is so called collected, then the role of the teacher comes in. It requires a different set of questioning technique to actually make them see the patterns and to be able to make connections.”</p> <p>Teacher 2: “...it’s more the strategic questioning and the skill of the teacher to question the child in such a way that it guides the child’s thinking and illicit out...but I feel that if a teacher isn’t that experienced and that strong in the content, she might not know the techniques of strategic questioning, so it’s very critical.”</p>	<p>Teacher control and Student agency</p>

As could be concluded from the findings above, teachers may come encounter paradoxical tensions when designing learning experiences and environments for their students. Their conceptions of the tensions varied depending on their years of teaching experience, position in the school and the ways in which they perceived the roles that teachers and students play.

Conclusions and implications

Through illuminating and bringing to forth the different dimensions of pedagogical paradoxes of learning design, our study seeks to highlight the interplay of the challenges and considerations that teachers are confronted with. The pedagogical paradoxes framework that we present serves as an initial concept of the paradoxical tensions. It does not claim to be exhaustive and we note that there may be subsequent finer layers or concentric circles above or below these dimensions. Further study of other paradoxes or other dimensions of paradoxes could provide deeper depth to our understanding of the tensions and decision pivot that teachers as learning designers may go through. It is hoped that by surfacing these paradoxes, it brings greater awareness of the various conflicting demands and contradictory tensions that teachers may experience. Accordingly, having such an understanding of the tensions could also help in developing professional development programmes or support for both pre- and in-service teachers. For instance, it could serve as a focal point for discussion on key decisions to be taken in designing a learning experience.

References

- Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal*, 93, 373-397.
- Boling, E. C., & Beatty, J. (2012). Overcoming the tensions and challenges of technology integration: How can we best support our teachers? In Ronau, R., Rakes, C., & Niess, M. (Eds.), *Educational Technology, Teacher Knowledge, and Classroom Impact: A Research Handbook on Frameworks and Approaches* (pp. 136-156). Hershey, PA: IGI Global.
- Beetham, H. & Sharpe, R. (2013) *Rethinking Pedagogy for a Digital Age: Designing for 21st Century Learning*, New York, NY: Routledge.
- Drexler, W. (2010). The networked student model for construction of personal learning environments: Balancing teacher control and student autonomy. *Australasian Journal of Educational Technology*, 26(3), 369-385.
- Eisenhardt, K. M. 2000. Paradox, spirals, ambivalence: The new language of change and pluralism. *Academy of Management Review*, 25, 703-705.
- Laurillard, D. (2012). *Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology*. New York, NY: Routledge.
- Lewis, M. W. (2000). ‘Exploring paradox: toward a more comprehensive guide’. *Academy of Management Review*, 25, 760-76.

- Luckin, R. (2010) *Redesigning Learning Contexts: Technology-Rich, Learner-Centred Ecologies*, New York, NY: Routledge.
- Marton, F. (1981) Phenomenography - describing conceptions of the world around us. *Instructional Science*, 10, 12-33.
- Marton, F. and Booth, S. (1997) *Learning and awareness*. Mahwah, N.J: Lawrence Erlbaum.
- Papachroni, A., Heracleous, L., and Paroutis, S. (2014). Organizational Ambidexterity Through the Lens of Paradox Theory: Building a Novel Research Agenda. *The Journal of Applied Behavioral Science*.
- Prawat, R. S. (1999). Dewey, Peirce, and the learning paradox. *American Educational Research Journal*, 36(1), 47-76.
- Tay, W. Y. & Ng, Z. Y. (2014). Paradoxical tensions of online peer assessment: A case study of a secondary school in Singapore. In Chen, G., Kumar, V., Kinshuk, Huang, R., & Kong, S. C. (Eds.), *Emerging issues in smart learning*. New York, NY: Springer Berlin Heidelberg.
- Warschauer, M. (2007). The paradoxical future of digital learning. *Learning Inquiry*, 1(1), 41-49.
- Westenholz, A. (1993). Paradoxical Thinking and Change in the Frames of Reference. *Organization Studies*, 14, 37-58.

Acknowledgments

We thank the two teachers from the participating school, in kindly supporting us in advancing our understanding of pedagogical paradoxes in learning design.

Joint Idea-Building in Online Collaborative Group Discussions

Yann Shiou Ong, The Pennsylvania State University, yannshiou@psu.edu
Marcela Borge, The Pennsylvania State University, mborge@psu.edu

Abstract: This paper defines the construct of joint idea-building in collaborative learning situations, and proposes a hierarchy of sophistication levels of joint idea-building activity based on the frequency of idea-building moves and the pervasiveness of extended idea-building moves. Criteria for the sophistication levels are informed by qualitative analysis of sampled collaborative online group discussions among college students tasked with discussing their responses to course-related questions. Our findings will be used to revise our existing assessment rubric for evaluating the sophistication level of joint idea-building activity in the context of online group discussions, as well as inform future work for improving the validity of our rubric.

Keywords: joint idea-building, assessment rubric, collaborative learning, online discussions

Introduction

Collaborative learning is a common theme among various research areas, including social, cognitive, developmental and educational psychology, the learning sciences, instructional design, and computer-supported collaborative learning (O'Donnell & Hmelo-Silver, 2013). One of the main group processes studied in collaborative learning is knowledge co-construction (Webb & Palincsar, 1996). Through collaboration, learners co-construct knowledge they did not have prior to the collaboration (Damon & Phelps, 1989).

In our previous study (Borge, Ong, & Rosé, 2015), we designed a rubric to evaluate the quality of collaborative processes in online group discussions of course-related questions. We encountered the challenge of defining the construct of *joint idea-building*, our conceptualization of knowledge co-construction, in a way that could help us distinguish between high and low performing collaborative groups. We consider the activity of joint idea-building as the extent to which learners act upon ideas introduced to the group by elaborating or extending these initial ideas. We examined the literature as a means to guide the development of communication coding frameworks and assessment tools to provide needed support for students. In doing so, we have found that existing literature provides limited elaboration of what joint idea-building activity looks like concretely. Much work has been done on other conceptualizations of knowledge co-construction, such as knowledge building (Scardamalia & Bereiter, 2006; Zhang, Scardamalia, Reeve, & Messina, 2009) and transactivity (Joshi & Rosé, 2007; Weinberger & Fischer, 2006), which we will discuss. However, after careful deliberation, we consider both constructs different than joint idea-building.

This paper aims to contribute to the existing literature and existing methods for evaluating collaborative activity by providing a detailed description of the joint idea-building construct. This includes a proposed hierarchy of sophistication levels of joint idea-building activity based on the frequency of idea-building moves and the pervasiveness of extended idea-building moves. The criteria defining the levels of sophistication are identified through qualitative analysis of a sample of six teams' (dyads and triads) online group discussions that are part of an online undergraduate level course. Our findings from the empirical data will be used to revise our existing rubric for evaluating joint idea-building activity as well as inform future work for improving the validity of our rubric.

Related work and ideas

Different conceptualizations of knowledge co-construction

There are two prominent conceptualizations of knowledge co-construction in the literature, knowledge building and transactivity. According to Scardamalia and Bereiter (2003), knowledge building emphasizes the production and advancement of public knowledge among a community of knowledge builders. Thus, the building of ideas, from the perspective of the knowledge-building framework, is at the level of the community, where the goal is to advance the knowledge of the community. On the other hand, transactivity focuses on how individual learners act upon each other's ideas at the level of turns of discussion and how they execute their collective reasoning processes (Berkowitz & Gibbs, 1983; Teasley, 1997). As Joshi & Rosé (2007) explain, transactivity combines the notion of idea extension and elaboration with aspects of argumentation and knowledge negotiation.

Both of these perspectives on knowledge co-construction differ from our conceptualization of joint idea-building. Our conceptualization of joint idea-building is at the level of an episode nested in a discussion within a small group. A conceptual goal of these discussions is for learners to foster deeper understanding of ideas prior to critiquing by developing and building the ideas jointly. An idea can be built through: (1) elaboration, such as providing details, using analogies and examples, or highlighting pros and cons; or (2) extension, such as connecting the idea to other ideas, generating a related idea, generalizing the idea, or synthesizing ideas.

Challenges in defining and evaluating the sophistication of joint idea-building activity

Our approach to evaluating the quality of joint idea-building arose from an attempt to help groups reflect on important markers of collaborative activity in order to improve upon them over time. For this reason, we needed a way to help collaborative learners recognize key aspects of high quality collaboration and assess the level of sophistication that their team exhibited around these aspects. To identify important processes, we used the existing literature to create a list of important behavioral markers, of which joint idea-building is one (Borge et al., 2015). We then had to unpack each of these markers into a scale of less and more sophisticated processes so that students could compare their existing processes to desired processes. Though we managed to reach a high level of inter-rater reliability with our assessment rubric, we still faced conceptual challenges in determining what counts as an idea-building move. Two questions that were difficult for us to resolve were: (1) should an idea-building episode include moves associated with knowledge negotiation, such as challenges to an idea and supports for the idea in response to the challenges; and (2) should questions posed by learners be counted as an idea-building move?

We acknowledge that more sophisticated collective cognition requires both idea-building (as part of information synthesis) and knowledge negotiation to co-occur (Stahl, 2006). However, we perceive idea-building as distinct from knowledge negotiation (Borge et al., 2015). Information synthesis is a collective, cognitive activity whereby group members work to share, collect, make sense of, and connect individual information into a cohesive, collective whole. On the other hand, knowledge negotiation is a collective cognitive activity associated with decision-making where alternative ideas are presented, acknowledged, and evaluated by the group (Borge et al., 2015).

Though these two collective cognitive activities can overlap in time, we argue that they should be evaluated independently from each other because it is important to examine the extent to which an idea is understood and developed by a team before the team begins to make decisions about it. We support this position with existing findings related to group processes: that how teams synthesize ideas from multiple members can affect the quality of collective decision-making processes (Borge & Carroll, 2014), that in the presence of competing ideas, teams have a tendency to reject or ignore ideas, even crucial ideas, because they fail to monitor and regulate activity to ensure that ideas are explored (Barron, 2003), that diverse ideas are essential to innovative thinking but only to the extent that these ideas are considered, explored, and synthesized by the group (West, 2007), and that tendencies to pick ideas without the development of shared understanding and equal consideration can lead to problems with interpretation, decision-making, learning, and performance outcomes (Barron, 2003; Borge & Carroll, 2014; Carroll, Borge, & Shih, 2013). Thus, while not every idea may need to be built upon, given what we know about group processes, it is likely that higher quality collaborative processes ensue when there is evidence that groups take time to extend and explore at least some ideas before moving on to a competing idea. Such behaviors may be indicative of a group's desire to understand alternative perspectives rather than simply debate them; this is a quality Stahl (2006) argues as necessary for high quality collaboration. Following this logic, we maintain that a question can act as a prompt to extend an idea-building episode (Berkowitz & Gibbs, 1983), thus it could count as an idea-building move. However, questions that center on introducing a competing idea, i.e., "Yes, but have you considered that idea X might be better?", are conceptualized as knowledge negotiation moves and therefore not counted as idea-building.

In our existing assessment rubric, the sophistication of a team's joint idea-building activity is evaluated on a score of "1" (least sophisticated) to "5" (most sophisticated) (Borge et al., 2015). The scores are based on two markers: frequency of idea-building moves and the pervasiveness of extended idea-building moves. Frequency of idea-building moves is categorized as: *extended idea-building* (five or more idea-building moves following an initial idea for a discussion topic), *limited idea-building* (one to four idea-building moves following an initial idea), and *no idea-building* (no idea-building moves following an initial idea). A score of "5" means a discussion includes at least two episodes of extended idea-building; a score of "4" means a discussion has one episode of extended idea-building; a score of "3" includes at least two episodes of limited idea-building; a score of "2" includes one episode of limited idea-building; a score of "1" means no idea-building occurred in the discussion. Criteria for the number of idea-building moves (measure for frequency of idea-building moves) and number of episodes (measure for pervasiveness of extended idea-building moves) were determined based on the instructor's estimate of reasonable expectations from an average performing team (assumed to receive a score of

“3”) and a high performing group (assumed to receive a score of “5”). For example, each member in a team of two or three members can reasonably be expected to contribute one or two idea-building moves, thus defining the criteria for an average performing team.

Many colleagues have rightly criticized our current approach, as we too recognize it may not be valid. However, our current measures provide us with a starting point from which to iteratively develop better ways of helping students and instructors make sense of and improve collective thinking processes. To iteratively improve the rubric, we are interested in empirically determining the abovementioned criteria based on empirical evidence from our existing data, which is the basis for the study reported in this paper.

Study aim and research questions

The broader aim of this study is to contribute to the existing literature by unpacking what joint idea-building activity looks like and providing descriptions of more and less sophisticated forms of this activity. Our findings will also be used to inform the criteria used to assess joint idea-building in our existing assessment rubric. Our aims lead to the research question: to what extent is the definition of sophistication levels of joint idea-building activity in our existing rubric aligned with empirical findings from sampled online group discussions?

Methods

Context and data sample

The study reported in this paper is the third iteration of a broader study by Borge et al. (2015). Participants in our study are students of a 16-week university level introductory online course on information sciences and technology. Students were introduced to concepts and research in disciplines including security and risk analysis, human computer interaction, emerging technologies, effects of technology on society, and informatics. Group discussion activities contributed 25% of students’ total grade, as development of collaborative reasoning practices was a course learning outcome. Students were required to complete five online group discussion activities throughout the course. Each activity required students to individually complete a pre-discussion activity after reviewing the discussion session materials (e.g., chapters from the course required text or supplementary materials). They then participated in an online group discussion with assigned team members (in dyads or triads) on a professional collaborative workspace with chat and document sharing features. Each discussion lasted approximately 45 minutes. The discussions were assessed using our existing rubric for collaborative discussion quality, which is based on the core capacities of information synthesis and knowledge negotiation; joint idea-building is an aspect of information synthesis (Borge et al., 2015).

The pre-discussion activity is designed to deepen students’ thinking about their required course readings. The activity consists five questions labeled as *topics* in our analysis: (Topic A) what part or concept in the reading was most difficult to understand and why?; (Topic B) what is one interesting argument about a concept in the reading made by the author and its implications for individuals and society?; (Topic C) what is the most important concept you have read and why?; (Topic D) which concepts discussed in the reading are most likely to impact your work, and how can they be applied to your current or future job?; and (Topic E) if you could ask the author one question regarding the reading, what would it be?. Students could decide the reading questions they wish to discuss during their online group discussions. They could also discuss other related topics of interest.

For our analysis, we sampled six teams’ online discussions from discussion session three. These discussions take the form of time-stamped posts identified by individual users who made the post. We selected session three because this session has the largest variation in idea-building scores, based on our existing rubric. This ensures we have examples of teams with varied performance and more likely varied patterns of idea-building. The six teams were sampled based on their joint idea-building score on the existing rubric: Teams 4 and 14 scored “3”; Teams 11 and 13 scored “4”; Teams 2 and 12 scored “5”. Furthermore, the sample comprises a mix of dyads and triads: Teams 2, 13 and 14 are dyads; Teams 4, 11, and 12 are triads.

Data analysis

Identifying idea-building episodes and moves

In establishing our existing assessment rubric, the research team held extensive discussions to reach agreement on how to determine the “boundary” of an idea-building episode and what is an appropriate unit of analysis for an idea-building *move*. The team agreed that idea-building episodes exist within the discussion of a topic. A topic is typically introduced to the team discussion by a member stating or suggesting which reading question to discuss (e.g., “Let’s start with question two.” or “Shall we go with question three first?”) A topic ends when a member pushes for the team to discuss another question (e.g., “let’s move on to the next question”) or a new topic is

introduced (e.g., “Ok, question four”). Several idea-building episodes may exist within a topic. An idea-building episode begins when a member puts forward an initial idea within a topic that is a response to the reading question. For example, a member’s response “The most difficult concept for me was the task-artifact cycle” is an initial idea to Topic A (what part or concept in the reading was most difficult to understand and why?). A different response to Topic A, “I found the messy desktop part most difficult” signals the start of a different idea-building episode within Topic A. An idea-building episode ends when the idea ceases to be built and a new topic is introduced. Exceptions are made when a subsequent post within the boundary of the new topic builds on an idea in a previous topic due to a time lag in making the posts (indicated by identical time-stamps or short time differences between time-stamps). Following the convention of conversation analysis, we define the unit of analysis as a turn made by a student (Sacks, Schegloff, & Jefferson, 1974). A student’s turn begins when she makes a new post and ends when another student makes a post. Thus, an idea-building move can comprise a single post or multiple concurrent posts by the same student.

Coding for idea-building moves

We coded for four main moves in our sampled discussion sessions. *Initiating a topic* (Topic) refers to a move where a main discussion topic is introduced, typically by stating a reading question or any question of interest. *Putting forward an initial idea* (Idea) is a move where an idea is put forward by a member in response to a Topic. Initial ideas are not counted as idea-building moves, but they initiate the opportunity for idea-building. Idea-building moves begin when the team adds on to an initial idea. These moves are described as follows. *Adding to an idea* (Add) is when a member adds to an existing idea or parts of it by elaborating or extending it. *Questioning* (Qn) refers to posing a question in relation to a previously introduced idea to seek further ideas or information. It excludes clarifications that do not seek to elicit further ideas. If a move includes both adding to an idea and questioning, the code is assigned based on the more significant move. For instance, if a student gives an example related to a previous idea (Add) and asks the team for their thoughts about the example (Qn), the move is coded as Qn since his question is potentially more significant in extending the idea-building episode. We identify such moves using a micro-coding framework with substantial inter-rater reliability, $\kappa = 0.74$, $p < 0.001$.

In addition, we wanted to identify when ideas evolve during an episode, as students may build on ideas related to an initial idea rather than focusing on the initial idea itself. Thus, branch ideas are indicated in the following situations: (1) ideas based off a previous idea, Idea 1, are labeled as Idea 1.1, Idea 1.2, etc.; and (2) initial responses to a question, Qn1 that seeks further ideas about Idea 1, are labeled as Idea 1.1., Idea 1.2, etc. Labeling for further branching of ideas is possible (Idea 1.1.1, Idea 1.1.2, etc.). Branch ideas count towards idea-building moves of the main, initial idea. For example, Idea 1.1 and Idea 1.2 are both part of idea-building moves for Idea 1.

The six sampled discussion sessions were coded independently by the first author and a student researcher. Both coders have experience evaluating the teams using the existing assessment rubric, and have participated in the aforementioned research team discussions about the idea-building construct. The coders met to discuss their independent codes to resolve any differences in the codes and arrive at a consensus set of codes, including where significant deviation from the main idea occurs and the labels of branch ideas should be used.

Statistical analysis of idea-building episodes

To answer our research question, we turn to the relevant descriptive statistics of the idea-building episodes that serve as measures of the two markers for joint idea-building activity in our existing rubric. Specifically, we are interested in the average number of idea-building moves per episode, which defines the categories for extended idea-building and limited idea-building, and serves as a measure for the frequency of idea-building moves. An idea-building episode comprises an initial idea and the additional idea-building moves following it. If an initial idea is not followed by other idea-building moves, then no idea-building occurs and it is not an idea-building episode. Non idea-building episodes were excluded from the statistics since we are interested in the average number of idea-building moves among episodes demonstrating idea-building activity. We also want to find the average number of episodes meeting the criteria for extended idea-building (to define a score of “5”) and for limited idea-building (to define a score of “3”), which is a measure for the pervasiveness of extended idea-building moves. These statistics will be compared with the criteria for the sophistication levels of joint idea-building activity in our existing rubric.

Visualizations of idea-building episodes

A free online mind mapping tool WiseMapping Version 3.0.2 (2014) was used to generate visual representations of the idea-building episodes based on each group’s discussion topics. Individual idea-building moves (Add, Qn,

or branch ideas) are represented as blocks connected to an initial idea (Idea). All initial ideas are connected to a topic (Topic). The colors of the blocks represent individual students who made the move.

Findings

Descriptive statistics of idea-building episodes

Among the six sampled group discussions, a total of 25 topics were discussed (excluding the discussion of one topic which only comprises knowledge negotiation activity). These topics resulted in a total of 55 initial ideas, 35 of which developed into idea-building episodes. Eighteen initial ideas were not built upon, while two ideas led only to knowledge negotiation activity. On average, each team discussed 4.2 topics and put forward 9.2 initial ideas, which resulted in 5.8 idea-building episodes. The average number of moves per episode is 4.2 (standard deviation = 3.1). We propose an episode with above average number of moves i.e. five or more moves can be considered an extended idea-building episode. The differences between extended idea-building, limited idea-building, and no idea-building episodes based on the frequency of idea-building moves are visually represented in Table 1 using examples from the sampled discussions.

Table 1: Definitions and visualizations of idea-building episodes based on frequency of idea-building moves. Discussion of Topic D includes three initial ideas, each with different frequencies of idea-building moves. Idea 3 demonstrates an *extended idea-building* episode, where an initial idea is built extensively over seven moves, including Qn and Add moves, and branch ideas (Idea 3.1 and Idea 3.2). Idea 1 shows a *limited idea-building* episode, where an initial idea is built over two Add moves. Idea 2 is an example of *no idea-building*.

Frequency of Idea-building Moves	Description	Visualization of Idea-building Episodes in a Topic Note: Each block represents an idea-building move. Blocks of the same color represent moves made by the same member.
Extended idea-building (Idea 3)	An initial idea is built over an extended number of moves (five or more moves)	
Limited idea-building (Idea 1)	An initial idea is built over several moves (one to four moves).	
No idea-building (Idea 2)	Initial ideas are shared but not built.	

To exemplify the difference between an extended idea-building episode and a limited idea-building episode, shown in the visualization in Table 1, we present an excerpt from Team 11's discussion on Topic D about the impact of a concept on the members' work.

Member	Message Content	Move
Chloe:	Which concepts discussed in the reading are most likely to impact your work? How can you apply the concept in your current or future job?	Topic D
Chloe:	The concept that I believe has already impacted my job is HCI in general. Especially the Task-Artifact concept. When I click on something on my desktop I expect it to open said object, just as I would expect and elevator door to open when I push the elevator button.	Idea 1

Daniel:	The evolution of HCI has a huge impact on my work. It might not fundamentally change what I do every day, but how I do it and the timeliness of it are both impacted by it.	Idea 2
Daniel:	I agree Chloe. How can any work not be impacted by HCI is some way?	Idea 1: Add
Chloe:	I honestly cannot say I have an answer to that question. I admit my job would suck without it and pretty much all everyday activities that involve technology.	Idea 1: Add
Brandon:	For me the area the concept that will impact me is the utilization of HCI to assist in social media & network analysis	Idea 3
Chloe:	Elaborate on this for me if you can Brandon.	Idea 3: Qn
Brandon:	I think as I observed through every lesson there needs to be an even balance between technology and the users, hence smart-watches allow an excellent balance.	Idea 3: Add
Chloe:	I'm not sold on smartwatches at all.	Idea 3.1
Daniel:	I love my smartwatch.	Idea 3.2
Chloe:	Mostly because they don't provide me with anything that my other devices cannot. However it has been rumored that Apple is working on developing an Apple car. If the Apple watch could be used as remote start then I could see a need. In fact if any watch had that option I could see a need.	Idea 3.1: Add
Daniel:	It allows me to "check" my phone without me then finding myself in a rabbit hole.	Idea 3.2: Add
Brandon:	I think HCI will cater [to] certain preferences and observations that the desired user wants and assists in their needs for social media.	Idea 3: Add

Of the three ideas presented by the members, Idea 3 is the one the team builds upon most extensively through a variety of moves. The team members request for further elaboration of the idea (Idea 3: Qn), provide more details and connect it to a real-life example (Idea 3: Add), share their personal opinions (Idea 3.1 and Idea 3.2), provide rationale for their opinions (Idea 3.1: Add and Idea 3.2: Add), then return to the original idea by synthesizing their perspectives (Idea 3: Add). In contrast, Idea 1 is extended to a limited extent by members generalizing the idea through a rhetorical question (Daniel's Idea 1: Add move) and extending the idea to other aspects of life (Chloe's Idea 1: Add move).

To determine the number of extended idea-building episodes (based on our proposed five or more idea-building turns) we can expect from a team demonstrating fairly sophisticated joint idea-building activity (i.e. teams with score of "4" or "5"), we consider the average number of extended idea-building episodes among the sampled discussions, which is 1.8 episodes. In contrast, the average number of limited idea-building episodes (based on our proposed one to four idea-building turns) is 4.0 episodes. Hence, we suggest that sophisticated joint idea-building activity (score of "5") comprises at least three extended idea-building episodes (higher than the average number of extended idea-building episodes), that is, five or more idea-building moves in at least three episodes. An average level of sophistication in joint idea-building activity (score of "3") comprises at least four limited idea-building episodes, that is, one to four idea-building moves in at least four episodes.

Discussion

Our group is working to develop concrete ways of evaluating collaborative activity in order to help students develop more sophisticated collective thinking processes. We also want to find ways to articulate learning theory and empirical findings into usable learning objects, such as assessments, that students and instructors can use to guide their socio-metacognitive development. This poses challenges for unpacking abstract constructs such as joint idea-building. Our existing rubric served as a starting point to evaluate joint idea-building, but we wanted to draw on real student collaborative activity to further inform our conceptualizations. Specifically, we wanted to compare the logically reasoned quantification of markers we currently use to determine sophistication of joint idea-building to actual quantification of those markers in real student activity.

Based on our findings, we observe that the criteria for number of moves defining an extended idea-building episode and a limited idea-building episode in our existing rubric concur with the empirical findings from the sampled discussions. However, the number of idea-building episodes expected for the levels in the existing rubric is lower than what students actually produced in the sampled discussions. Hence, we propose the following revisions to the criteria for joint idea-building scores in our existing rubric (revisions in **bold**): A score of “5” means a discussion includes at least **three** episodes of extended idea-building (five or more idea-building moves); a score of “4” means a discussion has **one to two** episodes of extended idea-building; a score of “3” includes **at least four episodes** of limited idea-building (one to four idea-building moves); a score of “2” includes **one to three episodes** of limited idea-building; a score of “1” means no idea-building occurred in the discussion.

Our work is always limited by the number of teams that can be evaluated. We recognize that the six teams we microanalyzed may not be representative of the population as a whole. However, we believe that the depth of our analysis can contribute to new ways of thinking about communication patterns and can inform larger quantitative studies, especially those working to develop automated methods for assessing collaboration quality. Our findings serve as anchors for future work, evaluation, and discussion of collaboration quality.

Conclusions and implications

In this paper, we conceptualized the notion of joint idea-building as being defined by two markers, frequency of idea-building moves and the pervasiveness of extended idea-building moves. Our findings from actual student online discussions provided evidence to inform modifications to our existing rubric for joint idea-building, and also suggest directions for future work.

In view of the proposed modifications to our existing rubric, a follow-up study is necessary to look at how the revised rubric impacts the sophistication of joint idea-building activity in future online group discussions. We are mindful about how the revision may affect students’ behavior, especially since our current instructional practice includes providing students with the same rubric as a means to help them evaluate and reflect on their own processes. While we want to modify students’ cognitive behavior towards more sophisticated joint idea-building, we also want to avoid students gaming the system by staging their discussions to achieve the maximum score with minimum performance, such as stopping their idea-building activity once the required number of idea-building moves are met as per the requirements stipulated in the rubric.

Our initial analysis of students’ existing idea-building episodes suggest the existence of other potentially important markers, such as the extent to which members engage in multiple turn-taking to build on an idea, and the number of initial ideas put forward for a discussion topic. To improve the validity of our rubric, future work will examine the extent to which our existing markers sufficiently capture the full range of idea-building sophistication, and whether other markers should be included. There is also a need for future work to qualitatively analyze the products of joint idea-building activity and evaluate the depth and richness of the ideas generated and synthesized, so as to determine whether the markers we have identified are related to the sophistication of the products of joint idea-building activity.

In order to support students to engage in more sophisticated joint idea-building activity, there is a need for future work to investigate the conditions that lead to more sophisticated idea-building activity. One possible condition is the type of discussion activity questions. Are there some questions that lend themselves to more sophisticated idea-building than others? Another possible condition is strategies for managing the discussions. Will focusing on a few topics instead of discussing all topics lead to more sophisticated idea-building? What about discussing one initial idea for a topic before moving on to another initial idea versus putting forward all initial ideas then selecting one to build on? Understanding of such conditions can help instructors make informed modifications to the discussion activity to promote more sophisticated joint idea-building activity among collaborative teams.

Through sustained work in these highlighted directions, we will continue to unpack the notion of idea-building, refine our existing conceptualization, and revise the cognitive tools we provide in our CSCL technologies to articulate theory as a means to help students develop socio-metacognitive expertise.

References

- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12(3), 307–359.
- Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly*, 29(4), 399–410.
- Borge, M., & Carroll, J. M. (2014). Verbal Equity, Cognitive Specialization, and Performance. In *Proceedings of the 18th International Conference on Supporting Group Work - GROUP* (pp. 215–225).
- Borge, M., Ong, Y., & Rosé, C. (2015). Design models to Support the Development of High Quality Collaborative Reasoning in Online Settings. To be included in the proceedings of the International Conference of Computer Supported Collaborative Learning (CSCL) 2015.
- Carroll, J. M., Borge, M., & Shih, S.-I. (2013). Cognitive artifacts as a window on design. *Journal of Visual Languages & Computing*, 24(4), 248–261.
- Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research*, 13(1), 9–19.
- Joshi, M., & Rosé, C. (2007). Using transactivity in conversation for summarization of educational dialogue. In *Proceedings of the Speech and Language Technology in Education (SLaTE) Workshop 2007* (pp. 53–56). Farmington, PA. Retrieved from http://www.isca-speech.org/archive_open/slate_2007/sle7_053.html
- O'Donnell, A. M., & Hmelo-Silver, C. E. (2013). What is collaborative learning?: An overview. In C. E. Hmelo-Silver, C. A. Chinn, C. K. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 1–16). New York, NY: Routledge.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In *Encyclopedia of Education*. (2nd ed., pp. 1370–1373). Macmillan Reference.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (pp. 97–115). New York: Cambridge University Press.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
- Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaboration? In L. B. Resnick, R. Salijo, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools, and reasoning: Essays on situated cognition* (pp. 361–383). Berlin: Springer.
- Webb, N., & Palincsar, A. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (3rd ed., pp. 841–873). New York, NY: Macmillan.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers and Education*, 46(1), 71–95.
- West, G. P. (2007). Collective cognition: When entrepreneurial teams, not individuals, make decisions. *Entrepreneurship: Theory and Practice*, 31(1), 77–102.
- WiseMapping. (2014). Retrieved from <https://app.wisemapping.com/c/maps/>
- Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge building communities. *Journal of the Learning Sciences*, 18, 7–44.

Acknowledgments

We would like to thank our undergraduate research assistant, Emily Hanson, for her contributions to this project, as well as other members of the research team for their constructive feedback on this paper. We would also like to thank the participating students for allowing us to examine their interactions. This project was supported by the National Science Foundation (IIS-1319445).

How Socio-Cognitive Information Affects Individual Study Decisions

Lenka Schnaubert, University of Duisburg-Essen, lenka.schnaubert@uni-due.de
Daniel Bodemer, University of Duisburg-Essen, bodemer@uni-due.de

Abstract: Metacognitive self-regulation theories assume that individual monitoring guides study decisions. However, self-regulated online learning is not done in isolation and inherently social. Group awareness research suggests that socio-cognitive information may be a strong asset to collaborative and individual learning. Integrating individual research traditions into a social setting, our experimental study ($N = 61$) investigates how visualizing socio-cognitive information influences core individual learning processes, especially the search for information, the learners' self-evaluations of knowledge and learning outcomes. While on the surface study behaviour seemed not to be affected by the availability of socio-cognitive information, more profound analyses revealed that learners provided with partner information did rely less heavily on initial self-evaluations, but adapted their evaluations and focused more on the partner information provided. Knowledge gain was not affected. In conclusion, social context can be an important factor in self-regulation emphasizing that individual and collaborative research traditions may complement each other.

Introduction

Open learning environments offer the opportunity for students to choose when and what to learn and how to search for what information, in short: to make important study decisions. Consequently, it is no surprise that research on self-regulated learning in online education increased tremendously in the last years (Tsai, Shen, & Fan, 2013). But online learning is not only self-directed, but also social by nature (Shea & Bidjerano, 2010). With a growing body of web-technology and more and more information available online, self-regulated learning is not done in enclosed spaces and, maybe more importantly, not done in isolation. Further, regulated learning is inherently social, because regulating learning means interacting with an environmental and social context (Järvelä & Hadwin, 2013). Learners may encounter differing opinions challenging the learners' knowledge and creating socio-cognitive conflicts, thus initiating further learning processes. Consequently, we can assume that self-regulated learning is severely influenced by its social context. In collaborative research, social context information is seen as a key prerequisite of meaningful social interaction and thus frequently visualized via (cognitive) group awareness tools which explicitly focus on providing socio-cognitive information, i.e. information on other learners' cognitions, to foster individual as well as collaborative learning processes and outcomes (Janssen & Bodemer, 2013). Integrating individual research traditions in a social setting, our study investigates how visualizing socio-cognitive information influences core individual learning processes, especially the search for information, learners' self-evaluation of knowledge, and learning outcomes.

Background

Metacognitive research suggests that self-regulated learners use their monitoring outcomes to guide their study decisions (Nelson & Narens, 1990). One thoroughly researched metacognitive concept is the individual confidence in one's knowledge or answers, which might act as a sensor for the need to re-study material (Thiede, Anderson, & Theriault, 2003) or influence how much time is spent on it (Dunlosky & Ariel, 2011). It can also support learners in prioritizing and sequencing their learning processes, which is especially important with time constraints (Son & Sethi, 2006). While this usually works well for good self-regulated learners, it relies heavily on the individual skills to monitor the learning progress or outcome (Thiede, 1999). Unfortunately, research also suggests that learners are often overconfident with regard to their knowledge (Pressley, Ghatala, Woloshyn, & Pirie, 1990). This might hamper learning progress due to ineffective study decisions, e.g., learners might feel confident about their knowledge and stop studying early even though they would have needed further study trials to reach their goals (Dunlosky & Rawson, 2012).

There are a number of external sources to inform self-evaluation processes. For example, external evaluations of learning operations or products provided by experts may help to adjust learning strategies and / or monitoring judgments (Butler & Winne, 1995). Conversely, metacognitive judgments also influence the way learners perceive and use externally provided information (Kulhavy & Stock, 1989). Confidence has been shown to influence feedback processing, as well as feedback effects (Butterfield & Metcalfe, 2006). One prevalent theory

is that unexpected feedback (pointing out errors committed with high confidence) motivates deeper elaboration of the feedback message and thus improves learning outcomes (Fazio & Marsh, 2009).

Another source of information informing the individual self-evaluative system is a community or group. Information from large as well as small learning communities may help learners to re-evaluate or even validate (falsify or verify) their knowledge. In collaborative settings, but also in individual online learning, learners may encounter competing and maybe even conflicting opinions of other learners. Comparing one's own knowledge with other learners' externalized knowledge related information can be a strong asset to learning, thus making comparability a key feature of group awareness tools (Bodemer, 2011). Being confronted with opposing points of view constitutes socio-cognitive conflict (Bell, Grossen, & Perret-Clermont, 1985). Within collaborative learning scenarios, socio-cognitive conflicts have been shown to foster learning processes and outcomes (Bodemer, 2011; Johnson & Johnson, 2009; Mugny & Doise, 1978), at least for epistemic conflicts (Darnon, Doll, & Butera, 2007). They are seen as an important motor for collaborative as well as individual learning (Mugny & Doise, 1978). If confronted with conflicting information about the learning subject, the learners own hypotheses are challenged and they are obliged to explain and maybe even defend or backup their position (Johnson & Johnson, 2009) or integrate the differing views (Darnon et al., 2007). Moreover, conceptual conflict has been shown to foster an active search for information due to an increase in epistemic curiosity (Lowry & Johnson, 1981). Within the context of self-regulated learning, cognitive conflicts suggest re-evaluations of one's own cognitions, e.g., they may introduce uncertainty (Buchs, Butera, Mugny, & Darnon, 2004). As a consequence, learners should be prone to initiate search processes to come to a satisfying solution, e.g., to validate one position (Buchs et al., 2004), but this might depend on the interplay of self- and partner-evaluations (Mugny, Butera, Sanchez-Mazas, & Perez, 1995).

While research done in these areas highly suggests that socio-cognitive information alters individual learning processes (e.g., search processes), existing studies only sparsely integrate individual research traditions on metacognitive self-regulation into a social or even collaborative perspective. Contributing in closing this gap, we investigated if and how individual study decisions during self-regulated learning (especially the search for information) are affected by socio-cognitive information about another learner. Thereby we focussed on the interplay of three potentially relevant variables: the individuals' self-evaluation of knowledge (own confidence), the other learners' self-evaluation (partner confidence) and their conflict status (does the partners' knowledge challenge (conflict) or support (consent) own knowledge). The following research questions were addressed:

1. How does the presence of partner information change the selection of additional information?
2. Does the presence of partner information change cognitive or metacognitive learning outcomes?
3. How do learners take their own confidence, the partners' confidence as well as the partners' answers in comparison to own answers (conflict status) into account when selecting additional information?

Methods

The study took place in early summer (May – June) 2014. Data was assessed in the course of a Bachelor's Thesis (Geerdes, 2014). 63 participants took part in the experiment, from which two had to be excluded due to computer failure, which left us with $N = 61$ participants included in our sample. They were all university students, mainly enrolled in a BA or MA course on Applied Cognitive and Media Science (47 female, 14 male). Their age ranged from 18 to 28 years with a mean age of 21.53 ($SD = 2.09$). All experiments were conducted in our research lab; instructions and materials were given via computer. Learners were randomly assigned to one of two research conditions: with and without socio-cognitive partner information available during learning.

Material and procedure

After welcoming and declaration of consent, the participants received some information about the procedure of the experiment and filled out a demographics questionnaire assessing age, sex, and university course. Then they were all presented with a 970-words text on immunology (adapted from material used in a study of Dehler, Bodemer, Buder, and Hesse, 2011, slightly shortened and re-written) and instructed to read the text carefully for up to 20 minutes. When they had finished, they were asked to answer 20 learning tasks, each consisting of a statement they were asked to judge as being true or false (cf. Figure 1). In addition, they were asked to give a confidence rating for each answer by stating on a binary scale if they were sure or unsure that their answer was correct. The answer was spatially coded (true: top, false: bottom), the confidence was color-coded (high confidence: filled-green, low confidence: hatched white-green) (cf. Figure 1). Afterwards, they were again presented with these 20 tasks as well as their answers and confidence ratings and were given up to 15 minutes to request and study additional information on as many tasks as they wanted to by clicking on a respective button provided for each task. They were also able to change their answers and/or confidence ratings. While learners

without partner information (*no partner information* condition) received only their own answers and confidence ratings, learners with partner information (*partner information* condition) additionally received bogus partner information generated by a fixed algorithm, which ensured that partner information was roughly balanced with regard to conflict status and partner confidence for each confidence level (ignoring the validity of the answers). After this second learning phase, participants were all asked to answer the learning tasks again from scratch, including confidence ratings. Finally, they took a knowledge test consisting of 19 questions (four options, single choice format), which assessed the learned concepts more deeply. Again, confidence in each answer was assessed, this time on a six-point Likert scale ranging from “not sure at all” (0) to “absolutely sure” (5).

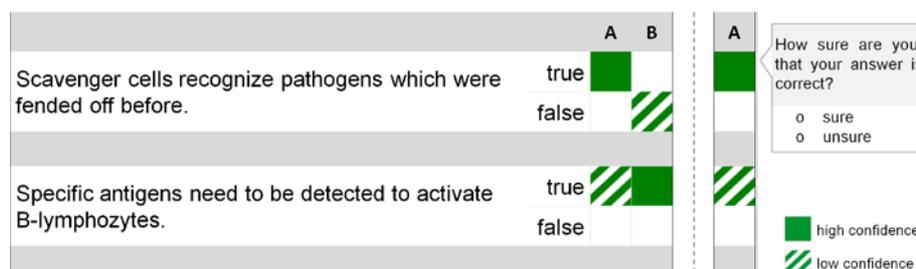


Figure 1. Examples of learning tasks and confidence ratings with (left) and without (right) partner information.

Dependent variables

In order to answer our research questions, we assessed how many and which additional information the learners requested during the second learning phase and in what order and how long they studied the information. To find out if information requests coincided with (low) confidence ratings, we computed within-subject correlations (a method used in metacognition research to describe regulation of study, cf. Thiede, 1999; Thiede et al., 2003). For the *partner information* condition, we also computed a correlation index between conflict status and information requests to capture its influence on learning behaviour. To better understand the interplay between own confidence, partner confidence and conflict status (resulting from a comparison between own and partner answers) on their influence on information requests, we assessed the percentage of information requests for each constellation for the *partner information* condition. Moreover, we assessed performance and confidence levels in the learning tasks and in the knowledge test by counting correctly solved or confident answers and computed a mean confidence for the post test. We assessed relative monitoring accuracy in the form of within-subject phi- or gamma coefficients for each participant (cf. Schraw, Kuch, & Gutierrez, 2013). To analyze the sequence of information requests in conjunction with information on individual confidence (both conditions), partner confidence and conflict status (*partner information* condition only), we ranked the information requests in order of first appearance and computed a mean rank per person for each value (e.g., low) of each variable (e.g., confidence). To eliminate the influence the number of appearances of each value has on its mean rank (mean rank increases automatically with the number of appearances regardless of selection strategy), we computed mean rank differences within each binary variable instead of using the individual values (e.g., $[\text{mean_rank}_{\text{HighConfidence}} - \text{mean_rank}_{\text{LowConfidence}}]$). The magnitude of the resulting figure informs about the extent learners give timewise priority to one value before the other, the algebraic sign tells us which one it is.

Findings

How does partner information change the selection of additional information?

To analyse and compare the selection of additional material, we matched the answering patterns of the learners (and their bogus learning partner) with the event of selecting additional information. To compare the selection agendas of the two experimental conditions, we first counted to how many tasks the learners in each group requested additional information and for how long they studied each one (cf. table 1). *t*-tests for independent samples showed no significant differences between the groups for neither the number of information requests ($t(59) = 0.50, p = .619, d = 0.13$) nor the mean study duration per request ($t(59) = 0.11, p = .914, d = 0.03$).

In a second step, we analyzed if the availability of partner information changed how learners chose the information to study. Without partner information 79.59% ($SD = 21.05$) of the information requested regarded items answered with low confidence, while with partner information it was only 63.78% ($SD = 23.17$). This difference was statistically significant ($t(59) = 2.78; p = .007, d = 0.72$). To further evaluate if learners really use their own confidence ratings to choose additional information, but also to see if conflicting opinions might influence these decisions, we computed within-subject correlations between the initial level of individual

confidence (high vs. low) or the conflict status (presence or absence of conflict) and the presence or absence of a request for information to each task, resulting in individual within-learner phi-coefficients (cf. table 1). *t*-tests on one sample confirmed a significant positive mean relation index between (lack of) confidence and information requests for the *no partner information* ($t(27) = 9.45, p < .001, d = 1.77$) as well as the *partner information* condition ($t(29) = 5.55, p < .001, d = 1.00$) and a positive relation between conflict and information requests in the *partner information* condition ($t(29) = 5.90, p < .001, d = 1.09$). *t*-tests for independent samples showed that mean phi-coefficients between confidence and information requests differed significantly between the two conditions, with the *no partner information* condition having significantly higher coefficients than the *partner information* condition ($t(56) = 2.67, p = .010, d = 0.71$), meaning that confidence ratings were more strongly related to information requests if partner information was not available. A *t*-test for dependent samples comparing the phi-coefficient within the *partner information* condition (own confidence vs. conflicts) showed no difference, meaning neither own confidence nor conflict status was more strongly related to information requests than the other ($t(29) = 0.29, p = .774, d = 0.05$).

Table 1: Descriptive statistics on dependent variables

	partner information						overall		
	available			not available			N	M	SD
	n	M	SD	n	M	SD			
overall study behavior									
...number of information requests	32	11.81	4.48	29	12.38	4.35	61	12.08	4.39
...study duration per information (in seconds)	32	14.58	8.68	29	14.36	7.32	61	14.48	8.00
regulation of study / within-subject phi-coefficient^{(*)(**)}									
...between information requests and own confidence	30	.35	.35	28	.59	.33	58	.47	.36
...between information requests and conflict status	30	.33	.30	--	--	--	--	--	--
performance / number of correctly solved items									
...in knowledge test	32	5.03	2.07	29	5.07	1.60	61	5.05	1.85
...in learning tasks pre	32	13.13	2.06	29	13.38	2.03	61	13.25	2.03
...in learning tasks post	32	15.13	1.91	29	15.45	1.68	61	15.28	1.80
confidence level									
...in knowledge test (mean confidence level)	32	2.02	0.82	29	2.01	0.85	61	2.02	0.82
...in learning tasks pre (number of certain items)	32	9.91	3.60	29	9.62	3.13	61	9.77	3.36
...in learning tasks post (number of certain items)	32	14.50	3.52	29	15.86	2.80	61	15.15	3.25
monitoring accuracy									
...knowledge test (gamma-coefficients)	32	.21	.45	29	.07	.42	61	.14	.44
...learning tasks pre (phi-coefficients)	32	.20	.22	29	.11	.25	61	.15	.24
...learning tasks post (phi-coefficients) ^(*)	31	.26	.23	26	.22	.30	57	.24	.26
sequencing of study process / mean rank differences^(***)									
...own confidence	32	2.20	3.45	29	3.63	4.02	61	2.88	3.77
...conflict status	32	3.96	3.02	--	--	--	--	--	--
...partner confidence	32	0.03	1.45	--	--	--	--	--	--

(*) due to invariability of one factor, no correlation indices could be computed for some participants

(**) positive values indicate that additional information was requested mainly to uncertain answers / conflicts

(***) positive values indicate that uncertain / partner uncertain / conflicting items were considered first

Does partner information change cognitive and metacognitive learning outcomes?

In order to find out if the presence of partner information influences learning outcomes, we compared how many learning tasks the learners were able to solve correctly pre and post learning and how many items they were able to solve correctly in the knowledge test (cf. table 1). A *t*-Test for independent samples showed no significant group differences in the knowledge test ($t(59) = 0.08, p = .937, d = 0.02$). Please note that performance in both groups was just beyond chance, indicating high test difficulty. A two-factorial ANOVA with repeated measures on the learners performance in the learning tasks showed a significant main effect of time ($F(1, 59) = 56.32, p < .001, \eta_p^2 = .49$) with the learners getting better from pre to post, but neither a significant main effect of condition ($F(1, 59) = 0.49, p = .488, \eta_p^2 = .01$), nor an interaction ($F(1, 59) = 0.02, p = .899, \eta_p^2 < .001$).

To see if partner information does rattle individual confidence and if being confronted with potentially conflicting information does enhance self-evaluation processes, we compared individual confidence levels as well as monitoring accuracy. We first compared how many learning tasks the learners solved confidently pre and post learning and how mean confidence differed between the groups in the knowledge test (cf. table 1). *t*-tests showed no difference in confidence levels in the knowledge test ($t(59) = 0.06, p = .954, d = 0.02$). In the learning tasks,

there was a significant effect of time with learners becoming more confident from pre to post ($F(1, 59) = 219.06$, $p < .001$, $\eta_p^2 = .79$) and a significant interaction between time and condition ($F(1, 59) = 5.07$, $p = .028$, $\eta_p^2 = .08$) with the confidence levels of learners with partner information not rising as much as without partner information from pre to post. There was no main effect of condition ($F(1, 59) = 0.50$, $p = .482$, $\eta_p^2 = .01$). Subsequently, we compared monitoring accuracy again with regard to the learning tasks pre and post learning (within-subject phi-coefficients between confidence and performance) and the post-test (within-subject gamma-coefficients between confidence and performance). Descriptive statistics are available in table 1. We conducted a Mann-Whitney-Test (due to violations of the normality assumption) to compare the gamma-coefficients between the groups, but found no significant differences ($U = 558.50$, $Z = 1.37$, $p = .172$, $r = .17$). Accuracy on the learning tasks showed a marginally significant effect of time ($F(1, 55) = 3.49$, $p = .067$, $\eta_p^2 = .06$), but neither an effect of condition ($F(1, 55) = 1.73$, $p = .194$, $\eta_p^2 = .03$), nor an interaction ($F(1, 55) = 0.26$, $p = .611$, $\eta_p^2 = .01$). It is worth mentioning that the correlation-coefficients were quite low in general indicating a weak linkage between self-evaluation and performance (low monitoring accuracy).

How do learners consider own and partner information when requesting information?

To investigate how learners take their own confidence, their partner's answer as well as confidence into account when choosing where and when they need additional information, we first focussed on the *partner information* condition and computed which answer patterns led to requests for additional information. Table 2 visualizes each pattern as well as the mean percentage (and standard deviation) of information requests.

Table 2: Mean information request percentage per answer pattern (*partner information* condition, $n = 28^{(*)}$)

conflict				consensus			
self uncertain	self certain	partner uncertain	partner certain	self uncertain	self certain	partner uncertain	partner certain
$M = 91.96$ $SD = 23.07$	$M = 78.87$ $SD = 36.04$	$M = 51.49$ $SD = 36.43$	$M = 65.77$ $SD = 34.79$	$M = 73.21$ $SD = 36.67$	$M = 65.48$ $SD = 43.49$	$M = 25.00$ $SD = 35.57$	$M = 16.96$ $SD = 33.37$

(*) due to highly unbalanced confidence ratings which did not allow for all constellations, four learners had to be excluded

We computed an ANOVA with our three binary within-subject independent variables describing each pattern (own confidence: high vs. low, partner's confidence: high vs. low, partner's answer: conflicting vs. consenting) and measured the impact on the actual percentage of information requests following each constellation (cf. table 3). Figure 2 illustrates the three-way interaction. Please note that the data was heavily skewed and not normally distributed. We used parametric tests, because methods to model multi-factorial non-parametrical data are scarce. Thus, the results of the inferential analysis should be treated with caution.

Table 3: Results of ANOVA regarding effects of within-subject variables on information request percentage

Effect	Independent Variable(s)	F	$df1$	$df2$	p	η_p^2
main effect	conflict	22.04	1	27	<.001	.45
	confidence	37.38	1	27	<.001	.58
	partner confidence	1.69	1	27	.205	.06
1 st order interaction	conflict * confidence	13.30	1	27	.001	.33
	conflict * partner confidence	2.49	1	27	.126	.08
	confidence * partner confidence	5.01	1	27	.034	.16
2 nd order interaction	conflict * confidence * partner confidence	6.09	1	27	.020	.18

As also observed while studying the phi-correlations regarding the first research question, these results strengthen the assumption that confidence as well as conflicts are highly responsible for information requests. Further, the results indicate that they interact in doing so. Looking at the plots in Figure 2, it seems that individual confidence matters more, if there is an agreement between learning partners and less, when there are conflicting opinions involved. In contrast, the statistically significant interaction between own and partner confidence on closer inspection seems to be explained by the second order interaction, thus we abstain from interpreting it, but focus on the second order interaction: in case of consenting opinions, partner confidence does not seem to impact information requests – the only information relevant is individual confidence. In the conflict case this is very

different. Here, both confidence variables seem to interact. While conflicts with unsure partner and a sure self triggered the least information request (an unsure partner disagreeing might not be regarded as relevant), a conflict with both partners unsure triggered the most. If the partner is sure of him/herself, the own confidence does not seem to matter too much – information request rate is quite high in both cases.

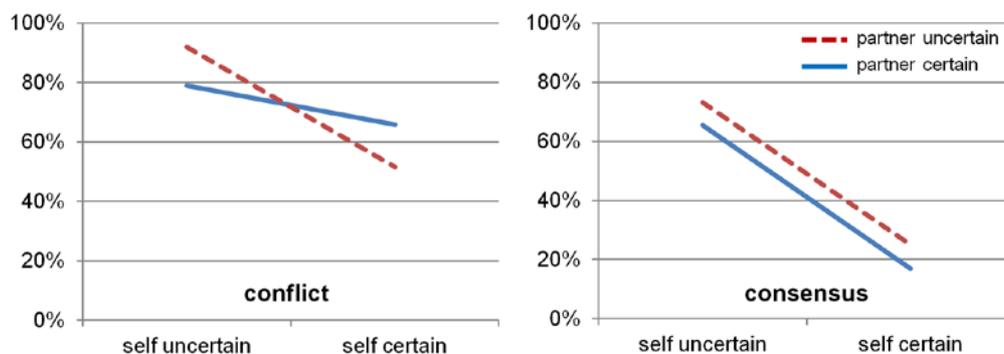


Figure 2. Three-way interaction effect (confidence, conflict status, partner confidence) on information requests.

In a second step we conducted sequential analyses to assess the order of proceeding for the learners in each group. We computed mean rank differences for every variable of interest (procedure described above) to calculate, if learners give precedence to a specific value, e.g., if they consider items with conflicting opinions before considering items with consenting opinions (cf. table 1). *t*-tests on one sample show that own confidence ($t(31) = 3.60, p = .001, d = 0.64$) as well as conflict status ($t(31) = 7.41, p < .001, d = 1.31$) influenced the order learners in the *partner information* group requested information – the mean rank differences differed significantly from chance – but partner confidence didn't ($t(31) = 0.10, p = .918, d = 0.02$). For the *no partner information* group, Wilcoxon signed-rank test (due to violation of the normality assumption) confirmed a similar effect for individual confidence ($Z = 3.67, p < .001, r = .68$). Taking the direction into account, learners requested information on items they were uncertain about or conflicting items earlier than on items answered with certainty or consenting items. A Mann-Whitney-Test between the conditions showed a marginal effect of condition on the mean rank difference for the (own) confidence dimension ($U = 593.50, Z = 1.87, p = .061, r = .24$), hinting that maybe the *partner information* condition used own confidence slightly less than the *no partner information* condition. Within the *partner information* condition we additionally compared the mean rank differences for the three dimensions own confidence, partner confidence and conflict status with *t*-tests for dependent samples to see which variable showed most influential with regard to order of processing. The results showed that the effect of conflict status was significantly greater than that of own confidence ($t(31) = 2.09, p = .045, d = 0.34$) or partner confidence ($t(31) = 6.95, p < .001, d = 1.23$). Mean rank difference for own confidence was significantly greater than for partner confidence ($t(31) = 3.31, p = .002, d = 0.59$).

Conclusions and implications

Our study aimed at describing how the presence of socio-cognitive information influences individual learning processes – even without the chance to communicate, collaborate or interact with the partner. More precisely, we looked at how partner information (information about knowledge and knowledge evaluations of a potential learning partner) influences the individual search for information and learning outcomes. We expected learners with partner information to access more information and study it longer motivated by the presence of conflicting opinions, but they didn't. Differing opinions might not have the same potentially surprising effect as external feedback, if its origin is unknown (Mugny et al., 1995). Additionally, learning outcomes were not affected by the presence of partner information, but it influenced self-evaluations – although not greatly. As expected, the confidence of learners with partner information available was rattled, presumably by conflicting opinions about the correctness of answers (Buchs et al., 2004), but unfortunately this was independent of performance – if a proper re-evaluation took place, it failed to enhance monitoring accuracy. Even more, this increase in uncertainty did not lead to an increase in study behaviour (e.g., number of information requests or duration of study), but it did change its focus: While in the absence of information on learning partners learners based their study decisions (what to study when) heavily on individual confidence (as frequently reported in self-regulation research, e.g., Thiede, 1999; Thiede et al., 2003), with information on a potential learning partner present, the situation became more complex. Analysing the individual relation between information requests and confidence and conflicts we found that without partner information, own confidence was the main influential factor, with partner information it became less so, with conflicts becoming more important. Timewise, conflicting opinions seemed to capture the

learners' attention resulting in slightly higher mean rank differences for the conflict dimension than for the confidence dimension and it also seemed to lessen the effect own confidence had on the search for information as group comparisons indicated (although this difference was not statistically significant). There exist different models on how learners allocate study time and choose and prioritize information to study, mainly resulting from a discrepancy reduction model (Dunlosky & Hertzog, 1998) or a region of proximal learning approach (Metcalf, 2009), with task constraints (Son & Sethi, 2006) and personal agendas (Ariel, Dunlosky, & Bailey, 2009) being important influential factors. Further studies should conduct more profound analyses in this area to broaden this existing research on individual self-regulation by integrating social scenarios. In our study, learners were reasonably free to select any information and they not only attended to conflicts or uncertainly answered items first, they did so primarily. Three-way interactions confirmed that –as expected from research on socio-cognitive conflicts (Lowry & Johnson, 1981)– conflicting opinions became a major influence on information search, while partner confidence had a somewhat unexpected effect: it was expected that the lower the partner confidence the lesser the experience of conflict and thus the lesser the search for information, but this was not the case. If the learning partner agreed with the individuals' opinion (consensus), information requests were solely based on individual confidence, disregarding partner confidence. If conflicts occurred, partner confidence interacted with own confidence to influence information requests. We can conclude that learners do take into account how the potential partners evaluate themselves, even though they are completely unknown. If they disagree and are sure of themselves, learners re-check their own information. If partners are uncertain about their answers, learners may still question their own answers, but especially and much more so, if they are unsure anyway. The latter case might indicate that maximum uncertainty (both uncertain) with conflicting opinions (no indication that any one answer is correct) changes our evaluation of the task, which might also lead to increased information search (a discrepancy reduction approach might indicate the farthest group distance from getting the correct answer and thus the most reason to attend to the task).

In conclusion, social context does seem to affect individual learning and we should reinforce our (theoretical and empirical) efforts to describe 21st century self-regulated learning as what it is – a communal process. Study decisions are strongly affected by the presence of socio-cognitive information. Focussing on relevant information might be enhanced by the presence of others, but may also become more complicated as more or unknown actors come into play. While this study was able to shed some light on how learners integrate knowledge of other's opinions as well as their self-evaluations to make study decisions in a highly controlled experimental setting, it was beyond the scope to analyse actual collaborative efforts. While we focussed on socio-cognitive information as a source of information to foster individual study decisions, the next logical step would be to explicitly integrate the notion of a learning partner as a source of further information – to interact, to ask, to explain, to question, in short: into an inherently collaborative learning scenario. Even though research on group awareness tools frequently incorporates socio-cognitive information into collaborative learning settings, combining it with metacognitive research by integrating self-evaluations of both partners as well as cognitive information on content-knowledge and incorporating methods of both research traditions may be a strong asset to research on collaborative as well as individual self-regulated learning.

References

- Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General*, 138(3), 432–447. <http://doi.org/10.1037/a0015928>
- Bell, N., Grossen, M., & Perret-Clermont, A.-N. (1985). Sociocognitive conflict and intellectual growth. *New Directions for Child and Adolescent Development*, 1985(29), 41–54. <http://doi.org/10.1002/cd.23219852905>
- Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior*, 27(3), 1079–1086. <http://doi.org/10.1016/j.chb.2010.05.016>
- Buchs, C., Butera, F., Mugny, G., & Darnon, C. (2004). Conflict Elaboration and Cognitive Outcomes. *Theory Into Practice*, 43(1), 23–30. http://doi.org/10.1207/s15430421tip4301_4
- Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65(3), 245–281. <http://doi.org/10.3102/00346543065003245>
- Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, 1(1), 69–84. <http://doi.org/10.1007/s11409-006-6894-z>
- Darnon, C., Doll, S., & Butera, F. (2007). Dealing with a disagreeing partner: Relational and epistemic conflict elaboration. *European Journal of Psychology of Education*, 22(3), 227–242. <http://doi.org/10.1007/BF03173423>

- Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, 27(3), 1068–1078. <http://doi.org/10.1016/j.chb.2010.05.018>
- Dunlosky, J., & Ariel, R. (2011). Chapter four - Self-Regulated Learning and the Allocation of Study Time. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 54, pp. 103–140). Academic Press. Retrieved from <http://www.sciencedirect.com/science/article/pii/B9780123855275000048>
- Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging*, 13(4), 597–607. <http://doi.org/10.1037/0882-7974.13.4.597>
- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271–280. <http://doi.org/10.1016/j.learninstruc.2011.08.003>
- Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, 16(1), 88–92. <http://doi.org/10.3758/PBR.16.1.88>
- Geerdes, K. (2014). *The influence of partner information on the individual search for information and learning outcomes [translated]*. University of Duisburg-Essen.
- Janssen, J., & Bodemer, D. (2013). Coordinated Computer-Supported Collaborative Learning: Awareness and Awareness Tools. *Educational Psychologist*, 48(1), 40–55. <http://doi.org/10.1080/00461520.2012.749153>
- Järvelä, S., & Hadwin, A. F. (2013). New Frontiers: Regulating Learning in CSCL. *Educational Psychologist*, 48(1), 25–39. <http://doi.org/10.1080/00461520.2012.748006>
- Johnson, D. W., & Johnson, R. T. (2009). Energizing Learning: The Instructional Power of Conflict. *Educational Researcher*, 38(1), 37–51. <http://doi.org/10.3102/0013189X08330540>
- Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1(4), 279–308. <http://doi.org/10.1007/BF01320096>
- Lowry, N., & Johnson, D. W. (1981). Effects of controversy on epistemic curiosity, achievement, and attitudes. *The Journal of Social Psychology*, 115(1), 31–43. <http://doi.org/10.1080/00224545.1981.9711985>
- Metcalf, J. (2009). Metacognitive Judgments and Control of Study. *Current Directions in Psychological Science*, 18(3), 159–163. <http://doi.org/10.1111/j.1467-8721.2009.01628.x>
- Mugny, G., Butera, F., Sanchez-Mazas, M., & Perez, J. A. (1995). Judgments in conflict: The conflict elaboration theory of social influence. In B. Boothe, R. Hirsig, A. Helminger, & R. Volkart (Eds.), *Perception, evaluation, interpretation*. Seattle: Hogrefe & Huber Publishers.
- Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology*, 8(2), 181–192. <http://doi.org/10.1002/ejsp.2420080204>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of Learning & Motivation* (Vol. 26, pp. 125–173). New York: Academic Press.
- Pressley, M., Ghatala, E. S., Woloshyn, V., & Pirie, J. (1990). Sometimes Adults Miss the Main Ideas and Do Not Realize It: Confidence in Responses to Short-Answer and Multiple-Choice Comprehension Questions. *Reading Research Quarterly*, 25(3), 232–249. <http://doi.org/10.2307/748004>
- Schraw, G., Kuch, F., & Gutierrez, A. P. (2013). Measure for measure: Calibrating ten commonly used calibration scores. *Learning and Instruction*, 24, 48–57. <http://doi.org/10.1016/j.learninstruc.2012.08.007>
- Shea, P., & Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers & Education*, 55(4), 1721–1731. <http://doi.org/10.1016/j.compedu.2010.07.017>
- Son, L. K., & Sethi, R. (2006). Metacognitive Control and Optimal Learning. *Cognitive Science*, 30(4), 759–774. http://doi.org/10.1207/s15516709cog0000_74
- Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, 6(4), 662–667. <http://doi.org/10.3758/BF03212976>
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, 95(1), 66–73. <http://doi.org/10.1037/0022-0663.95.1.66>
- Tsai, C.-W., Shen, P.-D., & Fan, Y.-T. (2013). Research trends in self-regulated learning research in online learning environments: A review of studies published in selected journals from 2003 to 2012. *British Journal of Educational Technology*, 44(5), E107–E110. <http://doi.org/10.1111/bjet.12017>

Development of Disciplined Interpretation Using Computational Modeling in the Elementary Science Classroom

Amy Voss Farris, Vanderbilt University, amy.farris@vanderbilt.edu
Amanda Catherine Dickes, Vanderbilt University, amanda.c.dickes@vanderbilt.edu
Pratim Sengupta, University of Calgary, pratim.sengupta@ucalgary.ca

Abstract: Studies of scientists building models show that the development of scientific models involves a great deal of subjectivity. However, science as experienced in school settings typically emphasizes an overly objective and rationalistic view. In this paper, we argue for focusing on the development of *disciplined interpretation* as an epistemic and representational practice that progressively deepens students' computational modeling in science by valuing, rather than deemphasizing, the subjective nature of the *experience* of modeling. We report results from a study in which fourth grade children engaged in computational modeling throughout the academic year. We present three salient themes that characterize the development of students' disciplined interpretations in terms of their development of computational modeling as a way of seeing and doing science.

Keywords: modeling; agent-based models; disciplined interpretation; epistemology; science education

Introduction

The development of a modeling-based epistemology in science is central to the development of scientific literacy (Nersessian, 2008; Lehrer, 2009). Models are fundamentally analogical forms (Giere, 1988), and studies of scientists building models show that the development of scientific models involves a great deal of subjectivity, and in many cases, is deeply intertwined with personal experiences and disciplined sensibilities about interpreting evidence (Keller, 1984; Ochs, Gonzales & Jacoby, 1996). However, science as experienced in school settings typically emphasizes an overly objective and rationalistic view (Lemke, 2001). This is particularly relevant for studies that involve learning science as modeling, because modeling is an act of design (Lehrer, 2009), and therefore, deeply interweaves “knowing” and “action” (Schön, 1995). This interweaving, as Schön (1995) argued, is deeply tied to subjectivities such as learning to see things from the perspectives of others, and engaging in reflective conversations with the situation. This is a far cry from common uses of computational modeling in the science classroom, which has traditionally followed the grossly “linear approximation” of teaching correct concepts through guided algorithmic refinement (e.g., White & Frederiksen, 1998; Boole & van Jooligen, 2013). Such images, grounded in technical rationality (Schön, 1995), leave out the development of necessary subjectivities, and this is the issue we address in this paper.

In this paper, building on Daston & Galison's (2007) notion of “trained judgment”, we argue for focusing on the development of *disciplined interpretation* as an epistemic and representational practice that progressively deepens students' computational modeling expertise by valuing, rather than deemphasizing the subjective nature of the experience of modeling. We report results from a study in which fourth grade children engaged in computational modeling by iteratively creating, presenting and evaluating their mathematical measures and computational models of motion and ecology throughout the academic year. In this paper, we will focus on their models of motion and investigate how they develop progressively more mathematically and computationally refined representations of motion as a process of continuous change. We present three salient themes that characterize the development of students' disciplined interpretations in terms of their development of computational modeling as a way of seeing and doing science: (1) the intertwined nature of designing models as communicative forms and developing deeper interpretations of numerical data; (2) progressive development of criteria for what counts as a “good” measure, by extending the tools of modeling beyond computational media; and (3) the shift from more normative and canonical forms of mathematization to invented forms of simulations as equally “accurate” and communicative models.

Theoretical background

Science studies scholars have argued that the development of disciplined interpretation is central to the production of scientific knowledge. There is always a gap between scientific representations and reality; scientific models are by their very nature non-veridical designed artifacts that are deeply influenced by the technological infrastructure used for inquiry and representation, as well as the purpose of representation (Galison, 1996; Daston

and Galison, 2007). These authors have argued that the epistemic stance of scientific work shifted from a falsely “objectivist” stance to “trained judgment”. This was evident in their comparison between the 19th century introduction of photographic technology where the machinic nature of photography created an impression that scientists could “get out of the way,” and let the photograph produce what became perceived as bare, uninterpreted, objective “facts.” In contrast, beginning in the early to mid-twentieth century, with the advent of the printing press that in turn widened the audience for scientific works such as atlases, the production of scientific images became necessarily more interpretive on the part of the scientist, with a clear goal of *enhancing the communicativity* of the images (Daston & Galison, 2007). This represents another moment in the history of objectivity: *trained judgment*. We further posit that the advent of computing as a key mode and medium of scientific inquiry further amplifies this epistemic stance. For example, a recent ethnographic study of scientific work in a biomedical engineering lab illustrates how the inherent interdisciplinarity of the practice of computational modeling results in new conceptual innovations, often by bridging the gap between disciplinary perspectives, as well as between theorization, dynamic visualization and experimental work (Chandrasekharan & Nersessian, 2014).

How can disciplined sensibilities about modeling develop, especially in the context of computational modeling, in an elementary science classroom? This is the central concern of this paper. We posit that answering this question involves two key issues: The first issue involves finding a suitable paradigm of computing that is intuitive and generative for young children. Scholars have argued that a particular form of computation—agent-based computation—can serve as an effective pedagogical approach that can help children bootstrap their own pre-instructional ideas and representational competencies in order to develop scientific expertise through modeling (Papert, 1980; Sherin, diSessa & Hammer, 1993; Sengupta, Kinnebrew, Basu, Biswas, and Clark, 2013). In agent-based computation, users construct programs by providing simple rules to a computational object or agent, such as a LOGO turtle, which then enacts the rules through movement in computational space. Programming the agent involves thinking like it, which enables the learner to engage in embodied and intuitive reasoning (Papert, 1980; Danish, 2014). In agent-based models (ABMs), simple, agent-level actions are repeated over time (in the case of generating continuous movement from discrete actions) and/or across multiple agents (e.g., in ecological phenomena). There is ample evidence in the literature that ABMs can support the development of representational competence in children (Sherin, diSessa & Hammer, 1993; Sengupta et al., 2015), and that pedagogies that emphasize the creation of representational conventions that can be understood by others is a kind of “selective pressure” that brings forth iterative representational innovations by the students (Enyedy, 2005). We therefore adopted agent-based modeling and programming as the *medium* of computation that children engage with in our study.

Given the rather short duration of most education research studies and the limited involvement of the teacher in the design and implementation of what happens during class, little is known about what happens when computational modeling practices develop as a long-term practice in the science classroom over an entire academic year. Herein lies the second issue: *we posit that disciplined interpretation is one such aspect of development that fundamentally involves long-term engagement of learners with the practice of modeling*. While some scholars have shown that children can indeed engage deeply with agent-based modeling to learn science over shorter durations ranging from a few days to a couple of weeks (e.g., Danish, 2014), we posit that a long-term focus is essential because scientific modeling is itself a long-term and complex endeavor in professional practice. It comprises complex interactions among theories, material means, the phenomenon of interest, the representational infrastructure, as well as the social contexts that shape these interactions, usually over a long period of time—often spanning several years (Pickering, 1995). A longer-term focus in the science classroom can provide insights into how students develop fairly stable and sophisticated *disciplinary dispositions* that involve an interplay between children’s intuitions, interpretations and actions on the world in order to progressively symbolize and refine the representations for scientific modeling (Lehrer, 2009). What does the development of children’s disciplinary dispositions look like when computation becomes a medium for scientific modeling throughout the academic year? In this paper, our goal is to provide illustrative cases, selected carefully from a year-long study that will illuminate key characteristics, and different forms of manifestations of a particular form of disciplinary dispositions that is often neglected in the science education literature—disciplined interpretation.

Methods

The ViMAP modeling environment

The modeling platform we used is ViMAP (Sengupta et al., 2015). ViMAP is an agent-based visual programming language that uses NetLogo (Wilensky, 1999) as its simulation engine. In ViMAP, users construct programs using a drag-and-drop interface to control the behaviors of one or more computational agents. ViMAP programming

primitives include domain-specific and domain-general commands as well as a “grapher” with multiple graphing windows, which allows users to design mathematical measures and compare across measures of different agent- and class-level variables. Figure 1 shows the programming interface and the graphing interface.

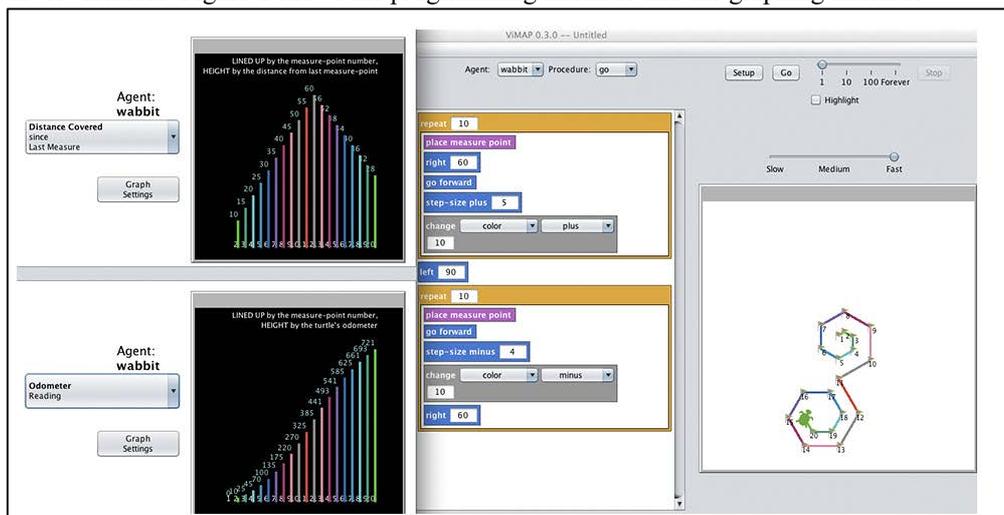


Figure 1. Screenshot of the ViMAP Modeling Environment (www.vimapk12.net).

Participants, setting, and data collection

The data that we have collected is in the context of a public school fourth grade classroom (ages 9 and 10) in an urban southeastern city. The study is a design study (Cobb et al., 2003) in which we worked in partnership with the classroom teacher to integrate agent-based programming and modeling within the existing math and science curriculum. Students carried out investigations of natural phenomena in kinematics and in ecology in modeling cycles that include modeling in ViMAP. Twenty-one students and their teacher participated in the classroom work, which was embedded in the regular curriculum. The teacher and the research team co-planned the activities based on the students’ progress and the teacher’s plans across the curriculum. During class time, the teacher played the primary teaching role, often adjusting the plans to meet the emerging instructional opportunities as plans were enacted. Two graduate-student members of the research team collected data. Ninety-five percent of students who attend the school are eligible for free and reduced-price lunch. The ethnoraical and gender composition of the student population in this classroom was as follows: African-American: 19, Latino: 1; Somali: 1; Male: 11, Female: 10. A sequence of the learning activities is shown in Table 1. In this paper, we only report the analysis of all modeling activities from October 14 until February 3. The research approach was both microgenetic and sociogenetic, because our goal is to understand changes in student thinking and how these ideas are shared and taken up in the larger class community. The forms of data collection include video records of all classes, detailed field notes for each day, collection of all of students’ models and non-computational artifacts, as well as interviews with student groups and individual students.

Table 1: Sequence of activities (Analysis reported in this paper through Feb 3)

Observations, pre-assessment, and interviews	Aug. 11 – Sept. 8	Researchers conduct observations, preliminary interviews with all students in the class
Survival Kits Geometry Unit	Sept. 9 – Oct. 2	Intro to ViMAP programming and modeling; Turtle geometry, centered around learning goals in perimeter, area, and angles of polygons; model sharing and revision
“Constant Speed” Robots	Oct. 14 – Nov. 20	Students develop understanding of speed as a rate of the distance traveled in a unit of time, including cycles of model sharing and revision; students used both ViMAP and physical modeling
Constant Acceleration and Gravity	Nov. 25 – Feb. 3	Students find ways to measure and model continuous changes in speed, using acceleration down a ramp and free fall as contexts; students used ViMAP, video analysis and physical modeling
Friction	Feb. 5 – Mar. 31	Students model processes of “slowing down” for Matchbox cars on different surfaces; students used both ViMAP and physical modeling
Interviews	Apr. 7 – Apr. 28	Mid-year interviews with all students

Modeling Ant Colonies	May 6 – May 13	Students model ant foraging, reproduction and predation in ant colonies in an embodied modeling activity, followed by programming in ViMAP-Ants. Students share and refine their models with 8 th grade mentors.
Post-Assessment	May 14 - May 19	End-of-year assessment and focus group

The role of the teacher and researcher: Teacher partnership

Teaching played a significant role in this study (Sengupta et al., 2015). Although a detailed analysis of teaching is outside the scope of this paper, it is important to note that the teacher reframed programming as a medium and activities for designing mathematical measures (i.e., units of measurement and graphs) of motion. The teacher co-designed (along with the researchers) and implemented learning activities that supported the interpretation and construction of mathematical measures using ViMAP as a way to explain a real-life phenomenon involving motion (e.g., walking and running). In these activities, she maintained an emphasis on connecting modeling in ViMAP to relevant out-of-computer modeling experiences, such as embodied and physical modeling activities. Furthermore, she created a culture for sharing and critiquing peer models, that is, their ViMAP programs, simulations and graphs. In this process, students began developing criteria for what features of their models would be worth sharing: the emphasis on *communicativity* acted as a selective pressure for model improvement; and, the class as a whole, normatively, developed criteria for what would count as a "good" computational model. These criteria originated in teacher-led class discussions as socially defined (voted by popular choice), but over time, became progressively more grounded in students' mathematical explanations of relevant aspects of their ViMAP simulations. This led students to use progressively more sophisticated computational abstractions, such as loops and conditionals, in order to make their models predictive (Sengupta et al., 2015).

Analytic approach

We conducted a thematic analysis (Miles & Huberman, 2004) in order to identify key forms of disciplined interpretations that learners developed during the phase of modeling motion. A theme captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set. In our study, at the highest level, each theme represents an interpretive judgment. Each theme, in turn, consists of sub-themes, which are sets of relevant *representational moves*, that is, actions undertaken by the learners that involve the creation, and/or editing of computational programs and other related representations, and *epistemic moves* (e.g., arguments about the validity or significance of certain representations). Over time, these representational and epistemic moves constitute, or lead to the development of an *interpretive judgment* (e.g., what counts as a "typical" measurement; what counts as a "good video"). These interpretive judgments developed through progressive refinement of models and moving back and forth across tangible, diagrammatic, and computational models of motion. Therefore, besides the learners' subjectivities, the judgments themselves are inextricably tied to the media involved in modeling, in addition to mathematical and physical ideas, and computational abstractions.

Findings

Interpreting numerical data and the emergence of ideas about error

In the first modeling cycle, the teacher wanted to design a context for students to define constant speed motion in terms of distances traveled per unit of time. The researchers initially wanted to use verbal descriptions of motion, as the best approximations of "constant speed" would require motors. The teacher, however, insisted on introducing physical objects, both computational and non-computational, as part of the modeling activity. Her goal was to "make things concrete," i.e., to transform the modeling activity from a virtual and conceptual one into a lived-in experience for her students. The students were therefore provided with Lego Mindstorms NXT robots programmed to move at a constant speed. We also provided stopwatches, adhesive Post-it flags, and seamstress-style measuring tapes, and asked students to measure the distances traveled in regular intervals of time. Students coordinated the placement of position-marking flags with a stopwatch in order to come up with mathematical measures and explanations of robots moving at constant speed. In Figure 1, a student is measuring the distance between flags to find the distance traveled in each three-second interval.

Students used these measurements to create computational models of the motion in ViMAP. In their initial models, each group's measurements for the distances traveled in each three-second interval were non-uniform due to challenges inherent in the act of measurement. However, none of the students problematized their data by considering the limitations of the devices for measurement or the imperfect coordination of the placement of flags with the stopwatch. Students thought of themselves as workers following an unquestionable procedure.

Some of them argued some about issues of fidelity *during* the measurement activity. However, once data collection was complete, no one thought critically about the potential problems in her or his data even though for most students, the data showed a wide variation in the measurements. Upon noticing this, the instructors designed an activity in which students watched a video of one group carrying out their measurement and data collection, and as a class, critiqued their work as shown in the video. The class replayed the video several times in order to notice and reflect upon successes and breakdowns in measurement. The ensuing discussion led to the first student talk about error: “It’s not that the robot was moving differently, it’s that *we* were making mistakes!” Making sense of the data with a focus on analyzing the *lived experience* of designing the measure helped students interpret data as *designed measures*, not independent of the challenges of measurement. This was particularly evident in their noticings of the various sources of error, as they replayed and re-analyzed the video. For example, some groups noticed that longer measurements of displacement were often coupled with shorter measurements, indicating that the timing of placement of the measure-flag shared by those measurements was likely off. Some groups also discovered errors due to misreading the measuring tape, and in some cases, due to the sticky flags being unintentionally moved by getting stuck to students’ shoes.



Figure 2. Measuring constant speed using adhesive paper flags as “measure flags” (left) and one student’s model of measured and “typical” step-sizes (right).

We then asked students to review their measurements in order to determine what they believed was a “typical” distance measurement for their robot to travel in three seconds. The teacher welcomed this as an opportunity to connect the modeling activity with learning about measures of central tendency in their math curriculum. The students iterated upon their existing computational models using a second computational agent in the same simulation to represent the motion according to their “typical” values (mean, median, or mode) for speed. Figure 2 shows one student’s model: her “measure-points” in the first-iteration were the following distances apart, measured in inches: 30, 36, 30, 26, 39, 30. Her second-iteration data shows six uniform measurements of 30 inches each. For each turtle, students programmed ViMAP to generate two graphs: one showing the value of each measurement (not shown), and another producing the total distance traveled by the robot. In the example shown in Figure 2, the graphs of the total distance traveled by the agent show a total distance of 191 inches for the measured data, and 180 inches for the adjusted, or “typical” data. Students also recognized that computationally, the typical model could be expressed as a loop, which the students and teacher appreciated as a more succinct program. A second affordance of the program for Iteration 2 is that the number of repeats can be changed to simulate the robot traveling for a longer period of time at the same speed.

In sum, making mathematical meaning of motion as processes of time-based change required the generation of and coordination among different *representational moves*, involving multiple forms of digital and paper-based, discrete-mathematical representations of the phenomena under study. Annotating video and photographic images in order to communicate and argue for the number of loops needed in their programs became a viable but emergent method for connecting among the representations, and can also be regarded as *epistemic moves* that grounded these representations within the disciplinary concepts. Students’ agency and involvement in creating connections across representations for the purpose of making meaning represents a key practice in model-based reasoning. Epistemologically, the connections among representations were a shared unknown, and it was up to the members of the class to come up with and refine generative ways to see, quantify, and model salient aspects of motion.

What counts as a “good” video for measuring acceleration?

Following students’ refinement of their descriptions of constant speed and (average) speed on inclines of varying steepness, we began to work on developing descriptions of acceleration. We provided clear acrylic tracks, marbles, and Lego bricks to build supports for the top and bottom of the tracks, as well as stopwatches and adhesive paper flags, in order to begin to measure acceleration. Students’ initial descriptions neglected processes of continuous change: the marble was “slow” at the top of the ramp, and “fast” at the bottom. When asked to measure how speed

was changing, students tried to reapply their method with the robots: they attempted to place flags at equal intervals of time, but they soon decided this was too difficult: the motion was *too fast* for the method used when measuring the speed of the (slower) robots. As a potential solution to this problem, the instructors introduced digital video as a new method for collecting and analyzing motion data. We made this design decision because authoring videos leverages out-of-school literacies for making and working with digital video, and, high frame-rate videos afford the possibility of slowing down recordings of motion that are otherwise too fast to measure.

Children’s ideas of what counted as a “good” video for measuring acceleration changed dramatically between the first and second iterations of their video recording and subsequent analysis. An example of student work is shown in Figure 3. Initially, their videos followed the marble in an action perspective, but this made measurement impossible because there was no frame of reference from which to measure the distances traveled. After attempting to measure the acceleration of the marble using the first round of videos, and after classroom-wide discussions, the class developed a norm for what counts as a “good” video for measuring motion in a frame-by-frame analysis: the camera has to stay still and the field of view must show the whole motion. Central to this was the realization that the viewer should be able to see the marble and identify the exact frame at which it was released. Students iteratively developed, shared and critiqued several videos, and progressively refined their measures, and over time, developed measures using physical and material means that used discrete mathematical representations similar to their ViMAP turtle’s “step-size”. One common form of measure involved placing a flag on their computer screens to mark the position of the marble at regular intervals of video frames. The pattern of the increasing distances between successive flags visually represented acceleration as a continuous process of change (Figure 3). The students’ *epistemic moves*, evident in the form of their explanations and concerns for what makes a video of motion useful for measuring speed represents the students’ understanding of important aspects of motion from a disciplinary perspective. These measurements, only possible through an innovative use of video, consisting of a set of eminently *representational moves*, yet deeply interwoven with the *epistemic moves*, were then used to make computational models that further communicate the mathematical pattern of change, as discussed in the following section.

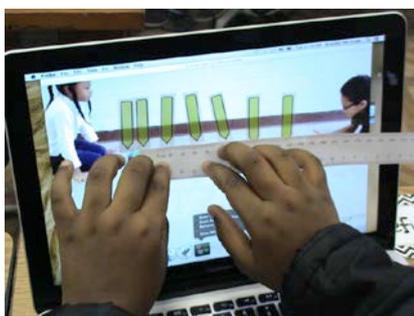


Figure 3. Adhesive flags mark the position of the marble at a regular interval of 10 frames.

Expanding views of “accuracy” to create visually communicative models

In the first two modeling cycles (Oct. – Dec.), conceptual understanding of the unit of “speed” played an important role in their state-mandated science curriculum, and the teacher emphasized distance traveled by the ViMAP turtle in one “step” as the representation for speed. During this phase of the study, the *goodness* of ViMAP models as a representation of motion was normatively evaluated by the class based on the match between the speed vs. time graph (“distance covered since last measure”), and the speed data that the model was designed to represent. Given that the graphs made the pattern of change explicit in these models, the students came to see graphs as the primary communicative devices, and the turtle enactment (i.e., the geometric shape generated by the turtle commands) was seen as merely the means to generate the graph. This was evident in multiple episodes of students’ sharing and presenting their models with the class. Over time, especially during Jan. – Feb., the instructors began to encourage students to further explore the ViMAP commands library, so that they would begin to deepen their use of the programming language. The goal here was to prompt students to re-envision and re-design their models using newer turtle variables, so that they could make their turtle graphics less literal and more visually and mathematically communicative. As a result of this instructional push, during the third modeling cycle, all students began to take a more design-oriented approach to producing models that communicate the most important ideas, and variation in student models emerged. In terms of *representational moves*, all the students in the class expanded their use of *variables* by using new commands to represent mathematically the gradual change in speed using one or more of the following variables: rotation, contrasts of color, pen-width, or relative size of agents. This was, in turn, motivated by and inextricably related to an *epistemic move*, that is—the students’ goals of making relevant features of the phenomenon (motion) more salient to the class during presentations in their ViMAP models. We

illustrate this change with the work of one student, Darien. His first model of constant acceleration is shown in Figure 4 (left). The figure shows the inscription made by the agent as it executes the associated commands. The model increases the distance traveled by the agent by two step-size units with each step. However, the enactment itself is limited in communicating the regularity of the increase—one would need to look at Darien’s code or at the graphs to understand the regularity of the change.

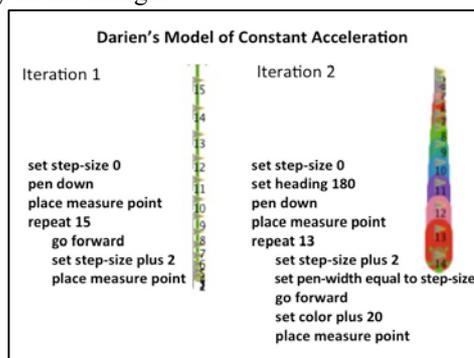


Figure 4. Darien’s first (left) and final (right) models of acceleration.

In his final model (Figure 4, right), Darien added color changes to visually differentiate the individual steps of the agent, and co-varied the pen-width with the step size using the command *set <pen width> equal to <step-size>*. Darien presented his model to the class, describing that the increasing pen-width of the ViMAP turtle is intended to communicate that the ball is getting faster as it falls. When he was sharing his work with the class, students asked questions about the representational significance of different aspects of his model, such as “Why does it look like a baseball bat?”, “Can we see the data”, and pointing out redundancies in his code (an initial *set <pen-width>* command, that was being overridden by the co-variation command). An excerpt of talk from Darien’s presentation is shown below:

- Akia: Why do you have *set pen width* equal to *step size*?
- Darien: To actually help the pen width [be] equal to the step size, so that way, the pen, the size will actually get bigger EV’RY STEP. [holds a hands slightly apart to show a space between, gesture beats and enlarges at syllables of ev-’ry step]
- Teacher: So when it gets bigger, does that show that your speed is increasing or decreasing?
- Darien: Increasing.

Darien wanted to show that pen-width was increasing with every step, and illustrates this with his hand movements as he speaks, beginning with his hands slightly apart, and enlarging the space on beats with the syllables “ev-’ry step”. One could argue that Darien’s initial model was more canonical, because it uses the commonly used representations of dot-traces and graphs. However, in his revised model, Darien’s goal was to make the process of a steady increase in speed explicit *without* the use of graphs. This in turn led him to using an interpretive move that involved using a computationally more sophisticated data representation—co-variation—in order to link the visual appearance of his model (*pen-width*) to a variable that was significant in terms of representing the underlying physics (*step-size*, or speed).

Conclusions and implications

We have argued here that when children can engage in long-term, extended cycles of modeling using agent-based computational platforms (e.g., ViMAP) and complementary forms of physical modeling, they can begin to develop disciplinary dispositions and sensibilities pertaining to scientific modeling that parallel the more mature interpretive work of scientists (e.g., see Daston & Galison, 2007). The long-term nature of the study allowed students to connect representational experiences across modalities, including their computational representations with their lived experiences of designing measures in the real world with physical objects. Through such experiences, children came to view data as designed measures, and their views of what counts as a “good” model deepened significantly as they engaged in cycles of sharing and refining their models to be progressively more

communicative. The emphasis on communicativity also led students to make deeper forays into programming and computational thinking.

Our work has implications for the praxis of computational modeling in the science classroom. There is now a growing body of literature that argues for the use of multiple and complementary forms of modeling in the classroom (e.g. Danish, 2014; Dickes et al., in press). In our study too, the students' representational and epistemic work were distributed across a range of computational and non-computational materials, using which they iteratively represented motion as a process continuous change. While modeling with ViMAP enabled the students to connect graphs of change over time to units of change (e.g., step-size), modeling with materials complemented this activity by enabling them to generate the phenomenon being modeled in the "real world", as well as to design the measure of change (e.g., step-size) using video analysis. It is also important to note that "making things concrete" using material forms was an instructional push initiated by the teacher, who co-designed these activities with the research team. We therefore believe that designing complementary forms of computational and non-computational modeling is critical for enabling teacher-adoption and appropriation of computational modeling in the K12 science curricula.

References

- Bollen, L., & van Joolingen, W. R. (2013). SimSketch: multiagent simulations based on learner-created sketches for early science education. *IEEE transactions on learning technologies*, (3), 208-216.
- Cobb, P., Confrey, J., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9-13.
- Danish, J. A. (2014). Applying an activity theory lens to designing instruction for learning about the structure, behavior, and function of a honeybee system. *Journal of the Learning Sciences*, 23(2), 100-148.
- Daston, L. & Galison, P. (2007). *Objectivity*. Zone Books.
- Dickes, A. C., Sengupta, P., Farris, A. V., & Basu, S. (in press). Development of mechanistic reasoning and multi-level explanations of ecology in 3rd grade using agent-based models. *Science Education*.
- Enyedy, N. (2005). Inventing mapping: Creating cultural forms to solve collective problems. *Cognition and Instruction*, 23(4), 427-466.
- Keller, E. F. (1984). *A Feeling for the Organism, 10th Anniversary Edition: The Life and Work of Barbara McClintock*. Macmillan.
- Lehrer, R. (2009). Designing to develop disciplinary dispositions: Modeling natural systems. *American Psychologist*, 64(8), 759.
- Lemke, J. L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of research in science teaching*, 38(3), 296-316.
- Nersessian, N. (2008). *Creating scientific concepts*. MIT Press.
- Ochs, E., Gonzales, P., & Jacoby, S. (1996). "When I come down I'm in the domain state": Grammar and graphic representation in the interpretive activity of physicists. *Studies in Interactional Sociolinguistics*, 13, 328-369.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc.
- Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. University of Chicago Press.
- Schön, D. A. (1995). Knowing-in-action: The new scholarship requires a new epistemology. *Change: The Magazine of Higher Learning*, 27(6), 27-34.
- Sengupta, P., Dickes, A. C., Farris, A. V., Karan, A., Martin, K., & Wright, M. (2015). Programming in K12 Science Classrooms. *Communications of the ACM*, 58(1), 33- 35.
- Sengupta, P., Kinnebrew, J., Basu, S., Biswas, G., and Clark, D. (2013). Integrating Computational Thinking with K12 Science Education Using Agent-Based Modeling: A Theoretical Framework. *Education & Information Technologies*, Vol. 18, 351 – 380.
- Sherin, B., diSessa, A. A., & Hammer, D. (1993). Dynaturtle revisited: Learning physics through collaborative design of a computer model. *Interactive Learning Environments*, 3(2), 91-118.
- White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and instruction*, 16(1), 3-118.
- Wilensky, U. (1999). *NetLogo*. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Acknowledgements

This research was supported in part by an NSF Early CAREER Award (PI: Pratim Sengupta). All opinions are the authors' and not endorsed by supporting institutions.

Mobilities of Criticality: Space-Making, Identity and Agency in a Youth-Centered Makerspace

Angela Calabrese Barton, Michigan State University, acb@msu.edu
Edna Tan, University of North Carolina at Greensboro, e_tan@uncg.edu
Myunghwan Shin, Michigan State University, shinmyu4@msu.edu

Abstract: We examine in-depth cases of youth makers from non-dominant communities in two makerspace clubs in two different mid-sized US cities. We argue that through their making practice, they are involved not only in “artifact making” (the prototypically viewed outcome of makerspace work), but also in space-making within and across the worlds of STEM, makerspaces, and community. The data suggest that such space-making fosters new forms of interaction among scales of activity, and supports the movement of ideas, resources, relationships and bodies in support of youths’ emerging practices and how they might be recognized for them. As the youth engage in their making practice, they inscribe new meanings for what it means to make within the worlds they inhabit, refiguring participation in these worlds and the possibilities for becoming within them. We used a mobilities of learning framework to guide our analysis.

Keywords: Science, youth, making, mobilities of learning

When little kids are playing outside football and it's getting too dark and they still keep playing and somebody might get hit in the head or something cause they can't see the ball really so I'm going to light up the football so you can see where it's going. (Stephan, artifact, interview, May 2014)

Introduction

In this manuscript we examine the stories of youth makers from non-dominant communities, and argue that through their making practice they are involved not only in “artifact making” (the prototypically viewed outcome of makerspace work), but also in space-making within and across the worlds of STEM, makerspaces, and community. Such space-making fosters new forms of interaction among scales of activity, and supports the movement of ideas, resources, relationships and bodies in support of youths’ emerging practices and how they might be recognized for them. As the youth engage in their making practice, they inscribe new meanings for what it means to make within the worlds they inhabit, refiguring participation in these worlds and the possibilities for becoming within them.

Our research questions are thus: 1) What making practice do youth from non-dominant communities, ages 11-14, take up in a afterschool community-based makerspace? 2) In what ways does their practice inscribe their spaces of making with possibilities for doing and becoming in making for community sustainability?

To answer these questions, we present two in-depth narratives of youths’ engagement in a community-based makerspace. The first case involves Samuel’s efforts to build the light-up football when he was in the 6th grade. The second case involves Jennifer and Emily’s efforts to prototype a heated, light-up sweatshirt in the 7th grade. While our telling of these narratives revolve directly around the youths’ making of artifacts, we hope to show how their practices for doing so alter the landscapes in which they work, and their opportunities to do and become within and across those spaces.

Findings indicate that the youth in our study have engaged in making practices that led to the creation of new artifacts that mattered to people in their communities. Further, youths’ making practices were undergirded in what we think of as “mobilities of criticality,” as they remixed and repurposed tools, practices and relationships from various communities towards space-making. In particular, we show how the youths’ making practices are rooted in community, and are reflections of their deep and critical knowledge of the needs their communities face within and across the spaces of making. We also discuss how the youth’s in-the-moment actions – a reflection of their making practices – served as critical “pivot points” in their design work (Holland, Lachiotte, Skinner, & Cain, 2001). The pivot points connected scales of activity, including STEM inquiry, making, community and action taking, in how they provided analytical foci for driving technically oriented design work, and opportunities for social negotiation towards new possibilities of doing and becoming in STEM, makerspaces and community.

Conceptual framework

We are interested in questions of youth engagement and identity work in making as it relates to how the spaces and places of making get re-organized, disrupted, and/or expanded through the youths' making practices. In particular, we are concerned with how youth's making practices take shape across multiple scales of activity simultaneously, but also over time – e.g., locally among peers in small group work in makerspaces as well as in the real and imagined spaces of STEM. Thus, we draw from mobilities of learning studies and social practice theory to frame our concerns. We are particularly interested in those studies that take a critical orientation, weaving in issues of power and positioning.

A mobilities of learning framework is centrally concerned with movement – of people, ideas, tools, resources, bodies and relationships – and how such movement shapes and reshapes the spaces and places of learning, and the social practices enacted and made possible therein (Leander et al, 2010). As individuals move through space and time, the sociohistorical narratives around them shift, shaping and reshaping how they inhabit or reinhabit space (Gutiérrez, 2012). For example, Kwan (2008) describes how Muslim American women's movements within public spaces have become restricted in the US since 9/11 in response to rising political narratives even though the actual physical access to these spaces remain unchanged. At the same time, she illustrates how such limitations in physical movement sit in juxtaposition with increasing access to new digital spaces. These arising digital spaces have become new homes for exposing oppressive narratives experienced by the women, as well as for opening up new opportunities for relationship building practices across cultural difference.

Examining the critical literacy practices of migrant youth in Southern California, Gutiérrez (2008) describes how youth use their “complete linguistic toolkit” – toolkits made up of linguistic practices of home and community, such as testimonio, in addition to the practices that are sanctioned in school settings – to navigate “the paradoxes of migration, immigration, and schooling” in the US. (p. 150). These hybrid practices helped students to link their past and present to an imagined future, and to reorganize everyday concepts acquired through social interaction in joint activities of school-based literacies. She suggests that these “rich interactional matrices of practice” lead to a new dialectic between the “the world as it is and the world as it could be” opening up new spaces for learning and transformation (p. 160).

In both of these studies, who individuals are and who they can be across the spaces –temporal, physical, and virtual – of their lives, and the practices they take up within and across those spaces, emerge from and transform the meanings of those space as constructed through social activity.

Mobilities of learning studies remind us that learning always takes place *somewhere*, both in “relation to history (time) and context (place/space)” (Bright, Manchester, Allendyke, 2013, p. 749). One thread of work that is particularly salient to our own work is that which examines space-making as a part of more expansive views of learning. We use the term space-making in ways similar to that of place-making (e.g., Cresswell, 1996, 2004; Massey, 2005; Lombard, 2014). An individual's opportunities to be and to become are shaped by place. At the same time, who one is also gives meaning to place; “places do not have intrinsic meanings and essences . . . the meanings of place are created through practice” (Cresswell, 1996, p. 17).

However, by drawing attention to spaces over places, we acknowledge the itinerant over the fixed nature of learning, where space reflects “a territory defined by practice-based learning, inhabited by a network of people, ideas, and objects in movement” rather than a fixed geographical area (Fendler, 2014, p. 787). We also use “space” to suggest that the possible platforms for being and becoming are not only solely contingent on the structural landscape of geographical places but are also tied to norms and power structures. “Space” also connotes for us the plurality of spaces (platforms for being and becoming) that are connected to a singular geographical place, e.g. youths' residential neighborhood.

Gutiérrez (2012) work on expansive learning helps to unpack the importance of movement across both vertical and horizontal dimensions of learning. Here, movement refers to the ways in which ideas, tools and practices are re-authored and re-mixed towards new possibilities for becoming in-practice across setting and over time (Engeström & Sannino, 2010). Through learning activity, new activity structures are produced as vertical and horizontal dimensions interact, leading to new forms of activity. Gutiérrez describes these new forms of activity as the kinds of hybridity that emerge as the tensions and contradictions that arise within and between activity systems. In these studies, *hybridity* refers to the novel combinations of different repertoires of knowledge and practice (e.g., science and peer/family/community) as individuals who horizontally move ideas and practices. However, it also refers to the hybridity that exists at multiple levels of the learning environment, where many activity systems come together (e.g., science, student, teacher, schooling, etc.). This perspective, thus, allows us to better understand youths' horizontal movement and hybridization toward making/engineering designs that are both meaningful from a disciplinary perspective and compelling to youth committed to their communities.

Methods

Our study was carried out as a critical ethnography over a two-year period. Critical ethnography was selected as our methodology because of its explicit focus on participatory critique, transformation, empowerment, and social justice. Critical ethnography is grounded in the idea that researchers can use the tools of ethnography to conduct empirical research in an unjust world in ways that examine and transform inequalities from multiple perspectives (Trueba, 1999). Critical ethnography provided an approach in which to “politicize” the interaction between actors and the social structures through which they act, grounded in the belief that these relationships are never neutral. This approach was important as we attempted to make sense of how youth, who are positioned in particular ways due to race, gender and class, engage in makerspace activities.

Our study is grounded in middle school youths’ experiences in two different makerspace contexts, Michigan and North Carolina, over the course of two and one years respectively. The makerspaces in both locations are housed in Boys and Girls Clubs [BGCs] (community-based clubs focused on youth development, homework help, and sports) in mid-sized cities, both facing some degree of economic depression. We have worked together with staff at the BGCs to establish these makerspaces, with the primary goals of supporting youth in developing productive identities in STEM, while also learning about making/engineering design in culturally sustaining ways. In both locations, we sought to engage youth iteratively and generatively in maker space activities and in community ethnography as one approach to embedding local knowledge and practice into making and engineering design.

Student and school sample

During 2013-2015, 36 youth participated, of whom 11 participated for 2 years (2013-2015), and the remaining 25 participated for 1 year (2014-2015). The youth were primarily from grades 5-8 (ages 10-14), and from lower-income families. Most are African American, although a few are white or biracial (See Table 1).

Table 1: Number of students in each site

Year	Location	Total Participants	Demographics
2013-2014	Michigan	14 youth	2 White 10 African American 2 Biracial
2014-2015	Michigan	21 youth	2 White (both returning) 17 African American (8 returning) 2 Biracial (returning)
	North Carolina	15 youth	14 African American 1 Biracial

Data were generated, 2013-2015, from artifacts, weekly youth conversation groups, and video analysis capturing youth interaction with STEM and community experts at various stages in their design process (See Table 2). In addition we used mid- and end- of year course artifact interviews, researcher field notes (per session), and youth created multimedia (e.g., video blogs) showing progress on their design to community members and STEM experts.

Table 2: Data Types Generation Strategies

Data Form	Specific Data Generation Strategy	MI (2yr)	NC (1yr)
Participant Observation	• Makerspace sessions/activities: Video recordings of twice weekly sessions and field notes in two sites	72hrs/yr	70hrs
	• Makerspace Community Events	8hrs	n/a
Conversation Group	• As a way to debrief what was happening in the club as well as to plan for future activities	30 hrs/yr	30hrs
Artifact	• Allowing youth opportunities to talk about their	4	3hrs

Think Aloud	engineering design work in detail (mid and end of year)	hrs/gp/yr	
Artifact Collection	<ul style="list-style-type: none"> Youth's sketch up notebook, 3D Google SketchUp model of design, worksheets, prototype, movie, etc 	ongoing	

Analysis

Data analysis involved multiple stages and levels of coding based on procedures for open coding and method of constant comparison (Strauss & Corbin, 1998). Our first pass involved reading through artifact interviews transcripts (conducted yearly at mid year and end of year) as well as our fieldnotes and the students' sketch-up notebooks kept during the course of their participation. The goal of this initial read through was to surface points and open codes of a) tensions and connections among the various youths' forms of engagement in making, b) critical design moments (e.g., sticking points, changes in direction, etc.), and c) generally how youth talked about and framed what it meant to participate. For example, in trying to open code for critical design moments, we noted times when youth made shifts in design, became deeply frustrated or disengaged, or otherwise more explicitly noted for us (e.g., artifact interviews) when they felt they were stuck or had important turning points. Weekly conversations were held between the authors on these insights as a way to work towards a more "expansive consensus"; that is to say that any differences in view were debated until new meaning was generated as a result of our differences. A detailed list of emergent open codes were kept with analytic memos attached to them, which we then brought to bear on other data sources, such as group conversation transcripts and various student artifacts not included in their sketch up notebook.

Our second pass involved identifying important resources and practices used by youth in their making, in relationship to the previously identified critical events, tensions and connections. With the help of our theoretical framework (mobilities of learning), we worked to make sense of what it meant for the youth to move, repurpose or remix the ideas, practices and resources they leveraged within these events. This axial phase of coding was used to uncover relationships and connections between the youths' making and the tensions that emerged from the data. In developing these coding schemes, we paid attention to how, and where, youth engagement appears greatest and the forms such engagement took, how they move ideas and resources across spaces, the different forms of learning, and the identity work that take place within and across these spaces. We took these data points as significant markers of equity – opportunities to access and activate traditional and nontraditional resources and to be recognized for doing so, as important to the making process and outcomes.

The relationships and connections identified in this second stage of coding, in turn, guided our selective coding, and became categories and themes, from which our example cases were selected for a final round of analysis and presentation. This final phase involved writing the narratives related to students' participation in the two makerspaces under study.

Findings

For the purposes of this proposal, we share one shortened narrative followed by a discussion of core findings.

In-depth vignette

- Interviewer: Samuel, why did you decide to make a light-up football?
- Samuel: Well, when little kids are playing outside football and it's getting too dark and they still keep playing and somebody might get hit in the head or something cause they can't see the ball really so I'm going to light up the football so you can see where it's going.

Samuel joined "M4C" a makerspace club at his local Boys and Girls Club during the fall of his 6th grade year. While he did not have friends in this club when he joined, and had never heard of "engineering" before, he said he wanted to join because he "kept seeing" what other kids were doing, and he wanted a chance to do "something like that" too. Samuel lived alone with his grandmother, after his mother ran into social and legal programs, a point that plays into his design work as we will see later.

Samuel designed a prototype of a "light up football" while working in the makerspace over a period of five months. His light up football has LED tube lights that wrap around the ball to provide maximum lighting with minimal added weight, friction, or power expenditures. Because the lighting is so efficient, it would also keep hands from getting burnt. The lights are powered with rechargeable batteries that can be recharged at a

solar docking station, limiting environmental impact and saving money. The football, itself, is constructed from nerf material to further minimize added weight and to reduce the possibility for injury if one were to be hit in the head. The batteries are stored in a pocket at the center of the ball, accessible by a small door, to keep it weighted properly and to minimize their potential contact with rain water and sweat.

The idea for a light up football grew out of Samuel's desire to make something that would be helpful to people in his community. As he states:

[My football] say about me that I really care about people. And I could, like, do stuff in the community so it could, like, do stuff together, like, peers can do stuff together, like, neighbors or school neighbors could like, go outside and do stuff together. . . Cause, like, some kids don't really play football, don't have no friends and stuff, so I go find people to help out a little bit.

Samuel's idea of care is nested in an understanding of the special needs of the young people in his community. Samuel knew that lighting was a concern at night due to limited working streetlights in his neighborhood. He also felt that the game of football was a positive peer activity that helped young people his age make friends and stay out of trouble. He knew that most families could not afford an expensive toy, and that inefficient designs were costly to the environment as well.

Samuel worked tirelessly on his design for five months seeking help from family, friends, and engineering and football experts alike. He was proud of his efforts. As he stated, "I was really proud 'cause it just made me feel good about myself so I could, like, acknowledge people what I could do. . . Like make what I did, a light-up football. I wanna make more stuff like that."

A light up football presented Samuel with many design problems of both technical and social consequence as well. For example, lighting a football requires power. As Samuel noted, powering the lights costs money. His initial solution was to use rechargeable batteries because "*mine's is rechargeable batteries so we can see all the time but so you won't have to keep going back to the store and buying, like, batteries to reuse.*" Saving both money and time by not having to return to the store to buy new batteries were both important in order for Samuel to keep the lights powered.

But, rechargeable batteries also addressed another design concern: environmental sustainability. Samuel was worried in particular about the problems created when non-rechargeable batteries are thrown into the trash. As he states, "Rechargeable batteries save energy and money so you won't have to, like, keep buying and buying batteries, so, [They] make the world greener. When you throw batteries away, those critters can get inside your trash, like the racoons, can like, take your batteries, take your trash and batteries out."

Powering the lights was, in fact, the "biggest design challenge" Samuel stated he faced. He noted that two batteries did not light the ball well enough, but more than two batteries, he felt, made the football too heavy, and too expensive. This insight seemed to spur Samuel to expand his design concerns to also include the weight of the football, and the location of the weight. Having his football like a "regular football" in terms of size, shape and weight, were all important, but all impacted by his desire to have a light up football.

Samuel sought input on these concerns from local football experts, which included a local football star. When recounting how these different experts helped him in his design, Samuel noted that the football star helped him to think about how to make the ball balanced, so that it would not be too heavy on one side or the other. To solve this problem, Samuel had to cut deeply into his prototype to place the batteries in the far center. As he stated, "Yeah. Yeah so I used that and so when I went back and tried to do it, I made sure that when I cut it, I made sure that it could be deep enough so it won't, like, make it so heavy. So it could be, not be so light, it could be just right. So *like a real NFL football.*"

There were many other design challenges that Samuel confronted as he progressed in his project. He sought out many different kinds of experts to help him out, from his mother to his peers, to football experts, to engineers, to the internet. All of these perspectives representing different kinds of expertise and needs mattered to him.

Discussion of findings grounded in vignette

Youths' making practices are rooted in community, and are reflections of their deep and critical knowledge of the needs their communities face within and across the spaces of making. Youth's in-the-moment actions – a reflection of their making practices – served as critical "pivot points" in their design work. The pivot points connected scales of activity, including STEM inquiry, making, community and action taking, in how they provided analytical foci for driving technically oriented design work, and opportunities for social negotiation towards new possibilities of doing and becoming in STEM, makerspaces and community.

Rooted in community

Practices as rooted in community. The youth in our study, in on-going ways, position themselves as a part of, or inside, the urban ecology, rather than outside of it. Their making practices, as rooted in a wide range of community spaces, draw upon expert knowledge on issues inside to these spaces, such as the funds of knowledge one has because of where they have grown up and with whom. These practices also incorporate insider positioning status, such as that which grants access to the social networks and contexts necessary for gaining deeper insights and access to resources when needed.

The youth brought to their investigations a wide range of funds of knowledge and relationships that played a role in how they defined the problem they wanted to work on, and its various dimensions. These funds are tied to particular community spaces where youth have insider status. For example, knowledge of where streetlights have historically not worked, why kids at their school get bullied, fashion, how to work with one's hands to build, or the reasons and impacts of major economic concerns of the home, all reflect their insideness – their membership and experiences in the community spaces that they inhabit. How the youth drew from these funds across spaces reflect their attempts to author interconnecting corridors for traversing between these community spaces, and their STEM-infused youth makerspace. These different points of intersection become meaningful sites of negotiation.

Practices as enactments of their deep and critical knowledge and care for the needs their communities face. There are nodes of criticality in many of the funds of knowledge deployed by the youth, and in how they sought to leverage these funds towards engaging more deeply in the technical dimensions of their work. All communities face risks that result from geographical, socioeconomic, and political challenges. However, the risks are greater for young people growing up in lower-income communities of color, where environmental and social injustices loom large.

We see such criticality enacted by these youth in their making work as tied to four domains in particular: *Economic* (e.g. making their designs affordable), *environmental* (e.g., designs that reduce their communities carbon footprint and support local ecologies), *social* (e.g., fostering positive peer relationships, healthy well-being, community ownership, and preventing bullying and gang activity), and *urban infrastructure* (e.g., providing lighting and warmth on cold, dark days).

For example, Samuel worried about dangerous peer friendships, such as gangs, and believed some of these peer-related challenges might be remedied with positive play, such as with football. Samuel persisted in refining his football so that it met the needs of a wide range of peers. He first sought peer input on lighting – weight and design. He then pushed for input on weighting and feel. He tested his football with peers his age and peers younger than him. He pressed them for feedback on what functionality they needed, which is why he ultimately sought to make sure his ball was waterproof. Each interaction required Samuel to consider many new technical factors in his design that he had not previously considered, but he was deeply motivated by how and why his football would serve his local peer community.

Pivot Points and their functions

As youth engage in such rooted making practices over time, their in-the-moment actions served as critical “pivot points” in their design work. The pivot points connected scales of activity, including STEM inquiry, making, community and action taking, in how they provided analytical foci for driving technically oriented design work, and opportunities for social negotiation towards new possibilities of doing and becoming in STEM, makerspaces and community. Here we refer to Holland and colleagues (2001) use of the term pivot to refer to “mediating or symbolic devices” not just to “organize responses but also to pivot or shift into the frame of a different world”. When youth leveraged their funds of knowledge, for example, towards work on their projects, they etched their insideness onto their engineering design, in ways that impacted the design process and how/where it unfolds, as well as their role in it. As pivots, these funds were not simply complementary to the youths’ engineering design, but *essential to both who they are and their design work*. Pivot points include tools (e.g. sewing machine, Google Sketch Up), relationships (e.g. Samuel’s ties to his cousins and peers) and the innovations themselves (e.g. Samuel’s light-up football), all of which are able to shift the nature of STEM engagement for the youth, and potentially transform their possibilities for becoming and being within particular spaces (e.g. Samuel’s peers and cousins engaging in safe play at night in their neighborhood).

The three key functions of pivot points are 1) Using funds as navigational indicators to secure a productive launching space to begin their making project; 2) driving technically oriented design work in dialog with community interests, and 3) facilitating social negotiations towards novel space-making endeavors to broaden possibilities for becoming in STEM.

1) Using funds as navigational indicators to secure a productive project launching space. The youth (including Samuel) took on a making project because they belonged to the same makerspace club. They were

not told what to make or how to make it. Rather they were charged with a fairly wide open task: design something that “uses portable energy” and that “attends to a particular community concern.” We have been concerned with how youth locate or author productive starting places for projects. Such initial location work can be challenging, for it involves social negotiations of who to work with, along with considerations of what challenges might be worth spending time on. In both cases presented in this manuscript, the youth leveraged their funds as navigational indicators to author a productive project launching space.

Samuel drew from a wide range of funds – M4C youth makerspace, family, peer, and residential community funds – in order to locate a productive project launching space. As Samuel noted in his interview, his light-up football was an idea he “thought, and thought and thought about” while home at his grandmother’s house unable to find transportation to the club nor able play outside after the dark. These considerations – limited streetlights, personal safety, and friendships – made that much more salient by his move to his grandmother’s care became points of negotiation for how and with whom he would work. Samuel switched group memberships twice, before he felt he had the space to tackle the issue he really cared about. Samuel’s initial ideas were legitimized by his cousins and peers who also knew that he had extensive experience playing football with friends, and had the expertise to design a football.

2) Driving technically-oriented design solutions in dialog with community interests. We also see imprints of youth’s rootedness in how they worked across scales of activity in their systematic efforts to refine their design constraints and evaluate possible solutions towards optimization. New design cycles were initiated on both technical and community terms. For example, youth-set end-point assessments required them to seek multiple perspectives, both community-oriented and technical. As community funds initiated more complex design conditions, Samuel needed to deepen his knowledge of energy systems and environmental and economic impact. Working with a mentor, Samuel figured out how to calculate power requirements of different lighting systems. He read information on the Internet on the affordances of LED lights, when his friends told him that bulky lights would not work on the football. He spent time figuring out how to assemble the components in a circuit with a switch. When he became concerned about the affordability of batteries as well as the impact on battery disposal on the local ecology, he thought about rechargeable batteries. But even then he had to figure out how to charge the rechargeable batteries in environmental friendly ways.

As Samuel began working on his design, he brought to bear a set of fairly *specific* community concerns, as discussed, to a fairly *vague* technical challenge. However, these specific community concerns helped Samuel to functionally break down the work he needed to accomplish from a technical standpoint – e.g., work on the lighting, the weight, and the shape of the football. Samuel’s funds of knowledge also gave him starting points for where, within his social network, he might look for feedback.

3) Facilitating social negotiation towards novel space-making: New possibilities of doing and becoming in STEM, makerspaces and community. The youth’s making practices iteratively and incrementally built on each other to expand their STEM expertise and rootedness in community. Both the merging and layering of STEM and their funds of knowledge onto and into each other were accomplished not only in the service of their design work, but also *in the attempt to change the real worlds in which they are working and becoming* (Fendler, 2014). That some of the youth have said they want to get smarter on these topics so that they can return to their community – not leave it as they move on up – speaks to this point well. In short, new possibilities in space-making operate both at the level of the making process, and the potential resultant impact of the youth-created innovation.

The playing field *in makerspaces (one area of space-making)* literally and figuratively transformed for the youth as they incrementally, but systematically, refined the problem and design they were working on in both technical and social ways, expanding their connectedness to others, and thus the access they had to ideas, tools, and resources for advancing their developing expertise. As Samuel walked through the main club rooms with his ball, kids gathered around him asking to test it out, and where and when they could buy one. His picture with the pro football star hangs on the wall, and other youth have since joined M4C to “do what Samuel did”, which included “making” things and “meeting famous people.” Becoming an expert involved a form of vertical development in his deepening scientific knowledge and practice, but it also involved interactive movement and engagement of such developing expertise with his cultural repertoires of practice (Gutiérrez, 2012).

What is more, the youth’s designs helped to transform the playing field *in community (another area of space-making)* for themselves and their peers. Samuel’s football will allow his peers to practice throughout the off-season, so that they can “get better at the - for next year - for the next season they play football.”

Equally as significant are the *playing fields of STEM, both real and imagined* (a 3rd area of space-making). The youth’s practices served as new tools to expand the very purposes and goals for engaging in science. At the heart of each youth’s design is an effort to work at the intersection of science and the public good, as a way to transform both. Their engagement with the problem was not simply motivated by individual

interest. Engagement was framed, in part, through collectively formed interests as they sought out feedback from a wide range of others, at the powered boundaries of race, power, care and danger. These tensions required greater engagement with STEM as they demanded more complex problems to be considered. At the same time, these tensions made possible recognition within STEM worlds while also exposing the challenges youth face in seeking recognition in these same worlds.

This kind of repositioning in STEM amidst these tensions is important. Dominant discourses position the youth as outsiders and non-experts in science and engineering. What we see is a reinhabitation of the spaces of STEM; one that deterritorializes STEM routines and practices, making as Perumal (2015) writes “physical entry into and living in previously forbidden places” a process of taking back and reclaiming the space of STEM in ways that recognize and care for the rootedness of young people (p. 26).

Conclusions and implications

For many of the youth with whom we work, gaining access to STEM is an uphill battle. We also see how the youth pushed back against these low expectations through engaging in mobilities of criticality. Their making practices, as rooted in community, allow for the reconstruction of the spaces of STEM, making and community, in how the movement of people, ideas, and relationships interrupt practices and ways of being.

The youths’ work suggests that leveraging both community insideness and scientific expertise is about much more than bridging these two worlds. While such bridging is important, it is in how this bridging makes possible new and more expansive opportunities to learn and to become in STEM, that we need to pay increasing attention to. Indeed, by engaging in mobilities of criticality, the youth speak back against the normative accounts that often frame their communities in deficit ways. At the same time, their enactment of their criticality through their making practices call attention to the prosaic and micro-level processes involved in making spaces – STEM, maker, and community – more habitable (Lombard, 2014).

References

- Bright, N., Manchester, H., Allendyke, S. (2013). Space, Place, and Social Justice in Education: Growing a Bigger Entanglement. *Qualitative Inquiry* 19(10) 747–755.
- Creswell, T. (1996). *In Place/Out of Place: Geography, Ideology, and Transgression*. Minneapolis, MN: University of Minnesota Press.
- Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*, 5(1), 1-24.
- Fendler, R. (2014). Becoming-Learner: Coordinates for Mapping the Space and Subject of Nomadic Pedagogy. *Qualitative Inquiry*, 19(10) 786-793.
- Gutiérrez, K. (2008). Developing a Sociocritical Literacy in the Third Space. *Reading Research Quarterly*, 42, 148-164.
- Gutiérrez, K. (2012). Re-Mediating current activity for the future. *Mind, Culture, and Activity*, 19, 17–21.
- Holland, D., Lachiotte, D., Skinner, D., & Cain, C. (2001). *Identity and agency in cultural worlds*. Cambridge, MA: Harvard University Press.
- Kwan, M.-P. (2008). From oral histories to visual narratives: Re-presenting the post-September 11 experiences of the Muslim women in the United States. *Social and Cultural Geography*, 9, 653–669.
- Leander, K., Phillips, N. & Taylor, K. (2010). The changing social spaces of learning: Mapping new mobilities. *Review of Research in Education*, 34, 329-394.
- Lombard, M. (2014). Constructing ordinary places: Place-making in urban informal settlements in Mexico. *Progress in Planning*, 94, 1-53.
- Massey, D. (2005). *For Space*, London: Sage.
- National Academy of Engineering. (2010). *Engineering, Social Justice, and Sustainable Community Development*. Washington DC: National Academies Press.
- Perumal, J. (2015). Critical pedagogies of place: Educators' personal and professional experiences of social (in)justice. *Teaching and Teacher Education*, 45, 25–32.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd Ed)*. Thousand Oaks, CA: Sage Publications.
- Trueba, H.T. (1999). *Latinos unidos: From cultural diversity to the politics of solidarity*. Lanham, MD: Rowman & Littlefield.

Acknowledgments

We thank participating youth. This work was funded by the National Science Foundation, DRL 1421116.

Supporting Teachers in Navigating Change Towards Science Practices Focus in the Classroom: Investigating Current Teacher Support for Science Practices

Nicole D. Martin, University of Wisconsin – Madison, ndmartin@wisc.edu
Sadhana Puntambekar, University of Wisconsin – Madison, puntambekar@education.wisc.edu

Abstract: A growing emphasis on teaching science as practice calls for a significant change from traditional science education. We must provide teachers with the appropriate support to navigate this change. The first step in supporting teachers to help students learn science practices is understanding current approaches for teaching science practices. We investigated the question: How do teachers support science practices in their classrooms? We analyzed how two teachers supported students by engaging and guiding students to participate in science practices; we focused on two key science practices: constructing and defending scientific explanations and analyzing and interpreting data. We found that teachers most frequently engaged students in practices but did not necessarily provide guidance for how to participate, explain why science practices are important, or describe how all the practices are connected. Teachers may need additional guidance to develop concrete teaching strategies to support science practices.

Keywords: science teaching, science practices, teachers as learners

Introduction

Science as practice: Teaching and conceptualizing science as practice has recently emerged as a prominent perspective in science education. This perspective views science as progressing beyond a reasoning process to encompassing how we understand, make sense of, evaluate, and represent the world around us (Lehrer & Schauble, 2015). While previous views of science as a reasoning process have focused on the logical construction of experiments, manipulation of variables, and interpretation of data (Lehrer & Schauble, 2015), the view of science as practice emphasizes a broader focus on the holistic process by which knowledge is constructed in the scientific community. This view of science is central to the new emphasis on science practices in education. Science practices span the entire process of science, from being able to generate and write a hypothesis, collect data, analyze data, understand how the data supports the hypothesis, and draw conclusions from the data and interpret the findings. The importance of these practices has been placed at the forefront of science education, as seen by their inclusion in the Next Generation Science Standards (NGSS). These new standards restructure previous attempts to shift science education towards inquiry-based, hands-on-science by “elaborate[ing] how to engage in the work of inquiry, and how this work is part of building knowledge” (Reiser, 2013, p. 5). More specifically, the NGSS break down the practice of science into eight components: asking questions; developing and using models; planning and carrying out investigations; analyzing and interpreting data; using mathematics and computational thinking; constructing explanations; engaging in argument from evidence; and obtaining, evaluating, and communicating information (NGSS Lead States, 2013).

While these practices are often presented as a list, they should not be conceptualized as step-wise, sequential skills. Rather, all the science practices are interdependent components of a single system of making sense about the world, and they must be viewed as such in order to fully understand the practice of science (Reiser, 2013; Lehrer & Schauble, 2015). This goal necessitates an integrated and holistic view of the science practices that build on one another, describing the process and reasoning of science. Teaching these practices in isolation would remove them from the context of the whole process by which the scientific community constructs knowledge; students cannot fully understand science as practice without a holistic approach to integrating science practices in the classroom. Students must learn to: ask scientific questions that drive the development of their investigations; analyze and interpret data from their investigations; and use this data to construct models and explanations about the phenomenon they are investigating. Ultimately, students must be able to develop convincing, well-supported arguments that use their data and explanations to answer their research questions and communicate their findings to the scientific community.

The science practices should not be conceptualized as a sequential list; however, it may be useful to think of them as hierarchically organized. Lehrer and Schauble (2015) advocate for such a hierarchical approach and place modeling as an over-arching, governing practice that is most central to the practice of science and supported by the rest of the practices. Others emphasize that engaging in argument from evidence is a central practice of science by which scientific knowledge is generated and learned (Berland & Reiser, 2009; Reiser, Berland,

Kenyon, 2012; Driver, Newton, & Osborne, 2000). We follow this lead and view engaging in argument from evidence as a culminating practice that unifies all of the science practices. Most highly interrelated in the hierarchical structure, engaging in argument from evidence is contingent on students' ability to analyze and interpret data and to construct explanations.

Changing teacher practice to support science practices: The growing emphasis on the importance of teaching science practices in the classroom calls for a change from traditional science education. The science practices we are now asking teachers to incorporate in their classrooms require significant changes in *teacher practice*, for which we must provide teachers with the appropriate guidance and support (Reiser, 2013, Windschitl, Thompson, Braaten, & Stroupe, 2012). The *teacher* in the classroom plays a hugely important role in how the curriculum is enacted (Remillard, 2005). Remillard (2005) emphasizes the importance of the relationship between the teacher and curriculum; she highlights that the curriculum that is actually enacted in a classroom is a complex combination of the formal curricular materials and the intended curricular aims of the teacher. Acknowledging the active roles teachers play in the enacted curriculum focuses attention on ensuring that teachers are well prepared and properly supported to create effective learning environments when they implement new curricula. We cannot expect a new way of conceptualizing science as practice to be successfully enacted in the classroom without actively supporting teachers in implementing science practices. We can design curricula to support students' learning of science practices, but this is not enough—we must understand *how to help teachers* support their students' learning of science practices

Tabak and Radinsky (2015) acknowledge that research on teaching is typically absent in the field of learning sciences and call for exploration in this domain as “a unique target for research on learning” (p. 345). We agree and seek to answer this call. In this paper, we begin to explore this question of how to best help teachers support students' learning of science practices in the classroom by first exploring how teachers currently support their students. In order to effectively help teachers navigate the shift to centering science education on science as a practice, we must first understand their current teaching approaches related to science practices. In the present study, we investigated how teachers supported science practices in the classroom. We focused on the central, overarching, and high-leverage (Reiser, 2013) practice of engaging in argument from evidence and the next most essentially related practices of analyzing and interpreting data and constructing explanations. We further focused on the holistic view of the science practices, which sees all of the science practices as interrelated and interdependent, emphasizing the relationship between practices. We aimed to answer the research question: How do teachers support science practices in their classrooms?

Particularly, we focused on the following mechanisms of support: opportunities to engage in practices (e.g. Driver, Newton, & Osborne, 2000; Cavagentto, 2010); guidance for how to participate in practices (e.g. McNeill, Lizotte, Krajcik, & Marx, 2006; Sandoval & Reiser, 2004); guidance for why science practices are important (Sandoval & Millwood, 2005); and guidance for how the practices are connected to one another (Erduran, 2014). Previous research has found these ways of supporting science practices to be beneficial for students; however, the large majority of these supports and forms of guidance have been material-based tools and interventions. There is much less empirical research on guidance coming from *teachers* to support science practices and even less research on how to help improve teachers' support for these practices in their classrooms. To address this gap in the literature and our research question, we used classroom observations of two teachers to investigate whether teachers simply had students engage in these practices as activities or if they actually provided guidance and supported students so they could learn *how* to approach analyzing and interpreting data, constructing explanations, and engaging in argument from evidence. We further investigated whether teachers discussed with students *why* these are important practices- why it is important as participants of science to analyze data, construct explanations, and use data to support explanations to create strong arguments. We finally examined whether teachers expressed the holistic nature of the practices to their students and made connections between the practices in the classroom to help students see how the practices are interdependent.

Methods

Participants and context

The participants in this study were two sixth grade science teachers, Mrs. Lloyd and Mr. Gordon. Both teachers taught at the same public middle school in a mid-sized, US Midwestern city. Mrs. Lloyd and Mr. Gordon had 13 and 23 years of teaching experience respectively, and they had both been working with the design-based, inquiry CoMPASS physics curriculum (Puntambekar, Stylianou, & Goldstein, 2007) used in this study for seven years. Their previous work with this curriculum is important because their experience entails that the teachers were familiar with ideas of having students justify claims, use data, and write explanations since these practices are embedded and emphasized throughout the curriculum. We chose two teachers in order to conduct a detailed, in-

depth analysis of the teachers' actions in the classroom over time. We chose to analyze two teachers instead of conducting an individual case study so we could compare differences between teachers. This was a ground-up study through which we wanted to create a good baseline for understanding teachers' support for science practices, as little is known about this question.

The curriculum was a 4-week design-based curriculum in which students learned the physics concepts of force, mechanical advantage, work, and energy by investigating how two different simple machines could help them accomplish challenges with ease. The students investigated how pulleys could help them solve the challenge of lifting a statue of their school mascot up to its pedestal and how incline planes could help them solve the challenge of getting a pool table into the back of a truck. Students worked together in small groups of three or four to conduct experiments using virtual computer simulations and to research the physics of simple machines using the CoMPASS eTextbook. The student-centered and inquiry-based nature of this curriculum positioned the teacher as a facilitator for student learning instead of as an authoritative source of information. The intended role of the teacher was to help facilitate discussion among students, monitor and assess students' conceptual understanding, and (most importantly for the context of this study) guide students to think like scientists and engage in the practice of science.

Data sources

We used classroom observations of the teachers implementing the physics curriculum to investigate how teachers supported science practices through pedagogy in their classrooms. All classes were video- and audio-recorded to analyze how teachers supported and discussed science practices involved in arguing from evidence in their classrooms. We followed Mrs. Lloyd and Mr. Gordon through the entirety of the 4-week unit in one of their science classes to capture how each teacher supported students' learning of science practices throughout the process of generating scientific questions, writing hypotheses, conducting experiments, analyzing and interpreting data, and constructing and defending explanations for final decisions to solve students' challenge. This resulted in seven videos (363 minutes) for Mrs. Lloyd and eight videos (358 minutes) for Mr. Gordon.

Analysis

We used a two-dimensional coding scheme to capture how the teachers supported students to learn science practices in the classroom. We focused on the central, overarching, and high-leverage practice of engaging in argument from evidence and two other essentially related practices: analyzing and interpreting data, and constructing explanations. Given the practical difficulty of distinguishing the practice of constructing scientific explanations and engaging in argument from evidence (Berland & Reiser, 2009), we discuss these two practices together as constructing and defending scientific explanations. This resulted in two science practice codes: 1) analyzing and interpreting data and 2) constructing and defending scientific explanations. We first coded the classroom videos for instances of these two science practices.

We then coded each of these instances for the type of support the teacher provided in order to better understand the quality of *how* teachers supported their students. We used four types of teacher support: a) engaging students in practice, b) guiding students in practice, c) discussing epistemic importance of practice, and d) connecting practices to present science as a holistic process. These support types distinguish whether the teachers were simply having students engage in the practice or if they were actively providing support to help teach students *how* to participate in these practices. These types of support also illustrate whether the teachers were talking with students about why these practices are important in science and whether the teachers were presenting science as a holistic process by making explicit connections between science practices. These support types are not to be considered in a rank order. While simply having students engage in a practice (a) does not provide the guided participation that students may need for learning how to participate in science as a practice, the remaining three types of support (b, c, d) are not necessarily increasing in quality or value. Table 1 illustrates what each of the four types of support codes would look like for the two science practices codes.

We segmented the videos into two-minute intervals for consistency of coding and provision of enough time for development of distinct instructional moves. We coded for the presence or absence of the types of support the teachers provided for the two science practices of focus within each segment. A second researcher helped us code 10% of the videos for each teacher and achieved 88% agreement overall and "substantial" or "almost perfect" kappa values (Stemler, 2001) for all but two of the individual codes. All disagreements were resolved through discussion and the first author coded the remaining videos.

We first quantitatively analyzed the overall patterns of support each teacher provided throughout the unit, considering both science practices under investigation together. To do this, we divided the total frequency of each type of support code by the total number of two-minute segments for each teacher to standardize frequencies across unequal video lengths. It is important to note that each two-minute segment could contain multiple codes

if multiple types of support or science practices occurred. We then compared the frequencies of support provided by the two teachers for the practices independent of each other. For this analysis, we conducted a Chi-squared test of homogeneity of proportions to compare the total proportion of each type of support for each science practice. We looked at the practice of analyzing and interpreting data and the practice of constructing and defending scientific explanations as independent practices in order to investigate differences in the support of science practices.

Table 1. Description of types of teacher support codes for both scientific practices investigated

		Type of teacher support codes			
		Engaging students in practice	Guiding students in practice	Discussing epistemic importance of practice	Connecting practices to present science as holistic process
Scientific practice codes	Analyze and interpret data	<ul style="list-style-type: none"> - Giving students opportunities or telling to look for patterns or relationships - Asking what patterns or relationships students see in their data or how one variable affects another 	<ul style="list-style-type: none"> - Pointing to parts of data table or directing attention to specific parts of data table to highlight patterns and/or relationships between variables - Directing students on how to look at data tables to find patterns 	<ul style="list-style-type: none"> - Discussing how scientists look for patterns and relationships in data to understand and make claims about how concepts are related - Discussing purpose of using data to support claims and answer research questions 	Making explicit connections between analyzing and interpreting data and: <ul style="list-style-type: none"> - Previous hypothesis; deciding if hypothesis is supported or not - Research question or hypothesis; deciding where to look in data chart based on question or hypothesis - Understanding data allows you to construct arguments based on that data
	Construct and defend scientific explanations	<ul style="list-style-type: none"> - Giving students opportunities or telling to provide (write or say) an explanation or argument - Promoting students to explain "why" - Asking open-ended questions that require students to give explanation 	<ul style="list-style-type: none"> - Showing students how to use appropriate data to support claims - Helping/prompting students to look at data to construct explanation/argument - Giving hints about what students should include in an explanation or argument 	<ul style="list-style-type: none"> - Discussing that scientists must justify or prove their statements (explanations) with data/evidence - Discussing the need to convince or persuade others (who may not be familiar with the data) of explanations by using evidence to back up claims 	Making explicit connections between an explanation or argument and: <ul style="list-style-type: none"> - Research question or design challenge you are trying to answer/solve - Previous hypothesis - Understanding data allows you to construct arguments based on that data

Results

Overall patterns of teacher support

To examine the types of support teachers most frequently provided for both science practices, we analyzed the total proportion of each type of support, looking at both of the practices together. Both teachers showed the same pattern in the type of support they provided students to learn science practices in the classroom.

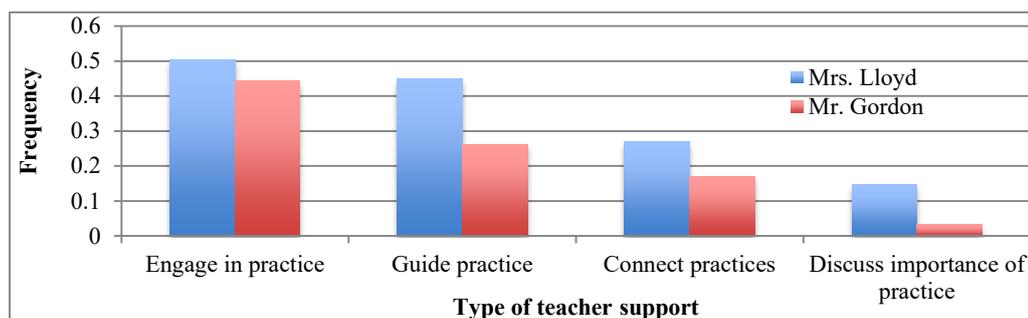


Figure 1. Frequency of different types of support teachers provided overall for both science practices.

Both Mrs. Lloyd and Mr. Gordon’s most frequent type of support was *engaging* students in the practice (0.50 and 0.44 respectively). Their second most frequent type of support was *guiding* how to engage in the practice (Mrs. Lloyd = 0.45, Mr. Gordon = 0.26), followed by *connecting practices to present science as a holistic process* (Mrs. Lloyd = 0.27, Mr. Gordon = 0.17). Both teachers least frequently *discussed the epistemic importance* of the practice (Mrs. Lloyd = 0.15, Mr. Gordon = 0.03). This common pattern of the type of support provided for students to learn science practices in the classroom can easily be seen in Figure 1. This figure clearly shows how both teachers (from most frequently to least frequently) engaged, guided, connected, and discussed importance of practices.

Comparisons between teachers

To better understand how the two teachers supported the science practices in their classrooms, we compared how often the teachers provided each type of support to see if there were any differences in their teaching practices. We analyzed the practices independently to further tease apart differences in support for the individual practices (see Table 2). We have organized the following comparisons by the type of teacher support code for clarity and included excerpts of teachers’ discourse to help illustrate the support teachers provided.

Table 2. Proportions (and frequencies) of teacher support across the unit and between teacher comparisons

Science Practice	Type of teacher support	Mrs. Lloyd	Mr. Gordon	Z score
Analyzing and interpreting data	Engaging students in practice	0.20 (37)	0.16 (30)	0.90
	Guiding students in practice	0.18 (34)	0.09 (17)	2.55 *
	Discussing epistemic importance of practice	0.06 (12)	0.03 (5)	1.73
	Connecting practices	0.14 (26)	0.08 (14)	1.99 *
Constructing and defending scientific explanations	Engaging students in practice	0.30 (56)	0.28 (51)	0.51
	Guiding students in practice	0.26 (49)	0.17 (31)	2.24 *
	Discussing epistemic importance of practice	0.08 (15)	0.01 (1)	3.64 *
	Connecting practices	0.13 (24)	0.09 (17)	1.13

* $p < .05$

Engaging students in practice

Mrs. Lloyd and Mr. Gordon spent the same proportion of time *engaging* students in the practices of analyzing and interpreting data ($z = 0.90, p = .37$) and constructing and defending scientific explanations ($z = 0.51, p = .61$). Engaging students in these practices involved giving the opportunity to and prompting students to analyze and interpret their data or construct and defend a scientific explanation based on their experiments in class. Mrs. Lloyd often engaged students in analyzing and interpreting data by asking students to talk to a neighbor about any patterns or relationships they saw in their data and then having students share their ideas with the class. Mr. Gordon similarly engaged students in analyzing and interpreting data by frequently asking students what patterns they saw in their data or telling them to look for these patterns. For example, he said: “What are you guys seeing so far? What kinds of patterns?” and “You should be looking at your data and discussing the patterns that you see.” In one example of engaging students in the practice of constructing and defending scientific explanations, Mrs. Lloyd gave students time to write their explanations on a designated page in their notebooks and said: “Ok, so I’m going to let you try this. It says, ‘We found that,’ well, what did we find...and then back it up here in the second part. So let’s work, see what you can do on your own.” All of these excerpts show examples of how both teachers gave students opportunities to engage in science practices without necessarily providing any additional instruction or guidance. These instances simply involved students “doing” each practice.

Guiding students in practice

Mrs. Lloyd spent significantly more time than Mr. Gordon *guiding* the practice of analyzing and interpreting data ($z = 2.55, p < .05$) and *guiding* the practice of constructing and defending scientific explanations ($z = 2.24, p < .05$). The following excerpts exemplify how Mrs. Lloyd guided students on how to participate in both of these practices. The first excerpt illustrates how Mrs. Lloyd guided students to analyze and interpret data. She had asked students to discuss patterns in their data, and in this excerpt, she provided explicit guidance on how they could look at their data charts:

See if you notice, because you guys tested out, 1, 2, 5 different kinds of pulleys. And as we went down this row [of the data chart], we added more, like, wheels to the pulleys – they became more complex. So do you notice anything as you sort of work your way down the chart? [Mrs. Lloyd points to a data chart projected on the classroom white board and moves her hand down each column] Do you see any patterns of things as we added more pulleys to our system?

Instead of only telling students to find patterns in their data, she directed students to consider looking down specific columns in their data chart in order to help them find a meaningful relationship. The second excerpt illustrates how Mrs. Lloyd guided students on constructing and defending scientific explanations. In this example, students practiced constructing an explanation in their notebooks about the relationship between height and potential energy and defending their explanation with data from their experiment:

We are looking at number two that compares height to potential energy. So for the first part where it says 'We found that,' you should explain what we found, that more height, what happened to potential energy? [*Students: 'Increased'*]. It increased. So you talk about that. Then you give data, at least two pieces of data. 'Cause we lifted our mascot to two different heights. So show an example at .1 meters and an example at .2 meters so you have data to back up what you are saying.

Mrs. Lloyd helped students see that they needed to state the relationship they found between height and potential energy, and she explicitly guided them on how to include multiple pieces of data to defend their ideas.

Connecting practices to present science as holistic process

Mrs. Lloyd spent significantly more time than Mr. Gordon *connecting* the practice of analyzing and interpreting data to other science practices ($z = 1.99, p < .05$), but both teachers provided similar amounts of support for *connecting* the practice of constructing and defending scientific explanations, ($z = 1.13, p = 0.26$). The following excerpt is representative of how Mrs. Lloyd made connections between analyzing and interpreting data in students' data charts to the overarching challenge and research question students were trying to solve: "Or also think about our challenge. Remember we wanted to increase our mechanical advantage and decrease our force. So what do you notice about mechanical advantage and maybe force. Take a look at those columns as well." Mrs. Lloyd helped students understand how their research question influenced where they should look in their data chart in order to analyze the data with a purpose; Mr. Gordon rarely made such connections. The next excerpt exemplifies the support both teachers provided for making connections to constructing and defending scientific explanations. As students constructed explanations about an experiment, Mr. Gordon had them return to the hypotheses they wrote before the experiment and said: "Look at what you wrote [for the hypothesis], and then go back and look at your data, and then go to your report out and circle whether you were confirmed or not. Then, we found out what? What did we find out about height and PE?" Mr. Gordon explicitly made the connection between the initial hypothesis about a research question, data analysis, and constructing and defending an explanation to help students see that these components were all related and in service to the scientific explanation.

Discussing epistemic importance of practice

Mrs. Lloyd and Mr. Gordon spent similar amounts of time *discussing the epistemic importance* of analyzing and interpreting data ($z = 1.73, p = .08$), but Mrs. Lloyd spent significantly more time than Mr. Gordon *discussing the epistemic importance* of constructing and defending scientific explanations ($z = 3.64, p < .05$). Mrs. Lloyd helped students understand the purpose of interpreting data to understand relationships and emphasized the importance of using data to support claims they present to other people; Mr. Gordon offered similar support by discussing the purpose and importance of interpreting data to support students' explanations: "Just having numbers [data] isn't enough. Explaining it means how does one thing affect the other...Did you put enough data to explain why?...Is the data you wrote enough to demonstrate fully your understanding?" Mr. Gordon tried to help students understand the purpose of interpreting data to show how science concepts were related and that including data in their explanation was important to demonstrate their understanding. For constructing and defending scientific explanations, Mrs. Lloyd most often discussed with her students how scientists must convince others of their ideas by constructing strong arguments using data:

"Now the next question says, 'We know this because.' How do we know this? Notice it says, 'Refer back to your data chart.' So this is one thing as a scientist you have to do. If you tell somebody something, they are gonna say, 'So what? Prove it! Prove it to me!' Now do we have data that can prove this? [*Students: 'Yes.'*] Yeah, we have a lot of data that can prove this. So when you are reporting out, you have to make sure that you include some data to show that what you are saying really, really is indeed true. It really is indeed true."

In this example, Mrs. Lloyd talked with students about the purpose for why they needed to use data to justify and defend their claims and emphasizes how important this practice was in science in order to prove their statements to other people. Mr. Gordon, on the other hand, rarely discussed this purpose with his students.

Discussion and implications

Given the increasing focus on science as practice in education, it is essential to understand how to help teachers navigate this shift and implement this focus in their classrooms. A first step in this process is assessing the current state of teachers' approaches regarding science practices. The goal of this paper was to investigate how teachers currently support students to learn the high-leverage and most central science practices in the classroom on a daily basis. We first review and discuss the findings from this study with respect to this goal. We then offer implications for supporting teachers as they continue to shift their classes to center on science practices and offer avenues of future research to address yet answered questions.

Overall, both teachers showed the same pattern of supporting students to learn how to analyze and interpret data and how to construct and defend scientific explanations. In order of greatest to least frequency, the teachers engaged students in practices, guided their participation in practices, connected practices, and discussed the importance of practices. While both teachers spent a similar amount of time engaging students in the practices of analyzing and interpreting data and constructing and defending scientific explanations, one teacher spent more time providing students with guidance on how to participate in these science practices, making connections between science practices that conveyed the holistic process of science, and discussing why these practices are important in science. The fact that both teachers in this study most often engaged students in the practices is consistent with previous ideas that teachers are not well prepared to support students to learn science practices in the classroom (Reiser, 2013). While one teacher appeared to use additional support strategies while engaging students, the other teacher most often engaged students with no additional support. These findings suggest that teachers do not necessarily use strategies beyond giving students opportunities to engage in science practices in order to help them learn. Giving students opportunities to participate in practices is an important part of helping students become active participants in the practice of science (Lave & Wenger, 1991), but other researchers suggest that having opportunities and being exposed to these practices is not sufficient for learning how to actually participate in science (Simon, Erduran, Osborne, 2006). Students likely need additional, explicit guidance on how to participate in these practices. While the teachers in the study did provide some guidance for how to analyze data and to construct and defend explanations using that data, the fact that Mrs. Lloyd provided guidance significantly more often than Mr. Gordon suggests that there are important differences, and likely some deficits, in how teachers approach helping their students to learn science practices. These differences suggest that while teachers are capable of providing the types of more explicit support students may need, some teachers, similar to Mr. Gordon, may not know how additional guidance can help students learn how to analyze and interpret data and then to use that data to construct a well-supported explanation.

Some teachers might currently guide students on how to participate in the critical data analysis and explanatory practices of science, but they may not be supporting students to truly understand science as a practice of building and refining knowledge about the world. The teachers in this study did not frequently make connections between science practices that explicitly showed students how the practices work together in service of constructing well-supported arguments to answer a question or solve a problem. Teachers' minimal time spent making connections between science practices in the classroom does not entail that they do not understand these connections themselves; not making connections could reflect a lack of intention to discuss these connections. Additionally, the teachers did not spend much time discussing why the practices are important in the field of science. Sandoval and Millwood (2005) suggest that students may need to understand the purpose behind these science practices in order to truly learn the practice. If students do not understand that the purpose of using data in an argument is to help convince an audience of their claims, they may struggle to construct well-supported arguments and fail to see how this practice is part of the practice of science as a whole. Thus, teachers' support may be inadequate for students to truly learn and understand the practice of science.

The lack of additional support in the classroom provides some insights into where teachers may struggle as they navigate implementing a focus on science practices in their classrooms. It appears that some teachers may not have readily available support strategies for helping students learn science practices. It may also be the case that teachers believe that doing inquiry-based science in the classroom will inherently teach students the practice of science, and they are not aware of the need for additional support strategies. Further, if teachers do have knowledge about strategies to teach science practices, they may be struggling to actually enact these in the classroom. Remillard (2005) emphasizes the complex relationship between what teachers want to implement in the classroom and what they actually implement in practice. From this lens, teachers implementing a new focus on science practices are in the middle of a complex relationship between their own understanding of science practices and strategies to teach these practices, as well as a wealth of additional curricular and environmental factors. Even if teachers' knowledge of science practices and teaching strategies were well-formed, there may be important factors associated with departing from traditional, teacher-centered classroom environments that might hinder the enactment of necessary support for students to learn the practice of science.

In offering implications for this study, it is important to reiterate the teaching experience and context of the two teachers. As mentioned previously, the teachers in this study were experienced science teachers who were familiar with having students justify claims, use data, and write explanations. They were not new to the idea of engaging students in the practice of science and thinking about what the practice of science looks like in the real world. Therefore, even teachers who have had several years of experience implementing innovative, design-based curricula that thoughtfully aimed to teach science practices showed need for further support in their own understanding and teaching of science practices. We should not assume that the process of teaching a curriculum designed to help students learn science practices will necessarily help *teachers learn* how to better support science practices. The idea that “learning by doing” may not be sufficient applies to teaching when we conceptualize teachers as learners. If students may need explicit guidance to learn science practices, why should not teachers also need explicit guidance to learn how to support their students’ learning of science practices?

The goal of this paper was to assess the current state of teacher support for science practices in the classroom. The findings from this study suggest that teachers may need additional support and guidance to explore and develop concrete teaching strategies that support their students’ learning of science practices. In future work, we aim to investigate what this support for teachers should be and interventions to support teachers as they help students learn the practice of science. In this future research, we will use interviews to investigate how teachers conceptualize science practices since their own understanding may influence how they help their students learn these practices. There is still much work to be done to achieve the goal of centering science education on helping students learn science as a practice, but this study provides important understanding of target areas where we can begin helping teachers to better support science practices education in the classroom.

References

- Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education, 93*(1), 26–55.
- Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K-12 science contexts. *Review of Educational Research, 80*(3), 336–371.
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*, 287.
- Erduran, S. (2014). Revisiting the nature of science in science education : Towards a holistic account of science in science teaching and learning. In *Proceedings of the Frontiers in Mathematics and Science Education Research Conference* (pp. 14–25). Famagusta, North Cyprus.
- Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.
- Lehrer, R., & Schauble, L. (2015). The development of scientific thinking. In L. S. Liben, U. Muller, & R. M. Lerner (Eds.), *Handbook of Child Psychology and Developmental Science* (Seventh., pp. 671–714). Wiley.
- McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students’ construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences, 15*(2), 153–191.
- NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: The National Academies Press.
- Puntambekar, S., Stylianou, A., & Goldstein, J. (2007). Comparing classroom enactments of an inquiry curriculum: Lessons learned from two teachers. *Journal of the Learning Sciences, 16*, 81–130.
- Reiser, B. J. (2013). What professional development strategies are needed for successful implementation of the Next Generation Science Standards?, (September). Paper prepared for K12 center at ETS invitational symposium on science assessment. Washington, DC. <http://www.k12center.org/rsc/pdf/reiser.pdf>
- Reiser, B. J., Berland, L. K., & Kenyon, L. (2012). Engaging students in the scientific practices of explanation and argumentation. *The Science Teacher, 79*(4), 34–39.
- Remillard, J. T. (2005). Examining key concepts in research on teachers’ use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246.
- Sandoval, W. A., & Millwood, K. A. (2005). The quality of students’ use of evidence in written scientific explanations. *Cognition and Instruction, 23*(1), 23–55.
- Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*(3), 345–372.
- Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education, 28*(2-3), 235–260.
- Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluation, 7*(17).
- Tabak, I., & Radinsky, J. (2015). Educators’ coaches, peers, and practices: Revisiting how teachers learn. *Journal of the Learning Sciences, 24*(3), 343–346.
- Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education, 96*(5), 878-903.

Acknowledgments

We thank both participating teachers and their students. This research has been supported by an ECR grant from the National Science Foundation to the second author (Grant # 1431904).

How a 6th Grade Classroom Develops Epistemologies for Building Scientific Knowledge

Christina Krist, Northwestern University, ckrist@u.northwestern.edu

Abstract: Current reforms in science education emphasize scientific practices as the means by which students develop and use scientific ideas. However, supporting students in engaging meaningfully in scientific practices is challenging because we do not know much about what students learn *about* the process of engaging in scientific practices, or the epistemic criteria guiding their work. In this paper, I characterize how classroom communities develop sets of epistemic heuristics by engaging in scientific practices over time. Specifically, I present how one classroom community's implicit answers to "*What kind of answer are we working to build?*" and to "*How does the idea we are trying to build relate to other phenomena and ideas?*" shifted throughout a unit. I argue that these shifts were designed into the curriculum, but required strategic work on the teacher's part; and that these shifts reflect epistemic learning.

Keywords: scientific practices, knowledge building, epistemology, communities of practice

Introduction

Current reforms in science education emphasize scientific practices as the means by which students develop and use scientific ideas (NRC, 2012). These reforms draw upon a situated perspective of learning in which students learn through participating in a community of practice to make progress towards a shared goal (e.g., Brown & Campione, 1996). In science classrooms, this shared goal is to build scientific knowledge that is useful for explaining the natural world. Scientific practices, then, are the ways that students engage to build that knowledge.

Supporting students in engaging meaningfully in scientific practices is challenging because we do not know much about what students learn *about* the process of engaging in scientific practices. Careful research has begun to characterize what student engagement in specific practices looks like over time (e.g., Schwarz, et al., 2009). However, students can still engage in the "doings" of the practices without an understanding of how and why those practices are useful for building knowledge (i.e., by rote). In addition to "doing" the practices, students should come to understand the "hows and whys" undergirding those practices: both the practical heuristics for how to build knowledge (Berland et al., 2015), and why those epistemic heuristics are useful (Manz, 2014).

In this paper, I aim to characterize how a classroom community develops sets of epistemic heuristics (i.e., epistemologies for scientific knowledge building) by engaging in scientific practices over time. Specifically, I examine how middle school students' consideration and use of two specific epistemic heuristics changes over the course of one content-area unit, and how those changes are supported by the teacher and the curricular context. By providing empirical evidence for students' knowledge-building work in practice, this study contributes to our growing understanding of how to support students' meaningful engagement in scientific practices.

Developing situated epistemologies-in-practice for scientific knowledge building

From a situated perspective, learning involves shifting how one participates in a community of practice. The community is characterized by their "joint enterprise," or their collective set of goals that gives the community a sense of what they are all about (Lave & Wenger, 1991). Applying a situated view of learning to schooling means that students should be learning disciplinary content in a context in which their tasks are guided by and shaping their understanding of the "joint enterprise" of the discipline. In science classrooms, that joint enterprise is building explanatory understandings of the natural world (Louca et al., 2004). When working to explain those natural phenomena, students are engaged in practices that help them make progress towards building explanations. Their engagement in these practices and the knowledge they develop both is guided by and shapes their understanding of their community's joint enterprise, or what it means and what it takes to build scientific ideas.

In this paper, I focus on understanding what students learn about what it means to build scientific ideas in science learning contexts that are organized to engage students meaningfully in science practices. Because what they are learning is epistemic—related to the nature of scientific knowledge and the work of building that knowledge—I turn to theories of epistemology to further focus my investigation.

Often, the epistemologies relevant for scientific work are phrased as unitary statements, such as "scientific knowledge is tentative." However, knowing these declarative statements has little effect on the nature of students' work. Instead, "practical epistemologies," or smaller pieces of epistemic knowledge that are

combinations of idea and action, guide the actual work that students do (Sandoval, 2005). These practical epistemologies consist of sets of “epistemological resources,” or “cognitive resources for understanding knowledge” (Louca et al., 2004, p.58) that are differentially activated depending on the context. For example, a student who thinks the goal of a task is to get the right answers on a worksheet may quickly scan through a book, looking for key terms. His activity is guided by notions that the ideas he is working to produce are already known and simply need to be stated. In contrast, the same student may carefully think through what is happening to make a cookie odor travel around a corner if he thinks the goal is to draw from what he knows about how matter behaves to work out an explanation. Thus, students always have an implicit answer to how they build science ideas, though their answer may not align with a disciplinary one. Learning how to build science ideas in disciplinarily authentic ways requires that students gradually, over time, engage in knowledge-building work that entails continual activation, or consideration, negotiation, and use, of the sets of epistemic criteria valued by disciplinary science.

The disciplinary answers for how to build scientific ideas take the form of *epistemic criteria*, or specific rules of thumb for how to construct and evaluate knowledge (Chinn, Buckland, & Samarapungavan, 2011). I focus on how students draw on resources for two criteria: the notion that scientific answers should provide mechanisms, and that they are working to build general models that explain multiple phenomena. Ideally, through engaging in work that entails continual activation of these criteria, students come to see how and why these disciplinary criteria are useful and productive for building scientific knowledge. This study aims to characterize this epistemic learning process, or how students’ use of various epistemic resources for these criteria changes over time.

To characterize students’ use of disciplinary criteria, I utilize Berland et al.’s (2015) Epistemologies-in-Practice framework. This framework identifies four epistemic criteria that are generative for both scientists’ and students’ knowledge-building, including the two that I focus on here: accounts should be *mechanistic* and accounts should be *generalizable* but built from *specific* phenomena and cases. The framework then broadens those criteria to the questions, or epistemic *considerations* (ECs) that those criteria—and many other non-disciplinary ones—serve as an answer to: *What kind of answer are we working to build?* and *How does the idea we are working to build relate to other scientific phenomena and ideas?* Thus, students tacitly consider and respond to these questions in making decisions throughout their knowledge-building process. Consequently, students’ implicit answers to these ECs, and therefore the epistemic resources guiding their work, are visible in classroom discourse and interaction organized around building scientific knowledge.

This study identifies and characterizes students’ epistemic work during their knowledge building activities over the course of one 6th-grade unit. My study focuses on a classroom with an expert teacher, Ms. L., who is using curriculum materials designed to engage students meaningfully in scientific practices (Krajcik, McNeill, & Reiser, 2008). I examine how Mrs. L.’s classroom community develops epistemologies for building scientific knowledge by investigating the following research questions:

1. How do the classroom community’s consideration and use of two epistemic criteria for building scientific ideas shift over the course of one 12-week unit?
2. How are these shifts supported by (a) the curriculum design and (b) the teacher? In what ways do these shifts indicate epistemic learning?

Methods

To investigate these questions, I conducted an instrumental case study to develop and empirically articulate the construct of interest—epistemic considerations in practice—and to provide a rich description of how this classroom community develops knowledge-building practices and norms. The primary data source for the study is a collection of video recordings of selected classroom lessons from one unit during the January-April 2013.

The curriculum for the unit organizes students’ work around a driving question: *how can I smell things from a distance?* This question sets the specific “joint enterprise” for the unit, the overarching question that they are working to explain. I selected lessons from this unit where the (intended) design of the curriculum provided opportunities for explicit knowledge-building work around the main scientific principles that the class was working to develop (e.g., lessons in which students were drawing and presenting models, or “jigsawing” interpretations from evidence to explain a phenomenon). In total, I selected 7 class periods, or approximately one every two weeks of the unit. The unrecorded class periods involved activities such as conducting investigations and interpreting data to answer sub-questions; reviewing readings; and taking quizzes or tests.

In order to see how the classroom community’s consideration and use of epistemic criteria changed over the course of the unit, I selected epistemically rich episodes, or moments in which the classroom discourse provided evidence of the students’ and teacher’s implicit answers to *what kind of knowledge they were working to provide* and *how the idea they were working to understand related to other scientific phenomena and ideas*. To select these episodes, I and a team of researchers content-logged the video according to activity types based loosely on the 8 scientific practices described in NGSS (e.g., “Developing and Using Models”; “Designing and

Carrying out Investigations”) as well as codes for general classroom activities such as “Free Time/Logistics.” We then “tagged” the video for any potential evidence of epistemic considerations (ECs) in student and teacher discourse. For example, when “tagging” for evidence of someone considering how the idea they were working to build related to other scientific phenomena and ideas, a coder marked any time a student or the teacher brought in another example (e.g., “It’s like Jello”) or used a generalization (e.g., “Well, there’s always dust in the air”). After content-logging and tagging, the research team used the distribution of tags as “sensitizing indicators” to select episodes for more in-depth analysis. We selected chunks of time that included a cluster of several EC tags within a single segment of activity. Episodes averaged 2:28 in length. The research team transcribed each episode by turn of talk. I then coded each turn for the two ECs of interest, Nature and Generality, as described next.

To characterize *Nature*, or the classroom community’s answer to “*What kind of answer are we working to provide?*”, I first coded each turn of talk within each episode for the elements of explanatory and other types of accounts. Elements of an explanatory account included describing the phenomenon, identifying factors (such as air particles) and unpacking factors (such as playing out the behavior of the air particles). Elements of other types of accounts included providing known-answer information, giving an illustrative example, describing details from personal experiences or a class activity, and imagining a hypothetical scenario. I then categorized each episode based on whether the majority of talk turns contained *explanatory* or *other* elements. If the majority of turns of talk contained explanatory elements, I characterized the episode as building an *explanatory account*. If the majority of turns contained other elements, I characterized the episode as building an *other type of account*.

To characterize *Generality*, or the classroom community’s answer to “*How does the idea we are trying to understand relate to other phenomena and ideas?*”, I first coded each turn of talk for whether the speaker was talking about a specific phenomenon, such as litmus paper changing colors; a general idea or generalization, such as the fact that all matter can exist in three states; or a representation, such as a model, that abstracted from the specific case somewhat. I then coded whether the discussion made connections between specific, represented, or general ideas, or if it simply focused on characterizing one type. If there were connections made, I coded the nature of the connections as either a connection between ideas that were both *known*, such as using a specific example like ice melting into water to illustrate the general principle about matter existing in three states; or a connection between ideas where one of the ideas was unknown and *built* during the course of the episode. For example, students drawing models to explain how and why balloons shrink in liquid nitrogen are *building* their explanation for that phenomenon.

In addition to coding for evidence of each of the ECs in discourse, I coded for teacher prompts for consideration in one of these epistemic areas. For example, a teacher could prompt for a particular type of account (coded as a prompt for Nature) by saying, “But why? Why can some people smell better than others?”

Findings and discussion

From the analysis of the selected knowledge-building episodes from one unit, I present how the classroom community’s answers to the two ECs of focus (Nature and Generality) shifted over the course of the unit. Taken together, these shifts demonstrate how the classroom community’s answer to the question, “How do we build ideas?” changed over the course of one unit. The shifts are: 1. Classroom talk increasingly contained elements of explanatory accounts and they increasingly constructed coherent explanatory accounts as the unit progressed; and 2. Almost all episodes contained connections between general and specific ideas. Early in the unit, those connections were between known ideas. As the unit progressed, students increasingly built ideas, primarily by making and comparing representations (models) and articulating general principles from those representations. I present each change in more detail and interpret them in the context of the curricular and teacher supports.

Design of the curriculum and distribution of epistemically-rich episodes

The unit of focus for this study was organized around the driving question, *How can I smell things from a distance?* Through cycles of observing phenomena, generating questions, developing initial models or explanations, conducting investigations, revising those models or explanations, and generating additional questions, students build explanations to three sub-questions that together answer the driving question. These questions, and the principles that students develop over the course of several lessons to answer them, are:

1. How does an odor get from the source to my nose? (Lessons 1-5). Principles: Substances are made of particles; in gases there is empty space between the particles; particles are moving constantly.
2. What makes one odor different from another? (Lessons 6-9) Principles: Every substance has unique properties; varying molecular arrangements of atoms give substances their properties, including odor.
3. How can a material change so you can smell it? (Lessons 10-16). Principles: The particle model explains states and phases of matter in terms of particle distance, movement, speed, and arrangement.

The lessons that I selected for observations were ones where students explicitly worked to build one of these principles. Though all of the activities in the unit were designed to help students work towards these ideas, there were various points at which they used an activity to begin pulling some of those pieces together. Table 1 represents the duration of various activity types present during each lesson and indicates the number of epistemically-rich episodes during each activity type (including time that was coded for multiple activity types). Note that a variety of activity types occurred throughout the observed lessons, but the activities during which most of the epistemically-rich episodes occurred varied and was not necessarily proportional to the amount of time spent on that activity.

Table 1. Distribution of activities during each class period.

Activity Types	L1		L3		L4		L6		L11		L12	
	Min	Ep										
Asking questions, making predictions	0:00		11:16	3	6:04	1	3:00	2	1:18	1	2:12	
Constructing explanations	8:06		0:00		15:07	7	14:42	5	12:01		8:40	5
Doing an experiment	7:36		0:00		4:50	2	9:00	1	3:26		0:00	
Discussing an experiment	4:36		0:00		16:08	4	18:00	3	11:33		3:01	
Discussing a phenomenon	7:06	2	11:15	4	1:55		0:00		0:53		7:00	
Drawing or presenting models	8:47		7:44	2	13:14		14:00	2	17:13		19:18	6
<i>Logistics or independent work</i>	16:10		14:29		2:48		1:00		11:40		13:14	
Total*	39:59	2	40:06	9	41:24	7	37:01	5	39:55	1	39:46	6

*Sum of activities and episodes may be more than the represented total due to double-coding of activity types.

Nature: Shifts in “What kind of answer are we working to provide?” over time

Of the 30 epistemically-rich episodes, only six were characterized as episodes where the classroom community was providing other types of accounts. These six episodes occurred during the first three observed lessons, as illustrated in Figure 1a. During these six episodes providing Other accounts, students were primarily recalling details from their personal experiences and providing known-answer information, as shown in Figure 1b.

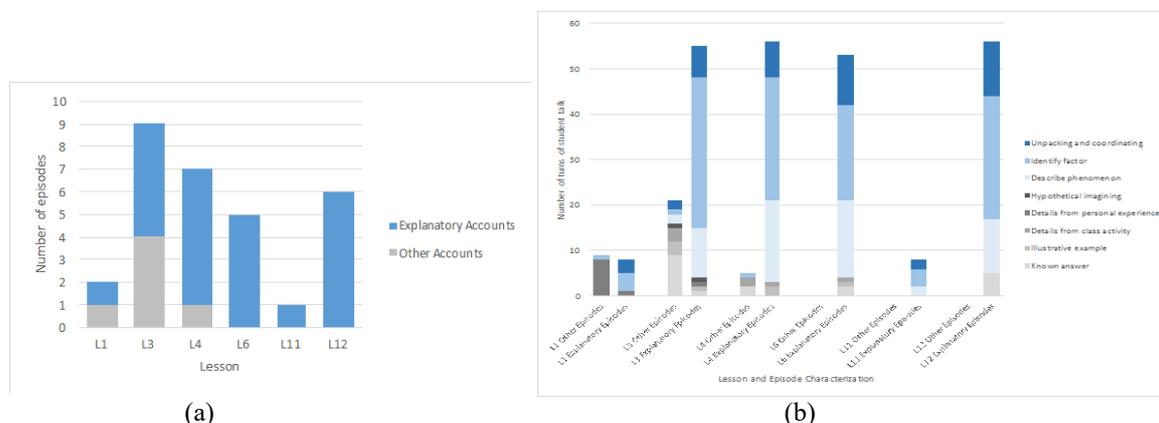


Figure 1. “Nature” characterization of episodes by lesson (a) and of student talk by episode type (b).

During the other 24 episodes, the classroom community was working to provide explanatory accounts. Student turns of talk contained all three elements of explanatory accounts, though they most often identified factors. Importantly, the turns of talk containing explanatory elements were not isolated student turns in response to a teacher question (e.g., T: What is air made of? S: Air particles). Instead, multiple turns of subsequent student talk worked to explain a phenomenon, often by weaving together both explanatory elements and elements of other account types. For example, in Lesson 3, Ms. L asked if all substances needed to be cold to “freeze.” Students responded, “No,” and then Ms. L asked them to recall the example from last night’s reading about candle wax. Stefanie responded by *describing her personal experience* with candle wax: “Once stuck my finger in some melted wax...I just waited for a bit, and then it started to dry up and turn hard.” After describing that experience, Noelle *described a similar phenomenon*: wax hardening after you blow out birthday candles. Megan and Peter then *identified two important factors* within the description of that phenomenon: “[The wax] hardens pretty quickly.” “Especially when it touches the cake.” Although not what Ms. L had hoped to identify (that temperature change

rather than absolute temperature mattered), they identified two factors relevant to room-temperature “freezing”: time, and contact with another substance (in this case, cake).

However, neither the number of turns of talk containing explanatory elements nor the types of explanatory elements they contained shifted much over the course of the unit. Students began describing phenomena, identifying factors, and unpacking those factors in Lesson 3 and they continued to do so in similar proportions through Lesson 12. Instead, what shifted were the number of elements within each individual turn of talk. In Lessons 1-6, each turn of student talk contained an average of between 1.1 and 1.2 elements. In the episode about the candle wax above, Stefanie, Noelle, Megan, and Peter’s turns would each be coded for one element. In contrast, the turns of student talk in Lessons 11 and 12 contained on average 2 and 1.5 elements, respectively. In other words, students provided more complex pieces of an account within a single turn of talk later in the unit.

To illustrate this difference, compare the individual turns in the episode about wax hardening to this exchange about squirrels finding nuts. After spending almost an entire class period tinkering with a computer model that visualized particle motion and temperature, Niral spontaneously identified *another phenomenon* for which temperature is a *relevant factor*: “Isn’t one of the reasons why squirrels look for their nuts before winter is because when the ground freezes it’s harder to smell nuts?” Katerina agreed with him, *unpacking and coordinating* the implications of that factor on particle motion: “Because when the nuts freeze, the molecules, you can’t smell them as easily [...] because when it’s warmer the molecules speed up and then it has more energy and can get to [the squirrel] faster, but in winter it can’t.” Each student provided multiple elements of the explanatory account.

Over the course of this unit, the types of accounts that the classroom community built during during epistemically-rich episodes shifted to become entirely explanatory. In addition, students gradually began incorporating multiple explanatory elements (e.g., identifying factors and unpacking them) in a single turn of talk. Taken together, these shifts suggest that the classroom community’s answer to “*What kind of answer are we working to provide?*” shifted from “definitions and facts” to “coherent mechanistic explanations of phenomena” that became more complex over time.

Generality: Shifts in “How does the idea we are trying to understand relate to other phenomena and ideas?” over time

In addition to identifying the type of account that the classroom community was working to build, I characterized how they were going about building that account with respect to the connections between general and specific ideas they were using to do so. As shown in Figure 2a, most episodes contained connections between general, represented, and/or specific ideas. In the six episodes that did not contain connections (in L3, L4, and L6), the class was characterizing a specific phenomenon while doing or discussing an experiment.

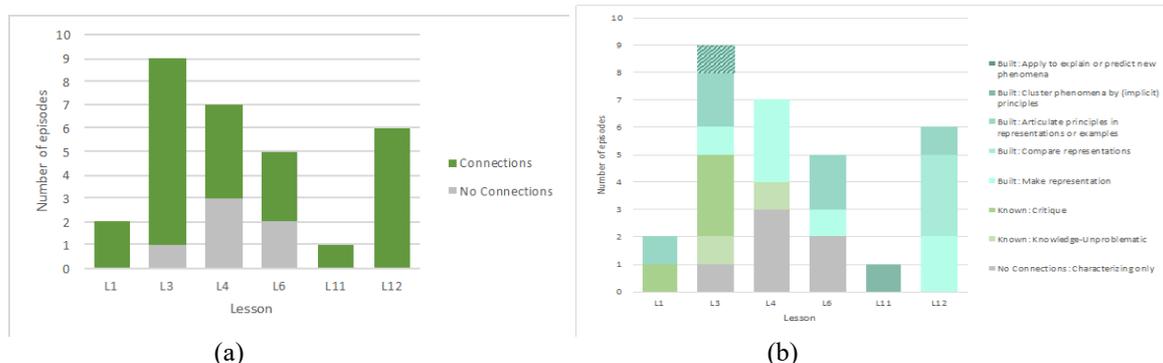


Figure 2. Number of episodes with connections (a) and connection accomplishments (b).

The nature of the connections that the classroom community made between general and specific ideas shifted over the course of the unit, as shown in Figure 2b. In about half of the episodes during the first three lessons, the classroom community made *known* connections, such as using specific examples to illustrate general principles. Interestingly, however, four of the six known-connection episodes treated the connection as a *critique* of either the example of the general principle. For example, during Lesson 3 while reviewing an idea from their reading, a student stated the principle that “All matter can exist as a liquid, a solid, and a gas.” Rather than accepting the principle as a correct and indisputable answer, though, Ms. L intentionally challenged it: “That is crazy to think about. So all matter can exist as a solid and a liquid and a gas. So I’m thinking of like, what would be hard to think about, like maybe rocks. A rock is a solid. I can’t—can I get that [points to the word “liquid”]?” In the discussion that followed, students brought in classes of examples that show that rock can melt: cement has

a liquid form, and lava is melted rock. They decided that a rock could become liquid, and possibly even gas, with enough heat.

These critiques of known-answer information often led to episodes where the class built or modified a general principle from their examples. Immediately following the rock discussion, Patrick asked about a pencil: “How can a pencil be gas?” The class talked through what would happen when you add heat to a pencil, and they decided it would probably start on fire before it would melt. They concluded by modifying their principle about the states of matter to say that maybe some things can go straight from a solid to a gas! Although burning is technically a chemical reaction and not a phase change, Ms. L’s framing of the principle as something that they could question opened the discussion for counterexamples. When a student provided one (pencils), their work through what would happen when heating it up led them to construct a sub-principle or boundary condition.

Towards the end of the unit, the majority of the episodes involved students building ideas: creating representations, comparing those representations, and articulating general principles from them. The discussion of students’ models for how a balloon shrank and grew back in different temperatures in Lesson 12 illustrates these patterns. After discussing the specific mechanisms portrayed in several individual balloon models, Ms. L asked the class, “What are we agreeing on? I mean like, what is it that we agree on?” Though there are several ways one could interpret this question, Ramona interpreted it as a call for generalizing across the models:

Ramona: Um. That, uh, when the balloon gets smaller, like, it's usually because that the molecules and particles are like, kind of compact, they're compacted together.

Ms. L: Mm-hmm.

Ramona: It kind of like, allows its space to like, like, kind of like, let loose in the balloon, and when it gets warmer, it kind of like, lets molecules kind of like split out.

Although she is using the context of the balloon, she gives several indications that she is talking about balloons shrinking in general: she uses the term *usually*, and she then speaks about what happens in present tense, indicating what happens in general rather than what happened during the specific instance they were modeling. Ramona identified the general idea that all the students (based on their models displayed on the board) seem to be in agreement that the cold balloon gets smaller because the air molecules are “compacted together,” unlike when molecules are warm and allow the empty space to “let loose” and the molecules to spread apart. This general pattern was built from their collective explorations of a specific phenomenon.

Overall, the shift to exclusively *building* connections between general and specific ideas by the end of the unit suggests that the classroom community’s answer to “*How does the idea we are trying to understand relate to other phenomena and ideas?*” shifted as well. From the beginning the classroom community demonstrated that the ideas they were trying to understand needed to make sense with other ideas and experiences. However, the nature of those connections between ideas and experiences shifted from, “We connect phenomena and principles to critique known information” to “We connect phenomena and representations to build general principles.”

How did features of the learning environment support these shifts?

In many ways, the curriculum was designed to support students in doing what they did: constructing general mechanistic accounts that were built from exploring and modeling several specific phenomena. This is encouraging, both that students are learning the content *and* that they are finding that content useful for building explanatory accounts. And undoubtedly, learning the content supported them in expressing complex mechanisms and articulating general principles from those mechanisms. Other curricular features also supported these shifts. First, the familiarity of the initial phenomenon—smelling something that was cooking before you could even see it—and the accessibility of phenomena used throughout the curriculum supported students’ identification of mechanistic elements. For example, by bringing in familiar examples such as wax hardening, students identified temperature and time as factors relevant to phase changes during Lesson 3, even though this “content” was not brought in explicitly until Lesson 10. These accessible contexts, along with Ms. L’s affirmation of students’ use of their experiences as both examples and counterexamples, made students’ everyday ideas productive resources for providing mechanistic accounts, making critical connections (such as questioning whether pencils melt), and applying principles to novel phenomena (such as squirrels smelling nuts). Importantly, students drew on everyday ideas throughout the unit, suggesting they were an integral part of their knowledge building practice.

In addition, the use of diagrammatic models supported students’ construction of mechanistic accounts, especially unpacking factors, as well as deep thinking about the specifics of a given phenomenon. By repeatedly drawing models for similar types of phenomenon (e.g., odor moving across a room; air compressed in a syringe; air in a balloon as it warms up and cools down), students’ models highlighted what was general across those

phenomena and they began using these general ideas (e.g., how molecules “usually” behave rather than how the specific air molecules in the cold balloon were behaving) to build explanations.

Finally, Ms. L worked very hard to gently problematize the kinds of ideas that she did *not* want them to be working to build. Early on, when students provided generalizations as known-answer accounts, such as, “Some people can smell better than others,” or “All matter exists as a solid, a liquid, and a gas,” she would affirm those answers, but was not satisfied with them. She would problematize a claim by calling into question the mechanism- - “Oh, interesting! So say more, How does that work?”—or by calling into question the reaches of its generality: “Excellent. That is crazy to think about. So I’m thinking like, what would be hard to think about. Like maybe rocks.” These affirmative problematizations, highlighting the “interestingness” and “craziness” of science ideas, gently led students away from known-answer accounts and stand-alone generalizations and towards constructing mechanistic accounts built across multiple phenomena.

How do we know this is learning (and not just a response to framing)?

These shifts are interesting, and appear supported by the context. However, the sets of epistemic ideas that students implicitly choose to apply when approaching a task depend on context and shift in response to teacher framing. So what counts as epistemic *learning* rather than a response to framing? One form of evidence of learning would involve seeing students respond to the same types of teacher prompts in different ways over time. That is, if a teacher is consistently prompting for mechanistic accounts throughout the unit and students eventually come to respond to those prompts with mechanistic accounts, they have learned something about what kind of answer they are working to build. Or, more convincingly, if a teacher is consistently prompting for *other* types of accounts (definitions, known-answer facts, stories, etc.) but over time students respond to those prompts with mechanistic accounts, they have learned something about the kind of answer they are trying to build.

Ms. L’s prompts for account type remained relatively consistent over the course of the unit, as shown in Figure 3. Her prompts for other types of accounts did decrease (with the exception of L11), though she was still prompting for other types of accounts more frequently than students were providing other types of accounts. In addition, about 50% of her utterances were prompts for mechanistic accounts *throughout the unit*, while the proportion of students’ utterances that provided mechanistic accounts steadily increased until they were providing mechanistic accounts in almost every utterance. This suggests that not only did students learn that they were trying to build mechanistic accounts, they found those types of accounts to be useful enough to continue providing them even when the teacher was prompting for other types of accounts.

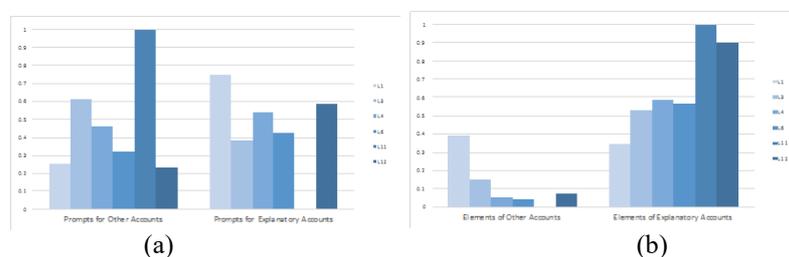


Figure 3. Proportion of teacher prompts for types of accounts (a) and proportions of student turns of talk containing evidence of account type (b) over the course of the unit.

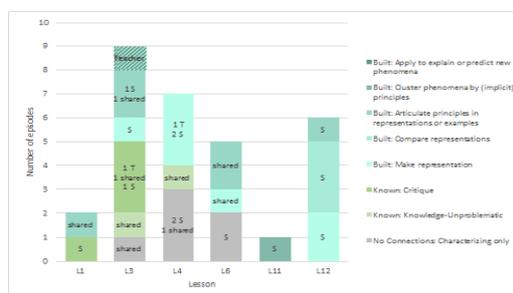


Figure 4. Distribution of generality work within each episode.

For evidence of learning about Generality, I noted who was doing the bulk of the work that led to the characterization of the episode. For example, did students begin using representations or general ideas after the teacher prompts them or does so herself? Or did they spontaneously do so, without teacher prompts or modeling?

As shown in Figure 4, the bulk of the work was shared between the teacher and the students during episodes in the early lessons. In these shared episodes, the teacher's influence tended to be towards the beginning of the episode: she would prompt for or model a connection between ideas, such as whether one can melt rock. Students then made substantive comments for the duration of the episode. In these early lessons, Ms. L was framing the type of generality work she wanted her students to do. In contrast, by Lessons 11 and 12, the students drove the episodes: they independently represented and compared ideas and spontaneously articulated principles and related phenomena. This suggests that students learned that they were trying to build general ideas from understanding specific phenomena, and that they found building ideas to be useful enough that they did it independently.

Conclusions and implications

I presented how, over the course of a 12-week unit, the classroom community's answer to "What kind of answer are we working to provide?" shifted from "definitions and facts" to "coherent mechanistic explanations of phenomena" that became more complex over time. In addition, students' answer to "How does the idea we are trying to understand relate to other phenomena and ideas?" shifted from "We connect phenomena and principles to critique known information" to "We connect phenomena and representations to build general principles." These shifts were designed into the curriculum, but required strategic work on the teacher's part; and these shifts reflect epistemic learning rather than in-the-moment responses to a specific framing.

This epistemic learning is key for students' deep understanding of what scientific knowledge *is*, what it can *do*, and how it came to be. Here, students learned to draw more consistently upon sets of epistemic resources that undergirded their engagement in knowledge building activities, which were often driven by the students themselves. Namely, this study provides empirical support for how students implicitly came to understand heuristics for how to go about building knowledge. In doing so, this study expands upon the Berland et al. (2015) framework by characterizing the range of the classroom community's answers to the Nature and Generalization considerations in greater detail; by demonstrating how, and at what grain size, a classroom community's answers shift; and by connecting the shifts in their answers to those considerations to specific features of the learning environment. This characterization of the development of a classroom community's epistemologies for science is a critical step in supporting teachers in engaging students in meaningful versions of scientific practices that engage students in authentic ways of knowing and doing science as members of a knowledge-building community.

References

- Berland, L.K., Schwarz, C., Kenyon, L., Lo, A., Krist, C., & Reiser, B.J. (2015). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*. E-view.
- Brown, A.L., & Campione, J.C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 289-325). Mahway, NJ: Erlbaum.
- Chinn, C.A., Buckland, L.A., & Samarapungavan, A.L.A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, 46(3), 141-167.
- Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, 92(1), 1-32.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge Univ. Press.
- Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, 39(1), 57-68.
- Manz, E. (2014). Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research*, 0034654314558490.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, 89(4), 634-656.
- Schwarz, C.V., Reiser, B.J., Davis, E.A., Kenyon, L., Achér, A., Fortus, D., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632-654.

Acknowledgements

Thanks to Kelsey Edwards, Dan Voss, Brian Reiser, and Henry Suarez for their helpful contributions. This research was supported by NSF ESI-1020316 and the NAEEd/Spencer Dissertation Fellowship Program. Contents are the responsibility of the author and do not represent the official views of the above organizations.

Metaphors Are Projected Constraints on Action: An Ecological Dynamics View on Learning Across the Disciplines

Dor Abrahamson, University of California, Berkeley, dor@berkeley.edu
Raúl Sánchez-García, Universidad Europea de Madrid, raulsangan@gmail.com
Cliff Smyth, Saybrook University, smythcliff@gmail.com

Abstract: Learning scientists have been considering the validity and relevance of arguments coming from philosophy and cognitive science for the embodied, enactive, embedded, and extended nature of individual learning, reasoning, and practice in sociocultural ecologies. Specifically, some design-based researchers of STEM cognition and instruction have been evaluating activities for grounding subject content knowledge in interactive sensorimotor problem solving. In so doing, we submit, the field stands greatly to avail of theoretical models and pedagogical methodologies from disciplines oriented explicitly on understanding, fostering, and remediating motor action. This conceptual paper considers potential values of ecological dynamics, a perspective originating in kinesiology, as an explanatory resource for tackling enduring Learning Sciences research problems. We support our position via an ecological-dynamics reexamination of the function of metaphor in the instruction of sports skills, somatic awareness, and mathematics. We propose a view of metaphors as productive constraints reconfiguring the dynamic system of learner, teacher, and environment.

Keywords: ecological dynamics, martial arts, metaphor, mathematics, music, somatic practice

Objectives: Bringing the body sciences into embodiment theory

The Learning Sciences are abuzz with federally funded research projects, special journal issues, symposia, workshops, and graduate-level reading courses all evaluating the putative embodied, enactive, embedded, and extended nature of individual cognition, learning, reasoning, and practice in sociocultural ecologies (Abrahamson & Eisenberg, 2012; Chandler & Tricot, 2015; Nathan, Alibali, Nemirovsky, Walkington, & Hall, 2015; Overton, 2012; Wilson & Golonka, 2013). This so-called “E-turn” in the Learning Sciences (*embodied, enactive, embedded, extended,...*) has been contributing original critiques of long-held theoretical models and pedagogical assumptions underlying much educational scholarship and practice (de Freitas & Sinclair, 2014; Roth, 2015; Spackman & Yanchar, 2013). In turn, the E-turn is stimulating the design of new technological platforms (Lee, 2015), instructional activities (Lindgren & Johnson–Glenberg, 2013), and research methodologies (Worsley & Blikstein, 2014), including the use of multimodal data gathering and analytics for investigating the impact of multimodal tasks (Abrahamson, Shayan, Bakker, & van der Schaaf, in press).

By and large, these E-turn efforts have so far drawn and oriented on classical intellectual spheres, familiar educational spaces, and canonical domains of knowledge, skill, and intervention. Namely, most educational researchers who have been inspired by the E-turn are looking to the cognitive sciences, writ large, so as better to understand the potential role that overt or covert sensorimotor activity may play in the teaching and learning of language arts and STEM content (Enyedy, Danish, Delacruz, & Kumar, 2012; Glenberg, Willford, Gibson, Goldberg, & Zhu, 2012; Landy, Brookes, & Smout, 2014). We submit that the E-turn could bear an even greater impact on the Learning Sciences if the field considered additional intellectual disciplines and cultural practices. *In particular, we argue that educational researchers who are designing and evaluating instructional activities centered on students’ motor actions should avail from engaging with empirically based robust theoretical models from disciplines that evolved explicitly so as to investigate and serve the development, practice, and remediation of human motor action.* By thus looking further afield, we argue, the field of the Learning Sciences stands to avail of theoretical frameworks that hitherto have played at best minimal roles in educational research yet are optimally poised to inform E-turn research. As a case in point, this exploratory conceptual paper will consider the potential value of ecological dynamics, a theoretical model originating in kinesiology and sports science, as an explanatory resource for tackling an enduring pan-domain LS research problem, the problem of how metaphors work in teaching and learning.

Metaphor appears to serve as an appropriate construct for investigating new horizons and resources for E-turn research. On the one hand, the analytic treatment of metaphor has for the main looked to traditional cognitive-science Cartesian epistemology (Ortony, 1993), namely the historical assumption that cognitive activity is inherent in propositional symbol processing located in a brain that is functionally isolated from its sensory input and action output. Even the cognitive-semantics theories of conceptual metaphor (e.g., Lakoff & Núñez, 2000)

and conceptual blending (Fauconnier & Turner, 2002), which look to embodied situatedness as the roots of complex reasoning, have been critiqued as unfit to explain evidence of simple human interaction, learning, and development (Gibbs, 2011, 2014). That is, whereas conceptual metaphor/blending theory offers models for ontogeny and praxis that are compelling in their descriptive and diagrammatic clarity, these descriptions are not capturing the empirical reality of mundane immersive being, such as a student dyad working together on an instructional task. On the other hand, the imagistic, situated, and often multimodal and dynamical semantic features of metaphor apparently lend this linguistic phenomenon to alternative treatments grounded in non-dualist, post-cognitivist, anti-representationist epistemologies (Chemero, 2009; Hutto, Kirchhoff, & Abrahamson, 2015). As such, a focus on metaphor may enable us to judge what the body sciences, and in particular the theory of ecological dynamics, might add to mainstream perspectives in the Learning Science.

We are thus arguing for the application of theoretical frameworks from the motor-action sciences to the analysis of educational interactions, and we have selected metaphor as a phenomenal context for building this conceptual argument. We pursue this argument by way of examining the potential of ecological dynamics to illuminate the function of metaphors in three domains of instruction: sports skills, somatic awareness, and mathematics. We contend that ecological-dynamics treatments of instructional metaphor shed new light on its mechanism, function, and effect on learning. In particular, we will propose a view of metaphor as a type of productive constraint projected into the environment. Instructors communicate metaphors verbally—often orally and gesturally, in close, real-time responsive coaching interaction—with the purpose of steering students to modify their engagement in a problem space. Making sense of verbal content is necessarily a tacit sensorimotor process of generating covert spatial-dynamical imagistic structures (Glenberg & Kaschak, 2002). These phenomenological entities are either anchored in actual environmental features or constitute new gestalts superimposed onto the environment (“projections,” Kirsh, 2013). From the ecological-dynamics view, we theorize these induced structures as productively constraining students’ process of solving a motor-action problem of practice. Given these new constraints, students must readapt to the changed environment so as to continue seeking to satisfy their assigned objective. They do so by developing new goal-oriented motor-action coordinations better suited to the modified circumstances. These coordinations are precisely the instructor’s target learning outcome for the activity. In domains of conceptual learning, such as mathematics, these new coordinations are then encapsulated via reflection, semiosis, and discourse to take form as the target content.

Theoretical framework

Instructional metaphors—limitations of current theory

When a violin teacher asks a student to play a musical phrase as though she is throwing a ball into a basket, how does that instructional intervention result in a modification in her playing (see <https://www.youtube.com/watch?v=st4-CcO4XwM>, focusing on 01:53 – 03:14; see Figure 1)?



Figure 1. Emerging artist Mártha Déak (on left) in masterclass with Maestro Maxim Vengerov (on right), working on Mozart’s Concerto No. 3 in G Major K216. Vengerov suggests Déak play the very first two notes of the violin part as though she is throwing a basketball into the basket—“tee-yum... tum!”

Common theoretical models treating the curious phenomenon of metaphor will have us believe that the student internally builds structural alignments between source and target domains—between the imagined actions of throwing a basketball and the imagined actions and audiated effects of generating a musical phrase—and then

externally performs the intact conceptual blend of these imagined actions (Fauconnier & Turner, 2002; Lakoff & Johnson, 1980; Miller & Williams, 2010). We view these explanations as problematic, because they: (a) rely on a cognitivist perspective that assumes a mind–body duality comprising internal representations in the head that are implemented externally via the body into the environment—a view that jars with philosophical arguments (Phenomenology, Merleau–Ponty, 2005/1945; Enactivism, Varela, Thompson, & Rosch, 1991) cognitive-developmental psychology theory (genetic epistemology, Piaget, 1968), and cognitive-psychology empiricism (Hauk et al., 2004); (b) ignore the dynamical, interactive, developmental, and sociocultural facets of human experiences (Becvar et al., 2010; Malafouris, 2013); and (c) are not conducive to empirical evaluation (Gibbs, 2011, 2014). What might an E-turn model of instructional metaphors look like?

Ecological dynamics

Research on embodiment has recently been discussed in the sports sciences from a Radical Enactive Cognition perspective (Hutto & Sánchez–García, 2015) building upon the research program of ecological dynamics. The *ecological-dynamics* research program (Vilar, Araújo, Davids, & Renshaw, 2012) blends *dynamic systems theory* (Thelen & Smith, 1994) and *ecological psychology* (Gibson, 1966) enabling sports scientists to explain the learning of physical activities as the self-organizing of subject–environment dynamical complex systems. The ecological-dynamics perspective conceptualizes human learning not as hermetic process in an individual’s head but rather as an entire system tending toward new task-satisfying dynamical equilibrium among the individual (the student) and the environment, where the environment may include natural and artifactual materials that may extend the student’s reach. The system may include also fellow students, who may engage in the collaborative co-enactment of the new practice, and a coach, whose role is modeled as follows.

Whereas human agents engaged in goal-oriented activity constitute a self-organizing dynamical system, still their behavior can be affected or “channeled” by introducing various types of constraints (Araújo & Davids, 2004, p. 50). It may be useful to iterate that the term ‘constraint,’ which is used colloquially to negatively value an undesirable impediment to action, here signifies a positively valued discovery or intervention that blocks out the myriad ineffective engagement options and thus steers the novice closer toward the possibility of engaging effectively. Notwithstanding, though, and regardless of the source or type of constraint, it is ultimately the novices who, within the envelope of constraints, must develop situated sensorimotor schemes appropriate to achieving the task objective under shifting circumstances. A constraint-led model frames the *non-linear pedagogical approach* to sports, based on introducing and modifying constraints in the learning environment (Davids, Button, & Bennet, 2008; Renshaw, Chow, Davids, & Hammond, 2010). Non-linear pedagogy offers coaches and athletes a viable alternative to direct instruction. Individual athletes are explicitly encouraged each to discover their own motor-action coordination solution through goal-oriented physical activity within an appropriately constrained learning environment. Empirical evaluation of constraints-led, discovery-based learning has demonstrated its effectiveness and advantages (e.g., in skiing, Vereijken & Whiting, 1990).

To the extent that constraints are instrumental to the phenomenon of learning, they become important in research on learning. For researchers attempting to model learning as the emergence of agents’ adaptive behaviors it has been useful to identify, articulate, and classify these various constraints. The leading kinesiologist Newell (1986, p. 404) identified three sources of constraints on the learning process: organism (biochemical, biomechanical, neurological, and morphological levels), environment (e.g., gravity, temperature), and task (goals and rules). Of these three sources of constraints, our interdisciplinary study focuses on the environmental, which in turn includes the subcategory of *augmented information*—a real-time instructional intervention, such as a teacher’s responsive suggestions on how the student should modify her performance (Newell, 1996, pp. 422-423). From the ecological-dynamics perspective, the teacher’s proactive, formative evaluation of students’ suboptimal performance is thus conceptualized as introducing a productive constraint.

Augmented information comes in a variety of modalities, including speech, gesture, and/or visual materials, but also direct physical contact. To the best of our knowledge, scholars of constraint-based pedagogy have not looked at augmented information communicated via metaphor or simile. We propose a theoretical contribution centered on clarifying the systemic function of augmented information delivered via metaphor or imagistic instruction. To emphasize, from an ecological-dynamics perspective we view the potential effect of instructional metaphor as productively constraining the learner’s solving of a motor-coordination problem.

Instructional use of metaphors, analogies, and similes has been documented in various sports branches (Masters, 2000; Lam, Maxwell & Masters, 2009). Yet it extends to other *overtly* embodied physical practices, such as dance (Bernard et al., 2006; Böger, 2012; Franklin, 1996; Kolter et al., 2010). For example, a variety of metaphors are used in a somatic practice called the Feldenkrais Method (Buchanan & Ulrich, 2001; Feldenkrais, 1972). Yet metaphor is prevalent also in the *covertly* embodied conceptual disciplines, such as mathematics (Abrahamson, Gutiérrez, & Baddorf, 2012; Núñez, Edwards, & Matos, 1999; Presmeg, 2006; Sfard, 1994).

Modes of inquiry: Comparing cross-domain paradigmatic cases

We three authors are bringing to bear theoretical constructs from our respective domains of investigation—mathematics, sports, and movement-based awareness (‘somatic’) practice. Acknowledging the obvious surface differences of these domains, our collaborative rationale here is to transcend these differences so as to articulate and refine a model at a semantic level that is general enough to illuminate phenomena in the three domains but not so general that it loses practical traction in any domain. Not an empirical study, this conceptual essay is more so a collective thought experiment examining the utility of the body sciences for the learning sciences.

Evidence: Ecological-dynamics analyses of instructional metaphors across three domains

We now present relevant case studies on the use of metaphors within didactical practices from our respective domains of investigation: martial arts (judo), mathematics (algebra), and somatic practice (Feldenkrais Method).

Case 1: Judo

Harai Goshi (“sweeping hip-throw”) is a widely known and often used judo coordination pattern for overcoming an aggressor. A superficial view of Harai Goshi might focus primarily on the leg action that “sweeps” the opponent off the ground. A systemic perspective, however, describes the coordination pattern as a “balance scale falling out of balance” (Sánchez-García & González, 2014). This metaphorical visualization of a Harai Goshi performance may not be available to a novice onlooker, and yet it is instrumental to successful replication of the coordination pattern. Therefore, instructors attempt to foster this visualization via coaching intervention.

Figure 2 shows a male judoka performing Harai Goshi. The judoka is leaning his torso forward so as to counter-lever the other end of the embodied beam that sweeps up the opponent toward horizontal orientation, at which point he will rotate and fling the opponent over and onto the ground. As such, a judo sensei is likely to offer a judoka who is practicing Harai Goshi the metaphorical transition information, “Add weight to tip the balance beam out of balance” (Sánchez-García & González, 2014). If he accepts this augmented information, the judoka projects it onto his systemic field of action, which includes the opponent, as a new environmental constraint that he must comply with in order to complete the task of overthrowing the opponent. Initially, this new metaphorical image may disequilibrate and unpack the judoka’s (ineffective) motor action coordination. With practice, though, he may figure out a new coordination that engages this image as an affordance for solving the problem. Reaching expertise, the judoka need no longer invoke the image, because the coordination becomes for him second nature. Unless he coaches others with the same image, he might forget it completely.



Figure 2. Instructional use of the “balance beam scale” metaphor in judo (sport)

Case 2: Mathematics

The mathematical subject matter of algebra presents many difficulties to students, including the notion of a variable as well as its symbolic notation in alphabetical characters (e.g., x , Kieran, 2007). A common curricular sequencing from arithmetic to algebra inflicts further cognitive entailments (Herscovics & Linchevski, 1996). In particular, novices to algebra are required to replace an *operational* conceptualization of the equal sign, as in “ $3 + 6 = ___$ [solve!]”, where the “=” denotes an imperative to perform arithmetic operations on the left wing of the equation and write the result on the right, with a *relational* conceptualization, as in “ $3x + 6 = 4x + 2$ ”, where neither of the two equivalent expressions on either side of the “=” is temporally antecedent to the other, the entire proposition does not bear any particular directionality, and it is not initially clear what operations should be carried out (Jones, Inglis, Gilmore, & Evans, 2013).

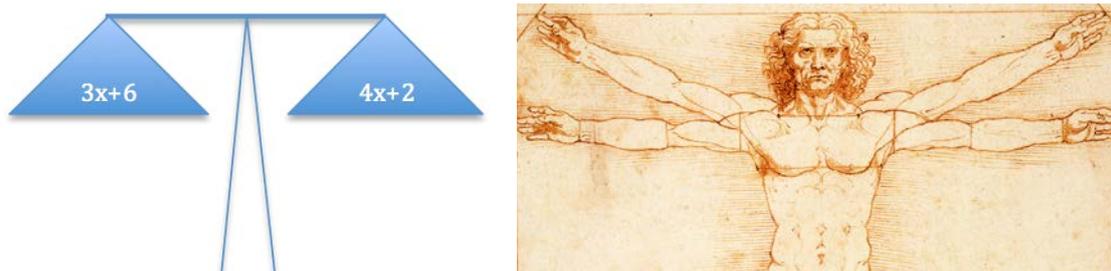


Figure 3. Instructional use of the “balance beam scale” metaphor in algebra (mathematics) juxtaposed with a fragment from da Vinci’s Vitruvian Man as illustration for the intended embodiment of balance images.

From the ecological-dynamics perspective, the phenomenon of metaphorical visualization and engagement is conceptualized as a system composed of an organism (the algebra novice), a task (“solving for x ,” i.e. the task of determining the numerical value of the variable), and an environment (the manifold ecology in which the organism is to complete the task). Algebra novices are viewed as impeded by contextually inappropriate environmental constraints that they themselves are projecting into the field of action, namely the operational visualization. That is, students’ robust arithmetic coordination patterns pre-constrain them to engage and manipulate features of the visual display in ways that block their access to a critical task-relevant affordance of the algebraic equation. The instructor who discerns behavioral evidence of the student applying these less effective constraints intervenes by offering alternative environmental constraints, and specifically by augmenting the algebraic proposition with the balance-beam structural dynamics (see Figure 3).

In passing we wish to stress from an enactivist perspective (Varela, Thompson, & Rosch, 1991) that “visualization,” a term oft used in mathematics-educational research, tacitly enfold action in addition to perception: To visualize is necessarily to see-for-acting, even if the perception never results in overt manipulation (Abrahamson, Lee, Negrete, & Gutiérrez, 2014); even when the symbols are inscribed on paper and therefore cannot be manipulated literally but only virtually (Landy, Brookes, & Smout, 2014).

Case 3: Feldenkrais Method

A common problem in human movement is how to enhance the efficient, comfortable, and safe use of the spine. This is not just a question of anatomy and physiology alone, as human movement is learned from infancy in the context of the community’s “form of life” for the individual (e.g., task demands, socio-cultural practices, the material-cultural environment; Rietveld & Kiverstein, 2014, p. 327). The most efficient movement of the spine involves the distribution of force and movement through the spine according to its overall shape, the structure of the vertebrae, and their relationships.



Figure 4. ‘Spine like a chain’ (illustration picture with superimposed image of a chain).

The overall pedagogical strategy here may be viewed via ecological-dynamics as including organism (the Feldenkrais student), environment (mat, clinic), and task (lifting the pelvis with the legs in a comfortable, efficient, and effective way). Feldenkrais pedagogy directs the learners’ attention to environmental constraints (gravity, surfaces, space), the changing relationships of/in the learner’s body, and felt qualities of action (effort, smoothness, reversibility). This “education of attention” (Gibson, 1966, p. 52) creates a ‘felt difference,’ which the student can utilize to produce more effective patterns of action (Ginsburg & Schuette-Ginsburg, 2010).

Feldenkrais teachers use instructional metaphors, whether mathematical, geometrical, spatial, mechanical, or imagistic, and these are believed to induce “global process gestalts or vitality contours,” “force-dynamic gestalts,” “causal (“if-then”) imagery,” and salient “node points” (Kimmel, Iran, & Luger, 2014, p. 17; see also Smyth, 2012). In particular, a teacher may suggest the student imagine the ‘spine as a chain’ (see Figure 4). The chain is a simple and well-known implement that transmits and shapes forces through the action of its links and articulations. It ‘maps’ well onto the structure and function of the human spine. With a student lying on their back, both knees bent and lifting their pelvis, the image of the ‘spine as a chain’ can be used to invite the student to constrain the movement of the spine (e.g., aiming to lift or lower one vertebrae, or ‘link,’ at a time). If the student accepts the suggestion to utilize the image, there is a tactile and haptic ‘sensing-for-action,’ as suggested by the enactivist perspective—a new anticipatory ‘feeling-for’ changing pressure in relation to the floor. In turn, the points of pressure of the vertebrae against the floor now become salient points of information. The movement itself may alter into a punctuated or step-like pattern, link by link, to better sense how the spine articulates in relation to the floor. Now physical effort-force (perceived muscular effort, pressure in joints and on feet) and spatial-movement dynamics (the height of the pelvis and parts of the spine) are modulated to implement this metaphorical image of the spine-as-a-chain in the learner’s action. This use of augmented information may thus afford a more differentiated and efficient use of the spine as well as an ability to sense the spine in a more differentiated way, potentially modifying future performance of this as well as related actions.

Conclusions and implications: A call to action

This explorative paper set out to consider whether theories originating in scholarship explicitly dedicated to motor development and control could contribute to educational research on teaching and learning in the disciplines (see also Beilock, 2008). Our motivation for this exercise was an assumption that motor-action models and methods should be relevant to any educational research program oriented on understanding the physical–dynamical roots of skill development, whether said skill is overtly embodied (e.g., surfing) or covertly embodied (e.g., physics). In particular, if we lay such stakes in the sensorimotor grounding of STEM concepts, it might behoove us to understand how sensorimotor activity is orchestrated with designed ‘fields of promoted action’ (Reed & Bril, 1996). Our rationale for this exercise was to: (a) select some motor-action theory; (b) identify an action-related pedagogical technique as a context for exploring the utility of the theory; (c) find appropriate paradigmatic phenomena from across the overtly and covertly embodied practices so as to test the robustness of this utility; and (d) evaluate for functional parity revealed across the practices via applying the theory. We chose the theory of ecological dynamics (Vilar et al., 2012) and its attendant constraints-led pedagogy (Davids et al., 2008; Renshaw et al., 2010), which originate in kinesiology and sports science, to examine the function of metaphor-based instructions across cases of martial-arts, mathematical, and somatic-rehabilitation practice. Our comparison suggested analogous educational process across the diverse disciplines.

This essay is no more than a Gedankenexperiment—we did not provide empirical data to support our claim. Nevertheless, we hope to have stoked a conversation on the potential of movement sciences to inform educational research. In particular, juxtaposing cases from covert and overt movement learning may sensitize educational researchers interested in STEM domains to invisible processes at play. Future empirical work may investigate the semiotic–somatic–semiotic processes putatively at play as students receive lexical information, enact it sensomotorically, and then reflect on this experience and document it using symbolic notation.

The hypothesis that instructional metaphors operate as productive constraints that learners project unto their task engagement could bear far-reaching implications for educational theory and practice. In particular, a proposal that learning across the disciplines is mobilized by projected imagery leads to an intriguing speculation: Some aspects of human reasoning in general could be modeled as the tentative projection of imagistic constraints on embodied cognitive actions. The very fundamental psychological construct of a frame could be recast from a dynamical-systems view as a constraint projected into a field of action. If learning is moving in new ways, so should the science of learning (Abrahamson & Sánchez-García, in press).

References

- Abrahamson, D. (Chair & Organizer) & M. Eisenberg (Discussant) (2012). You’re it! Body, action, and object in STEM learning. In J. v. Aalst, K. Thompson, M. J. Jacobson, & P. Reimann (Eds.), *Proc. of the 10th Int. Conf. of the Learning Sciences: Future of Learning* (Vol. 1, pp. 283-290). Sydney: ISLS.
- Abrahamson, D., Gutiérrez, J. F., & Baddorf, A. K. (2012). Try to see it my way: The discursive function of idiosyncratic mathematical metaphor. *Mathematical Thinking and Learning*, 14(1), 55-80.
- Abrahamson, D., Lee, R. G., Negrete, A. G., & Gutiérrez, J. F. (2014). Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM Math. Education*, 46(1), 79-93.

- Abrahamson, D., & Sánchez-García, R. (in press). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences*.
- Abrahamson, D., Shayan, S., Bakker, A., & Van der Schaaf, M. F. (in press). Eye-tracking Piaget: Capturing the emergence of attentional anchors in the coordination of proportional motor action. *Human Development*.
- Araújo, D., & Davids, K. (2004). Embodied cognition and emergent decision-making in dynamical movement systems. *Junctures: The Journal for Thematic Dialogue*, 2, 45-57.
- Becvar Weddle, L. A., & Hollan, J. D. (2010). Scaffolding embodied practices in professional education. *Mind, Culture & Activity*, 17(2), 119-148.
- Beilock, S. L. (2008). Beyond the playing field: Sport psychology meets embodied cognition. *International Review of Sport and Exercise Psychology*, 1(1), 19-30.
- Bernard, A., Steinmüller, W., & Stricker, U. (2006). *Ideokinesis: A creative approach to human movement and body alignment*. Berkeley, CA: North Atlantic Books.
- Böger, C. (2012). Metaphorical instruction and body memory. In S. Koch, T. Fuchs, M. Summa & C. Müller (Eds.), *Body memory, metaphor and movement* (pp. 187-199). Amsterdam: John Benjamins.
- Buchanan, P. A., & Ulrich, B. D. (2001). The Feldenkrais Method: A dynamic approach to changing motor behavior. *Research Quarterly for Exercise and Sport*, 72(4), 315-323.
- Chandler, P. & Tricot, A. (Eds.) (2015). Human movement, physical and mental health, and learning [Special issue]. *Educational Psychology Review*, 27(3).
- Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- Davids, K., Button, C., & Bennet, S. (2008). *Dynamics of skill acquisition: A constraints-led approach*. Champaign, IL: Human Kinetics.
- Enyedy, N., Danish, J. A., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *Int. Journal of Computer-Supported Collaborative Learning*, 7(3), 347-378.
- de Freitas, E., & Sinclair, N. (2014). *Mathematics and the body: Material entanglements in the classroom*. New York: Cambridge University Press.
- Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Feldenkrais, M. (1972). *Awareness through movement*. San Francisco, CA: Harper San Francisco.
- Franklin, E. (1996). *Dance imagery for technique and performance*. Champaign: Human Kinetics.
- Gibbs, R. W. (2011). Evaluating Conceptual Metaphor theory. *Discourse Processes*, 48(8), 529-562.
- Gibbs, R. W. (2014). Why do some people dislike conceptual metaphor theory? *Journal of Cognitive Semiotics*, 5(1-2), 14-36.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Ginsburg, C., & Schuette-Ginsburg, L. (2010). *The intelligence of moving bodies: A somatic view of life and its consequences*. Santa Fe, NM: AWAREing Press.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558-565.
- Glenberg, A. M., Willford, J., Gibson, B., Goldberg, A., & Zhu, X. (2012). Improving reading to improve math. *Scientific Studies of Reading*, 16(4), 316-340.
- Hanna, T. (1986). What is somatics? *SOMATICS: Magazine-Journal of the Bodily Arts and Sciences*, 5(4). Retrieved from <http://somatics.org/library/htl-wis1>
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2), 301-307.
- Hutto, D. D., Kirshhoff, M. D., & Abrahamson, D. (2015). The enactive roots of STEM: Rethinking educational design in mathematics. *Educational Psychology Review*, 27(3), 371-389.
- Herscovics, N., & Linchevski, L. (1996). Crossing the cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, 30(2), 39-65.
- Hutto, D. D., & Sánchez-García, R. (2014). Choking REctified: embodied expertise beyond Dreyfus. *Phenomenology and the Cognitive Sciences*, 14(2), 309-331.
- Jones, I., Inglis, M., Gilmore, C., & Evans, R. (2013). Teaching the substitutive conception of the equals sign. *Research in Mathematics Education*, 15(1), 34-49.
- Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 707-762). Greenwich, CT: Information Age Publishing.
- Kimmel, M., Iran, C., & Luger, M. A. (2014). Bodywork as systemic and inter-enactive competence: Participatory process management in Feldenkrais Method and Zen Shiatsu. *Frontiers in Psychology*, 5.
- Kirsh, D. (2013). Embodied cognition and the magical future of interaction design. *ACM Transactions on Human-*

- Computer Interaction*, 20(1), 3:1-30.
- Lam, W. K., Maxwell, J. P., & Masters, R. (2009). Analogy learning and the performance of motor skills under pressure. *Journal of Sport & Exercise Psychology*, 31(3), 337-357.
- Lakoff, G., & Johnson, M. L. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.
- Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from*. New York: Basic Books.
- Landy, D., Brookes, D., & Smout, R. (2014). Abstract numeric relations and the visual structure of algebra. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1404-1418.
- Lee, V. R. (Ed.) (2015). *Learning technologies and the body: Integration and implementation*. NYC: Routledge.
- Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, 42(8), 445-452.
- Masters, R. S. W. (2000). Theoretical aspects of implicit learning in sport. *Int. J. of Sport Psych.*, 31, 530-541.
- Merleau-Ponty, M. (1945/2005). *Phenomenology of perception* (C. Smith, Trans.). New York: Routledge.
- Miller, N. C., & Williams, R. F. (2010, September). *Building a better oarsman: Conceptual integration and motor learning in rowing instruction*. Paper presented at the 10th conference on Conceptual Structure, Discourse, and Language, University of California, San Diego, Sept 16-19, 2010.
- Nathan, M. J., Alibali, M. W., Nemirovsky, R., Walkington, C., & Hall, R. (2015). Embodied mathematical imagination and cognition (EMIC). In T. G. Bartell, K. N. Bieda, R. T. Putnam, K. Bradfield, & H. Dominguez (Eds.), *Proc. of the 37th annual meeting of the N.-American Chapter of the Int. Group for the Psych of Math. Ed. (PME-NA)* (Vol. 13, pp. 1352-1359). East Lansing, MI: Michigan State U.
- Newell, K. M. (1986). Constraints on the development of coordination. In M. G. Wade & H. T. A. Whiting (Eds.), *Motor development in children: Aspects of coordination and control* (pp. 341-361). Amsterdam: Martinus Nijhoff Publishers.
- Newell, K. M. (1996). Change in movement and skill: Learning, retention, and transfer. In M. L. Latash & M. T. Turvey (Eds.), *Dexterity and its development* (pp. 393-429). Mahwah, NJ: LEA.
- Núñez, R. E., Edwards, L. D., & Matos, J. F. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educational Studies in Mathematics*, 39(1), 45-65.
- Piaget, J. (1968). *Genetic epistemology* (E. Duckworth, Trans.). New York: Columbia University Press.
- Presmeg, N. C. (2006). Research on visualization in learning and teaching mathematics: Emergence from psychology. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present, and future* (pp. 205-235). Rotterdam: Sense Publishers.
- Ortony, A. (Ed.) (1993). *Metaphor and thought* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Reed, E. S., & Bril, B. (1996). The primacy of action in development. In M. L. Latash & M. T. Turvey (Eds.), *Dexterity and its development* (pp. 431-451). Mahwah, NJ: Lawrence Erlbaum Associates.
- Overton, W. F. (2012, May). *Relationism and relational developmental systems: A paradigm for the emergent, epigenetic, embodied, enacted, extended, embedded, encultured, mind*. Paper presented at "Rethinking Cognitive Development," the 42nd annual meeting of the Jean Piaget Society, Vancouver.
- Renshaw, I., Chow, J. Y., Davids, K., & Hammond, J. (2010). A constraints-led perspective to understanding skill acquisition and game play. *Physical Education & Sport Pedagogy*, 15(2), 117-137.
- Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological Psychology*, 26(4), 325-352.
- Roth, W.-M. (2015). Excess of graphical thinking. *For the Learning of Mathematics*, 35(1), 2-7.
- Sánchez-García, R., & González, A. (2014). Las cadenas hápticas como herramienta didáctica para la enseñanza de habilidades motrices (Haptic chains as didactic tool for the teaching of motor skills). *Retos*, 26, 138-142.
- Sfard, A. (1994). Reification as the birth of metaphor. *For the Learning of Mathematics*, 14(1), 44-55.
- Smyth, C. (2012). *The contribution of Feldenkrais method to mind-body medicine* (Master's thesis). (UMI No. 1536829). (Saybrook University Library: ProQuest Dissertations and Theses database.)
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind*. Cambridge, MA: M.I.T. Press.
- Vereijken, B., & Whiting, H. T. (1990). In defence of discovery learning. *Canadian J. of Sport Sci.*, 15(2), 99-106.
- Vilar, L., Araújo, D., Davids, K., & Renshaw, I. (2012). The need for 'representative task design' in evaluating efficacy of skills tests in sport. *Journal of Sports Sciences*, 30(16), 1727-1730.
- Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, 4(58), 1-13.
- Worsley, M., & Blikstein, P. (2014). Using multimodal learning analytics to study learning mechanisms. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 431-432). London, UK: Institute of Education.

Examining Tensions Among Youth, Adults, and Curriculum as Co-Designers in 4-H STEM Learning Through Design Programs

Steven M. Worker, University of California, Davis, smworker@ucdavis.edu
Cynthia Carter Ching, University of California, Davis, ccching@ucdavis.edu

Abstract: This research study explored the relationships between three types of co-designers in 4-H STEM learning through design programs: youth participants, volunteer educators, and the curriculum. Each co-designer was involved in a process of planning and making in order to accomplish a goal to satisfy requirements subject to constraints. The enactment of the learning environment revealed tensions between the co-designers evidenced in participation structures (negotiated between the curriculum and educators); abstract versus concrete approaches (between the educators and youth); and tensions amongst youth peers while engaging in small group design teams (between curriculum and youth).

Keywords: design-based science learning, making & tinkering, pedagogy, volunteer educators

Introduction

There is a recognized need for learning environments that support open-ended problem solving and encourage young people's motivation to learn science, technology, engineering, and mathematics (STEM). Making and tinkering (M&T) have gained interest for their potential to spark, sustain, and deepen young people's participation in STEM (Gutwill, Hido & Sindorf, 2015). M&T practices, related to pedagogical strategies from project-based learning (PBL) (e.g., Barron et al., 1998), and design-based science learning, help foster problem solving, creativity, and ownership; support youth in developing dispositions as designers (and makers); and promote STEM learning (Kolodner et al., 2003; Roth, 2001). M&T and design-based science learning relies on pedagogical approaches that emphasize planning, designing, and making shareable artifacts through multiple iterations of design, testing, and failure (Mehalik, Doppelt, & Schunn, 2008). Through these iterative and emergent practices, participants ultimately create shareable artifacts serving as external representations of knowledge (Punkte, van Eijck, & Jochems, 2012). Artifacts become concrete objects of talk available for critique and reflection (Roth, 1996). Research on instructional approaches integrating engineering *design* with science *inquiry* originated from work in schools (e.g., Learning-by-Design, Kolodner et al., 2003) and is related to M&T approaches often situated in out-of-school contexts (e.g., Honey & Kanter, 2013).

Research has often described idealized potential in M&T environments; for example, looking at leveling power dynamics between students and teachers (Schwartz, DiGiacomo, & Gutiérrez, 2015) or highlighting the potential of M&T to be more responsive to students' needs and thus, having potential to address inequity in STEM education (Vossoughi, Escude, Kong, & Hooper, 2013). There has been less research on pedagogy, and even less on the role of the curriculum (the sequence of facilitated activities), in shaping the learning environment for productive youth learning. One fruitful line of inquiry may be to study these spaces "in the wild," that is, in settings not organized or implemented by the researcher or by an experienced teacher. The dynamic nature of a learning environment, as well as the negotiations between actors, is messy, and greatly influences affordances for participation and learning. As such, multiple actors and artifacts shape the emergent nature of activities and sequences in the learning environment. We define "designer" as an actor in the learning environment exerting a substantial influence over the learning process: youth participants, educator, and curriculum.

M&T spaces abound in out-of-school time (OST). OST has been advanced as a space that nurtures youth interest, ownership, and learning of STEM (NRC, 2009). Many OST programs are organized by adult volunteers, offering youth voluntary participation and choice, and engaging youth over an extended period (McLaughlin, 2000). These volunteer educators often lack competence or confidence in implementing STEM education and must respond to structural constraints (e.g., voluntary participation, frequent absences, wide range of ages, organizational priorities). OST may help shed light on pedagogy employed thus clarifying tensions among major co-designers as the learning environment is co-constructed and how that, in turn, expands and/or limits opportunities for youth to learn and develop.

Theoretical points of reference

We employed sociocultural theory to make meaning of data and understand participants' values and meanings they ascribed to tools, practices, and expectations. Sociocultural theory places culture as a core concept in learning and posits that learning, development, and culture have a complex, intertwined relationship that contributes to a

child's capacity and agency to function in the world (Rogoff, 2003). Sociocultural perspectives typically employ *participation* as a metaphor to describe their intertwined nature of learning and development (Sfard, 1998). Our research, using a sociocultural approach (e.g., Lave & Wenger, 1991), focused on instances signaling deepening forms of participation and indicators of learning broader than standard-based definitions.

Methods

The purpose of the study was to describe the co-construction of learning environments by volunteer educators and youth in a naturalistic setting. The research questions addressed by this study were:

1. What are the pedagogical strategies employed by volunteer educators?
2. How do these teaching strategies influence youth (i) engagement in design practices, (ii) learning of STEM, (iii) strengthening tool competencies, (iv) manifestations of learning dispositions, and (v) development of psychological ownership?

By inquiring into these research questions, we sought to identify places of tension which revealed divergent goals and expectations between participants.

Research context

The research context was three 4-H educational robotics programs, emphasizing STEM education, organized by adult volunteer educators in Alameda, Santa Clara, and Solano Counties, California. The decentralized nature of the 4-H organization, and its reliance on volunteer educators, affords educators freedom and authority to design their own program goals, lesson plans, and make adjustments throughout the year. There are few educational standards for 4-H educational programming, although there are curricula available to aid volunteers in sequencing content and facilitating with experiential pedagogy. In this study, educators employed the *4-H Junk Drawer Robotics (JDR)* curriculum (Mahacek, Worker, & Mahacek, 2011). This written curriculum targeted middle school youth with a sequence of "messing about" science inquiry activities followed by engineering design challenges. Design activities invited youth to design, build, and test artifacts using common items (e.g., paperclips, rubber bands, craft sticks, tubing and syringes). The design challenges were framed in a non-competitive atmosphere encouraging youth to experiment and tinker with multiple solution pathways. While the domain was robotics, the curriculum has more in common with tinkering with its emphasis on learning through messing about with common materials. There were no pre-packaged kits (e.g., VEX, LEGO) and no computer programming.

Data collection and analyses

This research project relied on naturalist and qualitative multiple-case study methodology relying on participant observations and video clips, interviews with educators (Seidman, 2013), and focus groups with youth (Krueger & Casey, 2015). Data collection took place between 2014-2015 at three sites: Santa Clara, male educator and seven youth (three male, four female; between 11-16 years old) which met two-hours twice per month (six meetings); Solano, three educators (one male, two female) and eight youth (all male, between 10-12 years old) which met for an hour once per month (7 meetings); and Alameda, male educator and seven youth (3 male, 4 female; between 9-15 years old) which met for an hour once per month (six meetings). The final data corpus consisted of 17 field notes (with 139 minutes of video and 846 photographs), seven educator interviews (5 individuals; 273 minutes), three youth focus groups and two individual interviews (130 minutes). Data analyses was oriented by an inductive and comparative process beginning during field work in the form of analytical notes where we noted questions, emerging patterns, and themes (Merriam & Tisdell, 2016). Field notes were delimited using markers to segment data for deeper analysis. Analysis of focus group and interview data followed a systematic process of abstraction, delineating the transcripts using the same markers. This process of analyzing field notes separately from the individual data supported triangulation as we sought to identify concurrence and inconsistencies in educator's narrative from participant observations.

Findings

The unfolding of these 4-H learning environments was a complex progression involving interaction, negotiation, and contestation between three types of co-designers: the youth participants, the volunteer educator, and the curriculum. We identified these as designers because each was involved in a process of planning and making in order to accomplish a goal to satisfy requirements subject to constraints (the definition of design).

Designer: Youth participants. Youth-as-designers, in a traditional sense, saw them designing, building, and testing physical artifacts where they learned through an iterative design process (as intended by the curriculum). Youth were functioning in the role argued for by DiGiacomo and Gutiérrez (2014): "children should be the designers, rather than consumer, of the artifacts with which they interact" (p. 729). The learning

environment, as influenced by the other two co-designers, afforded opportunities for learning STEM content; strengthening tool competency; manifest learning dispositions: resilience, reciprocity, and playfulness (three important learning dispositions outlined by Carr & Claxton, 2002); and developing psychological ownership. Youth were also autonomous agents in the learning environment, where they decided with whom and how to interact with the materials, tools, peers, and adult educators while working in teams to design, construct, test, and redesign an artifact to fulfill pre-specified design requirements; i.e., some youth decided not to return in future meetings while others decided how they participated (e.g., “primary building”, “peripheral member”).

Designer: Adult volunteer educators. Educators were also designers, though their design goals differed from youth – they were designers of the learning environment with a goal to support youth learning. Educators worked within structural constraints, such as fluctuating attendance and the need to make activities fun, along with their own values and expectations for STEM education. The design work of the educator took place before and during the meetings. The practices of educators included selecting and adapting curriculum activities; facilitating an activity including allocating time for participation structures; framing and orienting youth to concepts and practices through varying pedagogical techniques- information transfer, autonomy support, metacognitive support, emotional coaching, and running interference. The participation structures and associated pedagogical techniques served to open and/or close avenues for youth to learn.

Designer: The curriculum. The curriculum was also a co-designer in the learning environment, with its interactions taking place both in real time by shaping interactions through its youth notebook and facilitator guide and through remote collaboration with the curriculum developer. The developer designed an intentional sequence of activities, including suggested participation structures and pedagogical techniques (learning activities, group sharing and reflection, and design & build), to generate expected learning goals for youth. The design work of the curriculum developer included responding to future ideas of activity requirements and constraints, completed long before the educator and youth came together. This idea has theoretical basis; the curriculum, a social and historical object, was a cultural tool “designed to foster collaboration and interaction in thinking among people participating in shared activity at a distance” (Rogoff, 2003, p. 274).

There were points of alignment between the design work of the three co-designers as well as areas of tension. In this paper, we share three examples that illuminate areas of tension between the co-designers: educators and curriculum; educators and youth; and curriculum and youth.

Participation structures: Tensions between the curriculum and educator

The type and sequence of activities, as well as the pedagogy, afford and/or constrain youth participation and learning. Youth were engaged in open-ended design practices, with opportunities for problem solving through failure and persistence (i.e., to manifest resiliency, playfulness, and reciprocity) are important in supporting the development of ownership, tool competency, and STEM content. Two co-designers were concerned with the type and sequence of activities to promote youth learning – the curriculum and the educator - however, at times, there were tensions between their solution pathways.

We employed participation structures as an analytical lens, in line with sociocultural theory, to describe patterns of activity, the allocation of time, and corresponding interactional norms (see Jordan & Henderson, 1995). From the data analysis, we identified six discrete participation structures: *lecture* (educator-led explanation of science or engineering ideas), *demonstration* (educator-led demonstration of a concept, tool, device, or artifact), *learning activity* (an educational activity facilitated by the adult with youth having “hands-on” time with manipulatives), *group sharing and reflection* (educator-facilitated group time where youth shared their artifact and received feedback from the adult and peers), *scripted build* (youth followed the educator’s instructions on assembling an artifact), and *design and build* (time for youth to design and build an artifact).

We calculated the amount of time spent in each participation structure over time by site (See Figure 1) using timestamped field notes. The allocation of time varied by site with Solano dedicating the most time (81%) to design and build, Santa Clara with the next most frequent design and build (46%), and Alameda with the least time in design and build (12%) but more time for scripted build (31%) and learning activities (33%). Only Alameda offered scripted builds and allocated the most time in lecture (22%). Santa Clara spent more time in group sharing and reflection (22%) than any other site. There was a stark contrast between sites in the allocation of time (e.g., Solano with much more design and build and Alameda with significantly less).

The tension was evident in the use of participation structures themselves. The *4-H Junk Drawer Robotics* curriculum only included three of the six participation structures – *learning activity*, *sharing and reflection*, and *design and build* – and did not include instructions for or recommend the *lecture*, *demonstration*, or *scripted build*. These three participation structures originated from the educators themselves. Interviews revealed that educators used the curriculum primarily as a legitimizer that this could be an appropriate 4-H program but then during the project as one of many resources to structure and sequence the learning activities.

However, all educators went “off book” for a variety of reasons, resulting in variation in how youth participated. At the same time, educators were dealing with pragmatic and structural constraints for which the curriculum did not attend. Educators were often adapting to structural constraints such as voluntary participation, by ensuring meetings were fun. The 4-H program, like many other community-based youth programs, is voluntary where youth have freedom and flexibility to participate.

Sawyer (Alameda): Because this still has to be fun. As much as I love teaching engineering and being exciting about this stuff we’re doing, if it’s not fun the kids won’t be back. I’m not there to entertain them per say. [The project] has to be moving and keep going. If we are spending a lot of time reading books or not doing anything, then the kids will get bored and they won’t come back to the next class. (Interview, 10/13/2014)

Sawyer recognized that while he wanted youth to learn engineering, he had to ensure meetings were fun for the youth, because if they were not, youth may not return. The nature and definition of “fun” was seen as offering hands-on activities. The term hands-on, however, was utilized in different ways by educators. One meaning focused on tinkering in service of reinforcing engineering learning while another meaning emphasized tinkering as a learning goal in itself. This reflects another reason for divergence between curriculum and educators, that the educator’s goals, interests, and values shaped their instructional practice with some spending more time in activities they thought valuable and in which they had an interest.

Robin (Solano): I’m a hands-on learner. ... some of today’s youth are the same way. They’re not – they have to do it to physically learn it, and that’s how I am. So I like to tinker and play with stuff. (Interview, 6/22/2015)

Robin stated that she valued hands-on design experience for its value in the service of tinkering. Further discussion with Robin revealed connections between her valuing hands-on learning and the goals she had for learning: having youth learn to be creative and developing a tinker mindset. This value was seen

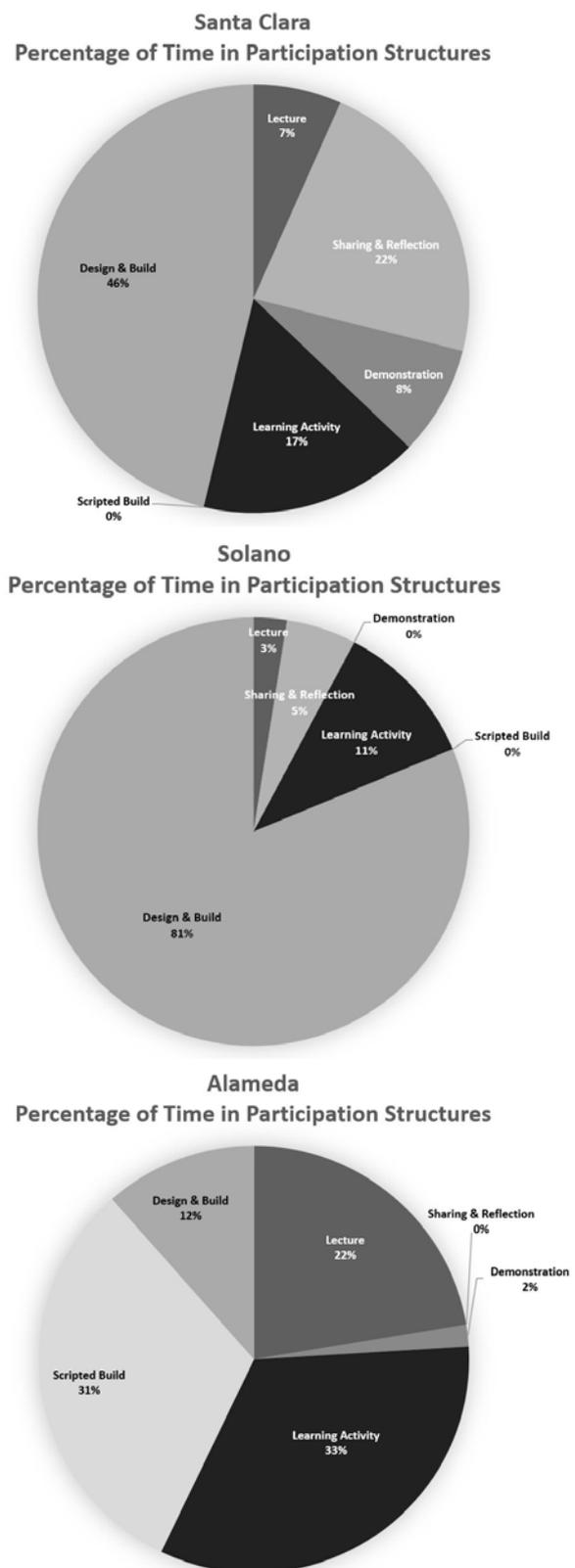


Figure 1. Time in participation structures by site.

in the allocation of time at the Solano site in *design and build* (81%), which was much more time than the other two sites.

Educators' competence and confidence in engineering were related to their approach structuring the sessions with an indication of an inverse relationship where competence and confidence in engineering were inversely related to pedagogies emphasizing metacognition and autonomy support and more information transfer approaches. In other words, educators who identified as professional engineers, who had been enculturated to professional engineering fields, were observed employing more information transfer pedagogical strategies.

Eugene (Santa Clara): I wanted to teach them something specific ... So it just naturally evolved and at the last meeting, I think the feedback from the kids wanted less talking, which is ok. We'll try to accommodate that. But I think there was a lot of learning that happened through that process. (Interview, 4/2/2014)

Further in the interview, Eugene reiterated his belief that youth needed engineering fundamentals, information, and concepts before they could successfully start building. He believed this approach produced better learning outcomes, but he recognized youth preference for more hands-on open build time.

The significance of this discussion is in how participation structures afford and/or constrain youth participation and learning. Data revealed that three participation structures – *lecture*, *demonstration*, and *scripted build* – primarily afforded STEM content learning and tool competency. The other three participation structures – *learning activity*, *sharing and reflection*, and *design & build* – afforded a broader range of learning. *Learning activities* were not about engaging in design practices or building, but rather, becoming familiar with a STEM concept though youth were observed manifesting a playful disposition, showing creative behavior in response to their peers and the educator. *Group sharing and reflection* was always in relation to something youth had built, and hence, served as evidence that youth had engaged in design practices, improved their tool competency, manifested dispositions (resilience, reciprocity, and/or playfulness), and developed a sense of psychological ownership. *Scripted build* was focused on youth following instructions to build the same artifact. Youth had opportunities to use tools - soldering irons, heat guns, and multimeters - to make something. An important distinction in tool competency between *design & build* and *scripted build* was that in the former, youth had freedom to choose when to use a particular tool, whereas in the later, there was an ordered sequence. We often observed youth using tools that we thought were not best suited for a material modification, and yet, this autonomy afforded youth opportunities to grow their tool competency, by using tools, learning how tools functioned, what their limitations were, all in the service of building a shareable artifact. During *design and build*, youth engaged in design practices where they generated ideas to meet the challenge (design) and then built and tested prototypes (build). This process afforded opportunities for youth to manifest resilience as they encountered challenges and failures, to contribute to each other (reciprocity), and be playful and experimental. Youth had opportunities to strengthen their tool competency and develop a sense of psychological ownership for their artifact.

Abstract or concrete approaches: Tensions between educator and youth

There has been longstanding friction between the developmental value of abstract/formal reasoning compared to concrete approaches (described as “bricolage” by Turkle & Papert, 1990). Those advancing making and tinkering commonly maintain the value of hands-on tinkering to promote creativity and motivation. M&T may contribute to a more supportive environment to integrate thinking and acting where participants make, explore, test, build, and redesign through cycles of iteration. However, school-based norms often find themselves weaved into out-of-school environments, and many of the educators in this study espoused abstract learning goals including the fundamentals of engineering, writing specifications (engineering notebooks), and transferring these principles to other areas in their lives. Across the study sites, I observed tension between educators' emphasis on learning activities to engage in more formal planning while youth preferred *thinkering* – a metaphor used by Roth (1996) to describe the “integral activity that had mental, material, practical, and social aspects” (p. 147).

The tension was observed through the educator's learning goals and emphasis on the planning aspects of design while youth expressed their desire for “less talking” and more “building.” For example, youth were asked to draw their designs on paper before beginning to build with the materials. This approach represented both the curriculum and educators' values of abstract thinking, namely, dividing the process into discrete components: first youth develop a representation of the artifact and then they handle materials and use tools to build the artifact. For example, at the Solano site, youth were asked to draw, individually, a gripper to be added to their group's arm. During group sharing, three youth had drawn devices using materials not available; one with “Dr. Octopus arms”, another with tongs to pick up the world, and a third with large metal claws. Two youth drew and shared designs incorporating with materials at-hand; one with a gripper using hydraulics while another with plate-like grippers. These two youth recognized constraints of materials which demonstrated awareness of design realities and

material properties. The excerpts highlight a shortcoming of abstract representation, namely, that when asked to draw a design, these youth engaged in playful and creative thinking – which does have a place in learning – not limited by the properties of materials. However, their designs were not feasible or indicative of materials constraints. Furthermore, the process of developing a representation did not attend to the types of tools or the processes youth would need in order to complete their artifact. Had youth been able to use materials to think and act, their designs would have been naturally self-limiting due to the constraints of the materials’ properties themselves. Youth engaging in tinkering with materials are naturally constrained by the materials and tools. Abstract representations can be fanciful and limitless, cultivating creativity, but tinkering with materials can be frustrating, promoting resiliency and persistence. In the context of a program emphasizing design practices, important learning concepts include constraints of materials which require analyzing trade-offs in design.

In focus groups, youth expressed a preference for tinkering and building. One youth summed his feelings for his least favorite activities: when “we did not get to build stuff” (Gordon in Alameda, 5-3-2015) In Santa Clara, youth preferred more build projects and less of the educator lecturing and demonstrating. One youth provided a response to the educator’s question around “what would make this project more fun?”

- Jason (Youth): Make your things at the beginning a little more shorter.
Eugene (Educator): Make my what things?
Jason (Youth): The presentations, since I’m pretty sure some of us were getting bored. ...
Eugene (Educator): Ok. Yep, I agree. I agree. I felt sort of like that. I like at the end you spent a lot of time building. But I think in the very beginning if I didn’t talk a lot it would be difficult for you, in my mind, to start working. Now, the last few sessions, I think we did a lot more building. And I think that was a lot of fun, I agree. Absolutely.

(Field Note Santa Clara, 3/26/2015)

Tensions were evident when educators front loaded content (like Eugene in the excerpt) or had youth draw designs before building- all around having youth work it out in their heads before they did it with their hands. Educators’ learning goals were a factor in allocation of time between participation structures and their pedagogical strategies. There was a contrast between more information transfer oriented participation structures (*lecture, demonstration*) and those affording more metacognitive and autonomy functions. The former was seen as “talking” and mostly focused on abstract concepts while the later involved youth tinkering and build time. Valuing abstract reasoning in this space served to reinforce school-based norms, something that may limit opportunities for the very thing those advocating for M&T value- creativity, open-ended problem solving, and shifting of power balances between teachers and students.

A primary builder: Tensions between youth and the curriculum

Providing opportunities for collaborative work has become something of a standard for project-based learning utilizing a peer-support social organization model. This is often rationalized as a reflection of professional practice (Puente et al., 2012), providing peer scaffolding (Rogoff, 2003), or tapping into other benefits of collaborative learning (Cohen, 1994), yet success has been found to depend on the quality of interaction (Barron, 2003). The *Junk Drawer Robotics* curriculum advanced small group learning wherein youth would work in small groups to share, design, and build together. Curriculum activities instructed educators to split youth into groups of two to four. The social organization recommended by the curriculum was implemented by educators at two of the three sites, however, the educator at Santa Clara had youth build individual projects.

Teamwork, advanced by the curriculum, and facilitated by the educator (at least at two sites), revealed another tension, evident when one youth spent more hands-on time with the artifact than their team members, making more modifications to the artifact without asking or explaining to team members, issuing instructions to team members, and acting as the spokesperson when sharing the team’s artifact. This occurred even though educators often verbalized their values: “You guys need to work as a team” (Joyce in Solano, 11/5/2014).

We labeled a young person fulfilling this informal role the “primary builder” to designate the youth who controlled the team’s artifact. The primary builder did not attempt to include other team members and was not receptive to their ideas; in essence, the primary builder controlled the artifact and excluded teammates. There were conflicts between group members, often disagreements around design decisions, especially in the earlier meetings, however, as meetings progressed, and a primary builder was established, there were fewer arguments evident, almost as though the primary builder had established themselves as the boss. The primary builder phenomenon

was not without frustration on the part of other members. This example from the last meeting of the project illustrates discontent for not being able to be more involved.

- Steven (Researcher): What was your least favorite activity?
Greg (Youth): It might be the fact that I was just sitting around just handing supplies to Jack while he built the whole thing. He basically built the whole thing. I just fastened everything ...
Steven (Researcher): Tell me more about that.
Greg (Youth): He was the mastermind of the whole thing. ... So he gets most of the credit for that. About maybe 75 percent credit for that.

(Focus Group Solano, 4/1/2015)

As quoted above, Greg's least favorite experience was not being as involved in the building of the artifact as he recognized that Jack did most of the work. This experience was echoed by other youth who did not serve as a primary builder in their respective groups. These other group members were often not included in significant design practices, yet continued to serve in legitimate, though peripheral, roles. For example, many of these peripheral members created a flag which may have been a way of signaling "this is mine, I contributed" even when the youth did not contribute to design of the primary artifact.

This is an insight into the tension evident between the curriculum's emphasis on collaborative work and this not happening in practice, even with prompting (though not scaffolding) from the educator. The hazard of a primary builder is in how it limits opportunities for other youth to engage in design practices, use tools, manifest dispositions, and develop ownership. Peripheral team members were not able to engage as fully in the design process, did not tinker as often, did not think through design decisions as often, and did not use as many tools. This certainly has consequences for the quality of learning afforded to peripheral team members. It also has disadvantages to the primary builder. Verbalizing design ideas with teammates facilitates communication, including conflicts, disagreements, and arguments, all of which may contribute to the generation of new perspectives and ideas.

Conclusions and implications

A particularly salient point illuminated by this paper is in how OST volunteer educators have to adapt to the structural constraints of their respective settings. In interviews, educators spoke frequently regarding voluntary participation in relation to their pedagogical practice, that is, having to consider that youth can "vote with their feet" (voluntary participation). Educators' recognized that while they wanted youth to learn engineering principles, they had to ensure meetings were fun, because if they were not, youth may not return. The nature and definition of *fun* was understood as offering hands-on activities; one educator was very emphatic – Sawyer: "Flat out, hands on." (Interview, Alameda, 10/13/2014). Maintaining youth engagement and interest was often seen as so crucial, that volunteer educators felt they had to make compromises in order to maintain the fun. A consequence are unplanned activities, like the Alameda *scripted build*, that preserved hands-on activities but at the expense of affording youth agency and opportunities to engage in design practices, strengthen tool competencies, exhibit resiliency, or improve feelings of ownership. These adaptations can also result in new and innovative activities, like a full-sized teeter-totter in Alameda. This activity could not have been included in a curriculum because of expense, technical knowledge to assemble, and space constraints, yet the educator's interest and passion prevailed, coupled with the flexibility provided by institutional context, and youth got to experience the concept of lever (and pivot point, fulcrum, torque) through a novel whole-body experience. In summary, in OST spaces, volunteer educators bring with them their own notions about effective teaching, their own interest and values, and through their pedagogical strategies, afford and constrain opportunities for youth to participate and learn.

M&T spaces, emphasizing design practices, may indeed hold idealized potential to "engender new types of participation structures where students and teachers can learn and become in practice, together" (DiGiacomo & Gutiérrez, 2014, p. 736). The results of this study, however, show how complex and problematic the leveling of power between teacher (and curriculum) and students may be in OST environments. Redefining relationships may be possible, but it is not necessarily a virtue of M&T in and of itself, but rather in the complex relationships and dynamic negotiations between co-designers. Studying teaching practices "in the wild," in settings facilitated by educators with varying competencies, illuminated the complex negotiations between three co-designers – youth, volunteer educators, and the curriculum – which explored questions about who can and does decide the nature of the activity structures.

This study contributes to our growing understanding of STEM learning through design in out-of-school time by illuminating tensions among major co-designers and how that, in turn, afforded and/or constrained

opportunities for youth to learn and develop. These co-actors were not without conflict, thus suggesting that these spaces and pedagogies do not exemplify STEM learning on their own, but neither do they preclude practices that deepen young people's interest and motivation for STEM learning.

References

- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12(3), 307-359.
- Barron, B.J.S., Schwartz, D.L., Vye, N.J., Moore, A., Petrosino, A., Zech, L., & Bransford, J.D. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *Journal of the Learning Sciences*, 7(3-4), 271-311.
- Carr, M., & Claxton, G. (2002). Tracking the development of learning dispositions. *Assessment in education: Principles, policy & practice*, 91(1), 9-37.
- Cohen, E.G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1-35.
- DiGiacomo, D.K., & Gutiérrez, K.D. (2014). Learning and becoming in an after school program: The relationship as a tool for equity within the practices of *making and tinkering*. In J.L. Poman et al. (Eds.), *Learning and Becoming in Practice: The International Conference of the Learning Sciences (ICLS) 2014* (Vol. 2, pp. 729-736). Boulder, CO: International Society of the Learning Sciences.
- Gutwill, J.P., Hido, N. & Sindorf, L. (2015). Research to practice: Observing learning in tinkering activities. *Curator: The Museum Journal*, 58(2), 151-168.
- Honey, M. & Kanter, D.E. (2013). *Design, make, play*. New York: Routledge.
- Jordan, B., & Henderson, A. (1995). Interaction analysis. *Journal of the Learning Sciences*, 4(1), 39-103.
- Kolodner, J.L., Camp, P.J., Crismond, D., Fase, B., Gray, J., Holbrook, J, Puntambekar, S., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design into practice. *Journal of the Learning Sciences*, 12(4), 495-547.
- Krueger, R.A. & Casey, M.A. (2015). *Focus groups: A practical guide for applied research* (5th Ed.). Thousand Oaks, CA: Sage.
- Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Mahacek, R., Worker, S., & Mahacek, A. (2011). *Junk drawer robotics curriculum*. Chevy Chase, MD: National 4-H Council.
- McLaughlin, M.W. (2000). *Community counts: How youth organizations matter for youth development*. Washington, DC: Public Education Network.
- Mehalik, M.M., Doppelt, Y., & Schunn, C.D. (2008). Middle-school science through design-based learning versus scripted inquiry. *Journal of Engineering Education*, 97(1), 71-85.
- Merriam, S.B., & Tisdell, E.J. (2016). *Qualitative research: A guide to design and implementation* (4th Ed.). San Francisco: Jossey-Bass.
- National Research Council. (2009). *Learning science in informal environments*. The National Academies Press.
- Puente, S.M.G., van Eijck, M., & Jochems, W. (2012). A sampled literature review of design-based learning approaches: A search for key characteristics. *International Journal of Technology and Design Education*, 23(3), 717-732.
- Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press.
- Roth, W-M. (1996). Art and artifact of children's designing: A situated cognition perspective. *Journal of the Learning Sciences*, 5(2), 129-166.
- Roth, W-M. (2001). Learning science through technological design. *Journal of Research in Science Teaching*, 38(7), 768-790.
- Schwartz, L.H., DiGiacomo, D., & Gutiérrez, K.D. (2015). Designing "contexts for tinkering" with undergraduates and children within the El Pueblo Magico social design experiment. *International Journal for Research on Extended Education*, 3(1), 94-113.
- Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (4th Ed.). New York: Teachers College Press.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4-13.
- Turkle, S., & Papert, S. (1990). Epistemological pluralism: Styles and voices within the computer culture. *Signs*, 16(1), 128-157.
- Vossoughi, S., Escude, M., Kong, F., & Hooper, P. (2013). *Tinkering, learning & equity in the after-school setting*. Paper presented at the FabLearn Conference, Stanford, CA.

Expanding Outcomes: Exploring Varied Forms of Teacher Learning in an Online Professional Development Experience

Maxwell Yurkofsky, Sarah Blum-Smith, and Karen Brennan,
mmy949@mail.harvard.edu, sgb839@mail.harvard.edu, karen_brennan@gse.harvard.edu
Harvard Graduate School of Education

Abstract: The field of teacher professional development has long been characterized by tensions surrounding the articulation and measurement of learning outcomes. These tensions are amplified as new technologies are used to provide online learning experiences that serve larger and more varied audiences. This paper uses a case study of the Creative Computing Online Workshop—a large-scale online learning experience for teachers—to explore the variety and breadth of teacher learning outcomes that can be considered meaningful. Through in-depth interviews with teacher participants we find that teachers make use of their learning in four distinct, but related, ways: engagement with new ideas, varied enactment in practice, rethinking of role and identity, and changing interaction with the world outside their classroom. We conclude by considering the implications of this expanded framework of teacher learning for future design and evaluation of professional development.

Keywords: teacher learning, professional development, learning outcomes, online learning, MOOCs

Introduction

Research about teacher learning has long asserted the importance of professional development but has often been frustrated in attempts to demonstrate the impact and outcomes of professional development experiences (Borko, 2004). This ongoing frustration has given rise to increasingly refined and narrow measures of the effectiveness of professional development, obfuscating the complexity and variety of potential learning outcomes—often with disappointing results (Hill, Beisiegel, & Jacob, 2013). While acknowledging the importance of research that seeks to demonstrate the value of professional development experiences in clear and quantifiable outcomes, we are concerned about what is overlooked by this approach. We therefore ask in this paper: What can we learn about the value of professional development by broadening both our definition of legitimate or substantive outcomes and the way in which we seek to uncover such outcomes in teacher learning?

To investigate this question, we present a case study of the Creative Computing Online Workshop (CCOW), an online professional learning experience focused on supporting teachers working with the Scratch programming environment in the classroom. We use in-depth, qualitative interviews to explore the outcomes from CCOW that teachers considered meaningful for their learning and teaching and compare these with outcomes emphasized by more traditional evaluations of professional development. Drawing on these findings, we ask how a broadened conception of outcomes might influence both the *evaluation* and *design* of professional learning experiences for teachers. We are particularly concerned with implications for the field of online learning, where new platforms such as Massive Open Online Courses (MOOCs) serve a large number of teachers with diverse backgrounds, needs, and interests.

We present our study of professional development outcomes from CCOW in four parts. First, we review the literature on professional development to identify both the traditional ways in which the outcomes of PD have been valued, understood, and measured and the ways in which these traditional approaches have been critiqued. We then describe the context of our research, CCOW itself, and our methodological approach to understanding teachers' own conceptions of their learning outcomes. In the findings, we highlight four types of teacher learning outcomes from CCOW. We conclude by considering these findings in light of relevant literature on teacher learning outcomes, as well as implications for further research and design of professional development.

Literature review

Over the past two decades, a number of foundational observational studies have cohered around a set of professional development characteristics associated with positive valuations by teachers (Garet, Porter, Desimone, Birman, & Yoon, 2001), increased self-reported use of particular instructional practices (Desimone et al., 2002), and enhanced teacher knowledge (Penuel, Fishman, Yamaguchi, & Gallagher, 2007). The characteristics noted by these studies include content focus, opportunities for active learning, and coherence with other learning activities (Garet et al., 2001). Underlying this consensus is a logic model, wherein professional development with

these characteristics results in changes in teachers' knowledge and beliefs, which in turn changes their practices and improves student outcomes (Desimone, 2009; Supovitz, 2001). Nevertheless, efforts to scale up professional development models that emphasize these characteristics have repeatedly been unable to demonstrate positive changes in student outcomes (e.g., Garet et al., 2011).

Broadly, there have been two categories of response by researchers to this problem. The first has been to tighten the alignment between various parts of this logic model. Wayne, Yoon, Zhu, Cronen, & Garet (2008) and Desimone (2009) each set a research agenda for the measurement of effective PD that focused on capturing the causal mechanisms between PD, teacher actions, and student learning. Following in this trend, a number of recent studies have measured the impact of professional development by evaluating teachers' adoption of practices based on a fairly narrow checklist of behaviors in specific instructional settings (e.g., Polly, 2011; Walker et al., 2011). Other similar approaches have focused on different aspects of the causal chain at different points of the research design (e.g., Hill, Beisiegel, & Jacob, 2013). While these kinds of research designs and questions represent important attempts to better understand the mechanisms through which professional development results in changes in teaching and learning, we are concerned about what may be lost when researchers focus exclusively on evaluation through these frameworks—and as a result, design professional development experiences with very specific outcome measures in mind.

Accordingly, we turn to the second category of response, which more openly embraces the complexity of both teacher and student learning processes. Some work has questioned the over-reliance on student achievement metrics in the evaluation of professional development (Cochran-Smith et al., 2012; Shulman, 1986). Others have questioned the time frame in which the impacts of a learning experience are considered, arguing that some effects of PD may only become evident after longer periods of time (Supovitz, 2001; Wayne et al., 2008), that sustainability itself is an important and valuable measure (Coburn, 2003; Zehetmeier & Krainer, 2011), or further that there are more indirect ways that teachers can internalize and enact new learning over time (Kennedy, 2005; Muijs, Day, Harris, & Lindsay, 2004).

Other work has continued this embrace of complexity by focusing on the socio-cultural and contextual nature of teacher learning and practice. Considering teacher knowledge less as abstract principles, and more as being situated in their local context (Putnam & Borko, 2000) raises questions about our ability to effectively measure what teachers across a variety of contexts have learned from a professional development experience. A multiplicity of factors influence how teachers make choices about what aspects of learning to implement in practice, in light of the particulars of their localized student needs (Ball & Cohen, 1999; Cuban, 2001; Guskey, 2002). These understandings have led to research that foregrounds teacher learning as embedded in organizations and communities, thus questioning individualized measures of learning (Johnson, 2015; Schlager & Fusco, 2003).

The emergence of MOOCs and other forms of online learning platforms as prominent sites of teacher learning has furthered the need to rethink the metrics we use to evaluate learning experiences. An ongoing tension within the field of research on MOOCs is how to define assessment (Haggard et al., 2013). Many MOOC designers and researchers challenge the application of traditional metrics to understand the learning of MOOC participants due to both the diverse and often unique intentions learners bring to these platforms and avenues for asynchronous and individualized engagement (e.g., DeBoer, Ho, Stump, & Breslow, 2014). Part of the purpose of the present study is to more fully understand the results of the interaction between the variability of learner intention and participation and the variability of what counts as meaningful learning.

Methods

This study builds on these approaches in seeking to broaden the conception of meaningful outcomes for teacher learning experiences. We use the Creative Computing Online Workshop (CCOW) as a case study to explore a central question: What outcomes do teachers identify as important to their professional learning experiences? By framing our central research question in this way, we are intentionally *not* applying a pre-determined definition of success or effectiveness, and instead using an in-depth qualitative interview approach to enable teachers to highlight which outcomes were most meaningful and valuable to them (Kvale, 1996). Rather than presenting a checklist of outcomes to be applied to future learning experiences, the findings from this case-study model are presented with the goal of illuminating possible alternative places to look in the hopes of understanding the range of ways that teacher learning can manifest (Shulman, 1983).

Research context

The Creative Computing Online Workshop (CCOW) was a six-week online course designed for teachers to learn about Scratch (<http://scratch.mit.edu>). Scratch is a free graphical programming environment that enables young people to create their own interactive games, stories, animations, and art—and then share their creations with others in an online community. CCOW was hosted from June 3 to July 12, 2013, and had international enrollment

of ~2,100 people, with 51% indicating that they intended to participate beyond “just browsing”. Both Scratch and CCOW were inspired by constructionist approaches to learning, which emphasize learning through designing, personalizing, sharing, and reflecting (Brennan, 2015). The design of CCOW reflected these principles as participants created their own projects with the Scratch programming language, shared and remixed one another’s projects, connected through the Scratch online platform and CCOW online forums, designed and pursued a final project motivated by their interests (ranging from curriculum design to hosting workshops to exploring forms of expression), and maintained online design journals that served as a record of and reflection on their participation throughout the workshop. Over the six weeks of the workshop, participants watched workshop videos 24,000 times, created 4,700 Scratch projects, and wrote 5,000 discussion posts in the course forums, Twitter, and via Google+. The teachers described in the present study represent a core group of participants whose engagement and activity in the workshop far surpasses the “average experience” indicated by such broad metrics, as is typical for MOOCs (DeBoer et al., 2014).

Data collection

The data analyzed here were collected as part of a larger study investigating what teachers found meaningful about their learning experience in CCOW (Brennan, Blum-Smith, & Yurkofsky, 2015). To obtain a sample of teacher participants who had completed the course, we identified 57 full-time teachers out of the 127 participants who had completed an exit survey and expressed a willingness to be interviewed. We intentionally sampled this group for diversity across curricular areas, ages taught, country of origin, programming and Scratch-specific experience, and extent of participation in CCOW. Of the 36 teachers that we contacted, 15 agreed to be interviewed. The majority of teachers in our final sample taught technology, lived in the United States, and had at least some programming experience. Six teachers in our sample worked in primary schools and the other nine in secondary schools. Eleven identified as “experienced” or “very experienced” with Scratch.

To investigate the question of what teachers took away from their learning experience in CCOW, we conducted semi-structured interviews with each of the 15 teachers. Interviews lasted around 90 minutes and were conducted in-person or over Skype or Google Hangouts by one of the authors. Interviews were audio recorded, transcribed by a third party service, and then checked for accuracy by one or more of the authors. We developed and followed an interview protocol structured into four major sections: (1) teachers’ demographic and professional information, (2) descriptions of meaningful learning experiences from the past, (3) discussion about which aspects of CCOW teachers found more or less meaningful, and (4) whether (and how) teachers made meaning of their learning in CCOW following course completion. The analysis and findings discussed below focus on this final part of the interview.

Our approach to data collection had three key limitations. First, our sample was biased towards US-based technology teachers who had prior experience with Scratch and programming. While this sampling certainly diminished our ability to understand the breadth and diversity of learner experience in the course, it should also be noted that there was individual and experiential variety amongst those we interviewed in terms of the context of their work and ways they made use of their learning. Second, because they were people who completed the course and elected to fill out an exit survey, it is also likely that the sample was skewed towards participants who had had positive experiences in the course. However, all the teachers with whom we spoke felt comfortable noting aspects of the course that had little or no influence on their practice. Finally, given the variety of ways, both formal and informal, through which teachers gain new knowledge, ideas, and practices, it can be difficult to precisely identify the experiences that resulted in specific pieces of new learning. We were unable to observe teachers’ practices either before or after their experience in CCOW and it is certainly possible that learning attributed to CCOW was, in fact, from another source. While the year that elapsed between CCOW participation and our interviews exacerbated this possibility, it also enabled us to see the ways that teachers reflected on their learning over time.

Data analysis

We began our analysis with line-by-line emic coding of the interviews and organized the emic codes thematically (Boyatzis, 1998). We considered relevant language in which teachers discussed their classroom practices in the time following CCOW, made reference to thinking or action that differed after the course, and reflected upon their approach to teaching and learning in the context of their course experiences. Following this initial coding, we attempted to articulate themes that would capture what teachers said they learned from their experiences in CCOW. Our review of the literature suggested that changes in teacher knowledge, skills, practices, and beliefs might be categories through which to consider the outcomes of teacher learning. As we analyzed teacher interviews, we saw these categories reflected but were also struck by unexpected outcomes such as ways that participants thought more broadly about the process of learning and their own role as a teacher. We searched

purposefully for examples of learning that did not match these initial findings, and returned repeatedly to the raw data to ensure that our emerging conceptualizations remained grounded (Maxwell, 2010). We used an iterative process of visual mapping and memoing (Luttrell, 2010) to relate these themes to one another before developing an overall framework that moved through four categories of learning outcomes: teacher and idea, teacher and practice, teacher and self, and teacher and world.

Findings

The following section will explore CCOW learning outcomes that teachers reported as meaningful. In *teacher and idea*, we look at how teachers took away particular facts, skills, and knowledge from the course. In *teacher and practice*, we consider the different ways that teachers applied those ideas to their own classrooms and students. In *teacher and self*, we explore how engagement with these ideas and practices led to changes in participants' thinking about how they approached teaching and learning in general. In *teacher and world*, we discuss how learning outcomes involved changing conceptions of learning and work beyond an individual teacher or classroom.

Teacher and idea

When reflecting on the impact of CCOW, many teachers discussed specific knowledge and skills that they gained from the learning experience. Although CCOW was designed as a creative and constructionist learning experience, one particular purpose of the workshop was to support teachers in building skills and knowledge around using and teaching Scratch in the classroom. Indeed, most of the teachers we spoke to expressed that they had participated in the course in order to gain this knowledge, particularly because of the recent release of Scratch 2.0. Justin, a middle-school digital art and design teacher, was interested in learning more skills about how to teach Scratch, and saw the workshop as a resource in his self-directed skill development. He reflected, "I think teaching-wise, I got more comfortable with different uses of screencasts and video lessons . . . it was like an orientation to Scratch 2.0, where it kind of like forced me to do some of the stuff I might have skipped from time to time." As Justin's description indicates, some of the ideas teachers gained from the course included particular programming tools and activities and ways to use them.

Others noted how engaging with activities in CCOW helped them learn new skills and ways of thinking about programming. Justine, an elementary instructional technology teacher, described how the Debug Its—programming tasks where participants had to fix a program that wasn't working—helped her to "read scripts, think iteratively, how to just decompose or pull things apart." When referring to the ideas that they learned from CCOW, there was considerable variety in how teachers articulated the value of that knowledge. In addition to simply increasing their knowledge base, some teachers focused on how building their own facility with Scratch enabled them to support more complex and in-depth student work with the program.

Teacher and practice

As teachers integrated ideas from the course into their existing practice, the nature of that integration varied, from direct reuse of course activities to changes in underlying approach to teaching. Some course activities—and the ideas embedded in their design—resonated with teachers and, in particular for those who taught Scratch in the classroom, were easily applicable to their own contexts. Aamir, a middle school computer science teacher, described how using "About Me" projects (an activity in CCOW where teachers created a program that introduced themselves to other course participants) allowed students to learn from models while developing a personal vision. Other teachers discussed using Debug Its and remixing activities, as well as showing students the "nuts and bolts" videos that the course designers had made to introduce core concepts and skills.

Beyond directly importing course activities into their classrooms, some teachers adopted less content-specific practices. Prompted by her own experience with keeping a design notebook—a place to record what she was doing during the course and how it was making her think—Sharon, an elementary computer lab manager, began to use more reflection-based activities with students. She designed prompts similar to those that guided her reflections on CCOW in order to help students develop metacognitive capacities and enable her to gain greater insight into their needs. Sharon, like other teachers, related these new practices to shifts in pedagogy, commenting that "I think I did a little too much lecturing . . . it worked better just to let them in there, and play around . . . it was finding that balance, and I think after CCOW, I found a better balance of how much information you give them to start with."

Some of the most interesting ways of turning ideas from the course into changes in practice came from teachers who taught *different* content and were less able to directly apply lessons or activities. For example, Laura, a sixth grade math and science teacher, found the experience with Scratch helpful in resolving an instructional dilemma. In describing the application to math, she reflected, "I could say—oh, use the Wait Block, but I don't. I

say—check and see what could help you . . . There’s just so many connections I think, with the habits of mind, the way that you work through something.”

It is important to stress that many of the teachers we spoke to were unsatisfied with the extent to which they had incorporated lessons from CCOW into their own practice. For some teachers, this was a result of their school’s stricter control over curriculum and pedagogy that discouraged classroom-level experimentation with new approaches. Other teachers expressed that they simply did not yet have the time to invest in substantive changes to their classroom practice, but were still hoping to use upcoming breaks to implement such changes.

Teacher and self

For some teachers, these changes in practice were connected to more fundamental changes in how they thought about what it meant to teach and be a teacher. Raphael, a high school math and physics teacher, already had a clearly articulated constructivist pedagogy, one in which he purposefully structured classroom activities to invite student inquiry and collaboration while minimizing his own role in transmitting knowledge. Yet, he described how meaningful the experience of reflection had been for him during CCOW, and what it meant to him to practice reflection in his own classroom. Noting that reflection had been “completely absent before,” he described his new approach: “I try to walk around the classroom, and I take a lot of notes about the way they’re working, and the way they’re interacting, the way they’re discussing, or not discussing to each other.” For Raphael, reflection had not explicitly been a part of his approach to teaching, but he saw it as valuable toward his goals of deeply engaging with his students’ understanding of content and supporting their exploration.

Many of the teachers that we spoke to highlighted how CCOW supported a shift away from the teacher as locus of authority and knowledge, towards students learning through exploration and collaboration. Laura, who discussed learning how to let students solve problems differently in math, described how she began to see her role as a teacher in science instruction change in the year following CCOW. “From the very beginning this year I just stood back. . . . Whereas last year . . . I would intervene a lot more. . . . Sometimes they need to learn by failure and I didn’t want to let that happen. . . . And it’s funny because this year’s kids are a little more rambunctious than last year’s kids.” Laura mentions changes in her practice, but also relates these changes to a new conception of herself as a teacher, as someone who is not a “hoverer” but feels comfortable standing back and letting students make mistakes. Some teachers emphasized that greater confidence in their programming knowledge and skills allowed them to be more comfortable with letting go of some control and allowing for greater student exploration, despite the increased potential for messiness and uncertainty.

Teacher and world

Some teachers articulated the value of their experiences in CCOW in terms of new or different ways of engaging with the world beyond their individual classrooms. While this theme was less widespread amongst the teachers we spoke with, its power for some speaks to the importance of broadening our conceptions of meaningful professional development outcomes.

CCOW served a legitimating function for some, giving them confidence and knowledge to advocate for creative computing with school administrators. Diane, an elementary director of educational technology, reflected that “the final project helped to clarify in my mind and to be able to validate it and to better be able to present it to my administration,” adding that she was able to use this clearer articulation of the kind of work she wanted to do to leverage the purchase of new computers. This kind of support was especially important given that many of the technology teachers we spoke with were often isolated from general school improvement work.

A number of teachers reported that their experience in CCOW helped them to make better use of the Internet to find ongoing professional support and information relevant to their teaching. As Kelly, a high-school computer studies teacher described, “I think I really learned about the power of the Internet for educators, of how connected we can be. . . . I usually go when I have a need, or I’m looking for something, or I’m going to show it to somebody else.” In this sense, ongoing access to information and peers working through similar endeavors helped teachers in actually implementing some of the instructional practices highlighted in CCOW.

For others, tapping into an online community of educators with similar ambitions and values had additional benefits. Aamir, who was the only technology teacher at a more rural school outside the United States, reflected on the impact of the online community that developed as a result of CCOW.

[The online course and community] also gives you the feeling that even if you are in a small village in the countryside, and any place on earth, you can be part of this huge community. It is very important. It is very important. It’s like you are more connected, you are part. You do exist. Very important. You do exist. You can speak with a lot of teachers in the countryside, I know

them and they feel like they really don't exist. With online courses, even if you are in the countryside, or very isolated, you can be part of a community.

As reflected in many of these experiences, it can be meaningful and powerful for teachers to locate their individual work, thinking, or approach within broader communities. The teachers that we spoke with came into the course with a range of experiences with Scratch, the world of computer science, creative computing, and constructivist approaches to teaching and learning. Some, like Raphael and Aamir, had a strongly developed pedagogical approach that was already aligned with CCOW and could locate that approach within conversations of theory and practice beyond their schools. For others, their CCOW experience helped them to see that the work they were doing, or wanted to be doing, was connected to larger communities. Diane, who noted the value of CCOW in advocating with her administration, reflected on how the course experience opened her up to a world of practice that she hadn't known about, but that deeply resonated with her vision of teaching.

Conclusions and implications

Our study builds on prior research that seeks to understand complex processes of teacher change (Ball & Cohen, 1999; Kennedy, 2005) and how best to design and evaluate professional development in light of these processes (Putnam & Borko, 2000; Muijs et al., 2004). Through in-depth interviews with teachers, we found that teachers were influenced by the ideas and practices they encountered in CCOW in a variety of interconnected ways—adapting their classroom practices, adjusting their roles as teachers, and/or securing organizational and informational resources to support their continued improvement.

This methodological lens allowed us to see a range of important shifts in teachers' behaviors and attitudes that might not have been captured by traditional metrics for evaluating professional development. Many of the ways in which teachers were influenced by specific practices in the course fall under what Spillane & Jennings (1997) call below-the-surface changes in instructional practice, which are more difficult to observe or evaluate but potentially more impactful (Coburn, 2003). The interconnections between the affective and cognitive dimensions of teacher learning in the course and the ways in which CCOW provided some teachers with the validation and confidence necessary to pursue changes in practice highlight the importance of the cultural surround in teacher development (e.g., Kirkpatrick & Johnson, 2014). Further, Cohen's (2011) discussion of the challenges of *ambitious teaching*, in which uncertainty and difficulty increase in the quest for deeper understanding, speaks to the messier terrain that teachers described themselves as being better able to traverse with heightened skills, validation, and confidence—benefits derived from both their learning experiences in CCOW, as well as their ongoing online interactions with peers.

The varied impact of professional development on teachers has important implications for the evaluation of professional learning experiences. Even when combined, traditional approaches to measuring the effects of professional development would have missed much of what teachers found meaningful about CCOW. Teachers did change their practices, but often in subtle or unexpected ways that the observational rubrics used in assessing many professional development programs might miss. While measured knowledge and skill likely increased for some teachers, for many teachers discrete knowledge gain had little to do with their changed approach to teaching. Additionally, as others (e.g., Johnson, 2015) have noted, traditional metrics rarely capture teachers' ability to share the practices they learned with others, advocate for resources that may support student learning, or learn more from colleagues in and outside of their school. Perhaps most importantly, because CCOW specifically values student outcomes that are difficult to capture on standardized tests (e.g., creativity, ability to collaborate with others, debugging), assessments of student outcomes might not capture what students gained from their teachers' engagement with CCOW. It is therefore worth considering what kinds of professional development experiences we may be labeling as "ineffective" because the outcomes they supported did not align with the metrics used to evaluate them.

This study also has implications for the design of online professional development experiences, and whether the traditional criteria used for evaluating professional development may constrain such design. Indeed, were CCOW designed to cultivate outcomes measurable by traditional evaluations of professional development, many aspects of the learning experience that influenced teachers' practices in important but unanticipated ways may have been cut or underemphasized. This raises the broader question of how to design for diverse outcomes, both those which have and have not been consciously articulated by course designers. One way in which the designers of CCOW left room for course participants to take the learning of the course in individually or locally meaningful directions was through the final independent project, where teachers took the concepts and skills about Scratch they had learned in the course and applied them in their own way. CCOW was also purposefully designed to build connections that teachers could access beyond the active time period of the course. As described by Kelly, learning how to make use of such connections can enable future learning by gaining greater knowledge and

confidence with utilizing the Internet as a tool for professional growth and development. Because the substance of such future learning cannot be entirely predicted by those designing a learning experience, designers must negotiate tensions that arise between providing teachers direct alignment of learning experience and current issues of practice and allowing for the possibility of new and unexpected ways of using that learning. In particular, while MOOCs can seem to leave room for varied learner intentions and styles through the privileging of learner autonomy, they leave open the question of how designers can anticipate the fact that participants may not always know what they want or need.

As a case study, the present work has only begun to explore the ways in which teachers might make meaning of, and take value from, their learning in a professional development experience. As CCOW was a constructivist learning experience around a content area (Scratch) with its own very particular embedded assumptions about learning, we wonder how the range of teacher learning outcomes would be different in a more traditional setting. Additionally, emerging technologies for online learning have the potential to complicate the picture of what teachers learn from professional development and how that learning can be measured. As such, there is an ongoing need for research that continues to investigate the relationship between the structure and content of the learning experience and the varied forms of learning that may result.

References

- Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond and G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3-31). San Francisco, CA: Jossey-Bass.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3-15.
- Boyatzis, R. E. (1988). *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA: Sage Publications.
- Brennan, K. (2015). Beyond technocentrism: Supporting constructionism in the classroom. *Constructivist Foundations*, 10(3), 289-296.
- Brennan, K., Blum-Smith, S., & Yurkofsky, M. (2015, April). *Supporting constructionism in the classroom: MOOC as platform for teacher learning*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. Retrieved from the AERA Online Paper Repository.
- Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3-12.
- Cochran-Smith, M., Cannady, M., Mceachern, K., Mitchel, K., Piazza, P., Power, C., & Ryan, A. (2012). Teachers' education and outcomes: Mapping the research terrain. *Teachers College Record*, 114(10), 1-49.
- Cohen, D. K. (2011). *Teaching and its predicaments*. Cambridge, MA: Harvard University Press.
- Cuban, L. (2001). *Oversold and Underused: Computers in the Classroom*. Cambridge, MA: Harvard University Press.
- DeBoer, J., Ho, A. D., Stump, G. S., & Breslow, L. (2014). Changing "course": Reconceptualizing educational variables for Massive Open Online Courses. *Educational Researcher*, 43(2), 74-84.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, 24(2), 81-112.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945.
- Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., ... others. (2011). Middle school mathematics professional development impact study: Findings after the second year of implementation. NCEE 2011-4024. *National Center for Education Evaluation and Regional Assistance*. Retrieved from <http://eric.ed.gov/?id=ED519922>
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice*, 8(3), 381-391.
- Haggard, S., Brown, S., Mills, R., Tait, A., Warburton, S., Lawton, W., & Angulo, T. (2013). The maturing of the MOOC: literature review of massive open online courses and other forms of online distance learning. *Department for Business, Innovation and Skills, UK Government*.
- Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research consensus, crossroads, and

- challenges. *Educational Researcher*, 42(9), 476-487.
- Johnson, S. M. (2015). Will VAMS Reinforce the Walls of the Egg-Crate School? *Educational Researcher*, 44(2), 117-126.
- Kennedy, A. (2005). Models of continuing professional development: a framework for analysis. *Journal of In-Service Education*, 31(2), 235-250.
- Kirkpatrick, C. L., & Johnson, S. M. (2014). Ensuring the ongoing engagement of second-stage teachers. *Journal of Educational Change*, 15(3), 231-252.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage Publications.
- Luttrell, W. (2010). Interactive and reflexive models of qualitative research design. In W. Luttrell (Ed.), *Qualitative educational research: Readings in reflexive methodology and transformative practice* (pp. 159-163). New York, NY: Routledge.
- Maxwell, J. A. (2010). Validity: How might you be wrong? In W. Luttrell (Ed.), *Qualitative educational research: Readings in reflexive methodology and transformative practice* (pp. 279-287). New York, NY: Routledge.
- Muijs, D., Day, C., Harris, A., & Lindsay, G. (2004). Evaluating CPD: an overview. Retrieved from <http://wrap.warwick.ac.uk/55785>
- Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, 44(4), 921-958.
- Polly, D. (2011). Examining teachers' enactment of technological pedagogical and content knowledge (TPACK) in their mathematics teaching after technology integration professional development. *Journal of Computers in Mathematics and Science Teaching*, 30(1), 37-59.
- Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, 29(1), 4-15.
- Schlager, M. S., & Fusco, J. (2003). Teacher professional development, technology, and communities of practice: Are we putting the cart before the horse? *The Information Society*, 19(3), 203-220.
- Shulman, L. S. (1983). Autonomy and obligation: The remote control of teaching. In L. S. Shulman & G. Sykes (Eds.), *Handbook of teaching and policy* (pp. 484-504). New York, NY: Longman.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Spillane, J. P., & Jennings, N. E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record*, 98(3), 449-481.
- Supovitz, J. A. (2001). Translating Teaching Practice into Improved Student Achievement. In S. Fuhrman & National Society for the Study of Education. (Eds.), *From the capitol to the classroom: standards-based reform in the states* (Vol. 100th, pt. 2, pp. 81-98). Chicago: National Society for the Study of Education: Distributed by the University of Chicago Press.
- Walker, A., Recker, M., Robertshaw, M. B., Olsen, J., Leary, H., Ye, L., & Sellers, L. (2011). Integrating Technology and Problem-based Learning: A Mixed Methods Study of Two Teacher Professional Development Designs. *Interdisciplinary Journal of Problem-Based Learning*, 5(2).
- Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting With Teacher Professional Development: Motives and Methods. *Educational Researcher*, 37(8), 469-479.
- Zehetmeier, S., & Krainer, K. (2011). Ways of promoting the sustainability of mathematics teachers' professional development. *ZDM*, 43(6-7), 875-887.

Acknowledgments

We would like to express our sincere appreciation to all of the Creative Computing Online Workshop participants, particularly those who were so generous with their time to respond to surveys and participate in conversations with members of the research team. This material is based upon work supported by Google and by the National Science Foundation under DRL-1019396. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Google or the National Science Foundation.

Investigating Effects of Embedding Collaboration in an Intelligent Tutoring System for Elementary School Students

Jennifer K. Olsen, Human-Computer Interaction Institute, Carnegie Mellon University, jkolsen@cs.cmu.edu
Nikol Rummel, Institute of Educational Research, Ruhr-Universität Bochum, nikol.rummel@rub.de
Vincent Alevan, Human-Computer Interaction Institute, Carnegie Mellon University, aleven@cs.cmu.edu

Abstract: Intelligent Tutoring Systems (ITSs) are beneficial for individual students learning in several domains, including mathematics where they have been used to support both secondary and elementary students. Collaborative learning may be beneficial to include in ITSs, particularly for conceptual knowledge. There is little work on collaborative ITSs, and it has mostly focused on older students. We aim to extend this work to elementary school students, by extending an ITS for fractions so it supports collaborative learning. We also build upon our previous work to further investigate the complementary strengths of collaborative and individual learning. In our study, 189 elementary school students worked with a conceptual or a procedural fractions ITS, and either individually or collaboratively. Students in both ITS conditions learned, but there were no differences in learning between individual and collaboration. However, the students working collaboratively spent less time on the tutor, indicating potential benefits of collaborative learning on efficiency in this setting.

Keywords: collaborative learning, intelligent tutoring systems, primary school

Introduction

Intelligent Tutoring Systems (ITSs) have been very successful in supporting students' learning as they work individually to solve problems (VanLehn, 2011). Although this individual work has been shown to be beneficial, there may be additional benefits from students being able to work collaboratively, especially when they are acquiring conceptual knowledge (Mullins, Rummel, & Spada, 2011). Previous research has shown that integrating collaboration into an ITS environment can be conducive for learning (Walker, Rummel, & Koedinger, 2009). But most of this work has been done with older students and has not been extended to elementary school students where collaboration is expected to be more challenging. We aim to investigate if the benefits seen with older students can apply to elementary school students and how best to utilize collaboration and individual learning within an ITS to support students.

Although ITSs have been shown to be beneficial to students in many domains (Murray, 2003), they have been shown to be particularly successful with mathematics (Ritter, Anderson, Koedinger, & Corbett, 2007). For elementary school students, fractions are challenging (Moss, 2005), but the learning of fractions can be successfully supported through the use of an ITS, showing that ITSs can be beneficial for young learners (Rau, Alevan, & Rummel, 2012). ITSs are beneficial to students by providing them with cognitive support as they solve a problem (VanLehn, 2011). ITSs provide step-by-step guidance for students both through the use of immediate feedback on steps and through on demand hints. That is, students will know right away when an error occurs and they can decide to request help from the system to figure out how to do any problem-solving step correctly. ITSs also track individual students' knowledge growth, which enables them to implement a cognitive mastery approach with individualized problem selection. ITSs often focus on helping students acquire problem-solving skills, and by integrating collaboration, we may be able to support students in explaining their reasoning, as is specified in the United States common core standards (Common Core State Standard Initiatives, 2015).

Collaboration is often beneficial for learning (Slavin, 1989; Lou, Abrami, & d'Apollonia, 2001) and is a way to get students to express their reasoning (Hausmann, Chi, & Roy, 2004). Collaborative problem solving may help students develop a deeper conceptual understanding (Teasley, 1995) through mechanisms such as co-construction and explanation-giving (Hausmann, Chi, & Roy, 2004; Chi & Wylie, 2014). Through these mechanisms, collaborative learning may be beneficial as an addition to standard ITS technology, especially for conceptual knowledge where students often need to develop a deeper understanding of concepts in the domain (Mullins et al., 2011).

Although most prior research on ITSs has focused on students working individually, there has been some work combining ITSs and collaboration with high school students (Baghaei, Mitrovic, & Irwin, 2007; Walker et al., 2009; Diziol, Walker, Rummel, & Koedinger, 2010). Walker, Rummel, and Koedinger (2009) found that students working with a tutor that had been redesigned to support peer tutoring (i.e., the tutoring system provided

support to the student in the role of the peer tutor) achieved learning gains at least equivalent to those working individually, demonstrating that collaboration can be successfully combined with ITSs. However, much of this work has been done with secondary and college students rather than elementary school students. For elementary school students, learning collaboratively may be challenging, especially in STEM domains such as mathematics (Mercer & Sams, 2006). Elementary school students often do not have fully developed social skills, making collaborative activities more challenging. Also, elementary school students may not have developed the vocabulary to discuss complex math concepts and relations. Despite these challenges, collaboration may still be effective for elementary school students by allowing them to make their thinking explicit and to practice their ability to talk about mathematics (Chi & Wylie, 2014). Even outside of ITSs, few studies have investigated whether computer supported collaborative learning (CSCL) can have a positive impact on learning with elementary school students. The studies that have been conducted in this area have either compared the use of a CSCL setting to face-to-face collaborative learning (i.e., not supported by computers) or have focused on technology interventions that mix individual and collaborative learning tasks without comparing learning collaboratively to learning individually as we propose to do (Chen & Looi, 2013; Lazakidou & Retalis, 2010; Tsuei, 2011). This research has shown positive impacts of young children working in small groups and with computers that can be extended to the use of ITSs.

In addition, collaborative and individual learning may have different strengths. For example, individual and collaborative learning activities may be better for acquiring different types of knowledge, such as conceptual and procedural knowledge (Mullins et al., 2011). Conceptual knowledge is the implicit and explicit understanding of the principles in a domain and how they are interrelated (Rittle-Johnson, Siegler, & Alibali, 2001) and may be better acquired through collaboration (Teasley, 1995). Procedural knowledge is the ability to be able to perform the steps and actions in sequence to solve a problem (Rittle-Johnson, Siegler, & Alibali, 2001) and may be better acquired through individual work with the opportunity of more practice. Understanding the relative strengths of individual and collaborative learning, which we focus on in this paper, may ultimately help in designing instructional conditions that effectively combine individual and collaborative learning.

In a previous study with elementary school students (Olsen, Belenky, Alevan, & Rummel, 2014a), we did not find differences in learning gains between students working collaboratively or individually for either conceptual or procedural knowledge. However there were several weaknesses in the study that we are addressing with the current study. In the previous study, the procedural tasks may have been too difficult for the students. Also, the total instruction time was only 45 minutes, which may not have been enough for differences in learning gains to emerge. The current study took place in a more realistic classroom setting (the previous study was a pull-out study), allowing the students to speak more freely and to work with the tutor across multiple sessions. The study involved 189 students and ran over five class periods to provide time for multiple units to be completed. To test our *hypotheses* that a collaborative ITS can effectively support young learners, and that collaborative and individual learning have different strengths for conceptual and procedural knowledge, we used two separate 2-condition between-subjects designs. Each compared a collaborative and an individual learning conditions, one with a conceptually-oriented set of tutor problems, and one with a procedurally-oriented set of tutor problems. This design allows us to compare between individual and collaborative learning within both conceptually-oriented and procedurally-oriented tutors. The design is not a true 2x2 design. The conceptual and procedural tutors cannot be directly compared because they support different learning goals.

Methods

Before we present the study in more detail, we describe our ITS designed to support fractions learning.

Tutor design

Informed by prior work on the Fractions Tutor (Rau et al., 2012) and our previous collaborative tutors for equivalent fractions (Olsen et al., 2014a), we developed a new ITS for a range of fractions units. To compare students working collaboratively to students working individually, we built two parallel versions of the tutors for use in our study, one for collaborative learning and one for individual learning. The ITS versions were built with the Cognitive Tutoring Authoring Tools (CTAT), extended to support collaborative tutors (Olsen et al., 2014b). For both collaborative and individual learning, we created two sets of tutor units, focused on either procedural knowledge or conceptual knowledge. For both sets, the units covered were naming, making, equivalent, least common denominator, comparing, adding, and subtracting fractions. For each unit, there were eight problems. All of the problems within a unit were of the same type and focused on the same procedure or concept. The problems were divided between graphical representations with two problems focusing on circles, two problems focusing on rectangles, and four problems focusing on number lines. Although the conceptually-oriented and procedurally-oriented tutors covered the same units, the materials were fundamentally

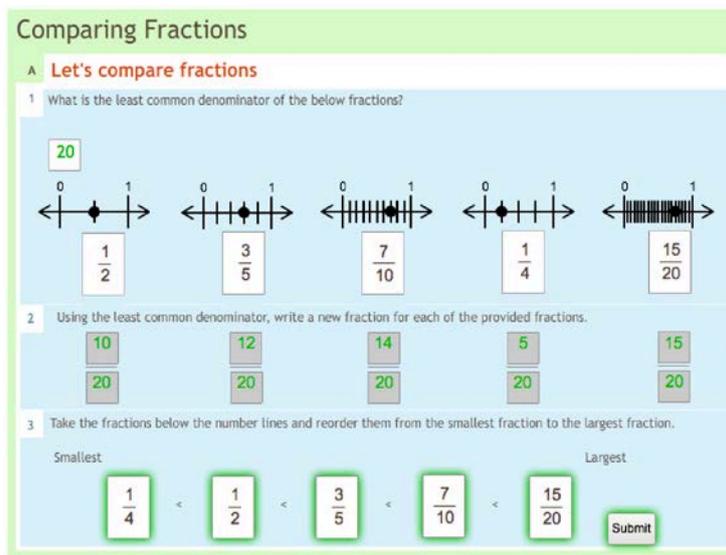


Figure 1. An example of an individual procedural problem for comparing fractions. The students go through the steps of converting fractions as a general mathematical procedure for solving this type of problem.

different to match different learning goals. For the procedural problem set, the students went through the steps needed to solve the problem. For example, for the comparing fractions unit, students would first find the least common denominator for the fractions they were trying to compare (see Figure 1) and then convert all of the fractions using this common denominator. Once they had completed this process, they were asked to put the fractions in order, from smallest to largest. For the conceptual problem set, the students were asked to fill in answers to see a pattern in the fractions, and then to fill in the blank for sentences to generalize the patterns to fractions concepts. For example, for the comparing fractions unit, the students were asked to mark if fractions were larger or smaller than one another. These fractions consisted of one where the denominators were the same, one where the numerators are the same, and one that is mixed (see Figure 2). The students were then guided to understand that when numerators are the same, then the fraction with the smaller denominator is larger, and when the denominators are the same then the fraction with the larger numerator is larger. These fill in the blanks were reworded for every other problem so students could not just memorize the slot that each answer should go into.

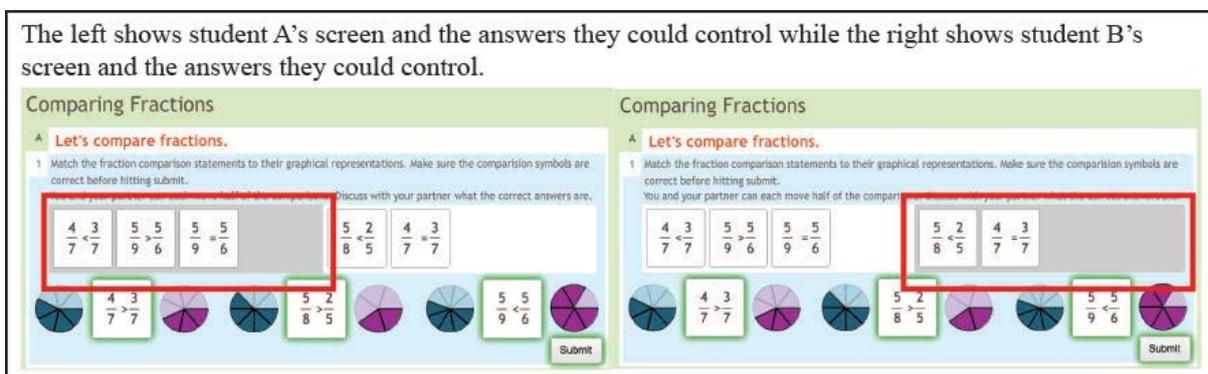


Figure 2. An example of a collaborative conceptual problem for comparing fractions. In panel A, each student is responsible for half the answers while in panel B the students share the responsibility for the questions.

In addition to the cognitive support normally provided by an ITS (step-level guidance for problem solving) that is used in both the individual and collaborative tutors, the collaborative tutors are also supported with embedded collaborative scripts for each tutor problem to provide social support for students (Kollar, Fischer, & Hesse, 2006). The collaborative tutors support synchronous, networked collaboration, in which collaborating students sit at their own computer and have a shared (though differentiated) view of the problem state and different actions/resources available to them. The students sat next to each other and communicated through speech. The embedded scripts supported collaboration through a distribution of responsibility to create accountability and interdependence (Slavin, 1989). The students were responsible for different parts of the problem (see Figure 2).

Often in the problem, each student had some action that only they could take for the step to be completed. This supported both students in needing to contribute. In Figure 2, each student has a set of the compared fractions. The students cannot fill in all three of the answers by themselves because they do not have access to all of the correct answers. They need to work together with their partner to fill in all of the blanks and submit the problem. This distributes the responsibility for each of the steps across both students. In addition, some steps created interdependence between the students. An example is for equivalent fractions that if one student entered a numerator, this would influence what the correct denominator would be for their partner. The embedded scripts provided support for the collaboration to be directly integrated into the ITS. Besides the embedded collaboration script, all collaborative and individual tutors were designed to be identical to allow for a comparison. During the time with the ITS, students who were collaborating sat next to each other but worked on different computers. They communicated through speech, which was recorded.

Experimental design and procedure

We conducted a study with 189 4th and 5th grade students from two schools across two school districts. The students came from a total of nine classrooms and five teachers. The experiment took place during the students' regular class periods. All students worked with the fractions ITS described above, either on conceptually-oriented tasks only or on procedurally-oriented tasks only, and working either collaboratively all the way through or individually all the way through. Thus, we had 4 conditions, which, as discussed above, we analyze as two separate 2-conditions designs. Students were assigned to individual and collaborative conditions based on class for a quasi-experimental design. This random assignment based on class was done to limit the disruption to the class. There were five classes that were assigned to work collaboratively and four classes that were assigned to work individually. Within each class, teachers paired their students based on students who would work well together and had similar math abilities. These pairs were then randomly assigned to work on the procedurally-oriented or the conceptually-oriented problem sets. Within the class there was an even split between students working on conceptually-oriented problems and students working on procedurally-oriented problems.

During the study, if a student's partner was absent in the collaborative conditions, the student would be paired with another student working on their same type of problem set for the remainder of the experiment. If there were two students that needed partners who had worked together before, they were paired together. If there were an odd number of students who needed a partner, then one student would work individually for the day. The teacher informed student pairings with a new partner when there was more than one option. When students started with a different partner from the day before, they would begin on the problem set at the place of the student that was less far in the problem sets.

The study ran across five days for each class where each day the students had 45 minutes to work with the tutor. On the first day, the students took the pretest individually. When they completed the pretest, they moved onto a tutorial that gave some instruction on how to interact with the tutor; otherwise, this was done on the second day. The students then worked with the tutor for the next three days in their condition. On the fifth day, the students took a posttest individually.

Test items

We assessed students' knowledge at two different times using two equivalent test forms in counterbalanced fashion. The tests targeted both conceptual and procedural knowledge types for all conditions. Each test had 20 questions, 10 procedural and 10 conceptual where six were isomorphic with the main six fractions units and four were near transfer targeting the four upper level fractions units. For each question on the test, the students were able to get a point for each step completed correctly. On the tests there were 23 possible conceptual test points and 68 procedural test points. Because of the discrepancy in points for the different types of knowledge, for both conceptual and procedural test scores, the percentages were used for all analyses.

Results

Due to absenteeism during the study, only 146 of the 189 students were used for the analysis. Students were excluded if they missed either the pretest or posttest. They were also excluded if they missed more than 1 day of working with the tutor. In the collaborative condition, students were excluded if they had more than a total of two partners during the tutors. There was no significant difference between conditions with respect to the number of students excluded, $F(3, 185) = 0.72, p = .54$. There was also no significant difference between conditions on the pretest score for either the conceptual test items, $F(3, 142) = 0.49, p = .69$, or for the procedural test items, $F(3, 142) = 0.68, p = .57$.

As discussed, we used a separate 2-condition between-subjects design to compare an individual and a collaborative learning condition for conceptually-oriented and procedurally-oriented tutor problem sets. For the

analysis, the data was thus treated as two separate data sets within which students working collaboratively or individually could be compared. Out of the 146 students used in the analysis, 70 students worked with the procedurally-oriented ITS, and 76 students worked with the conceptually-oriented ITS.

Pre/posttest learning gains

To investigate whether students learned using our tutors, and if there was a difference in learning between the students working collaboratively and students working individually within the two tutor problem sets, we used a multilevel approach to take into account the repeated measures of the pretest and posttest and differences between teachers. Within this analysis, we treated all students as individuals. We conducted a hierarchical linear model (HLM) with student at the first level and teacher at the second level. At level 1, we modeled the pretest and posttest scores, and at level 2, we accounted for random differences that could be attributed to the teacher. We did not include dyads as a level because of the added complexity of some students working with no partner (i.e. individuals), some students having one partner, and some students having two partners because of absenteeism. We are aware of non-independence issues such as common fate and reciprocal influence that may impact our results (Cress, 2008). We measured the effect size with Pearson's correlation coefficient (r) where 0.1 is a small effect size, 0.3 is a medium effect size, and 0.5 is a large effect size.

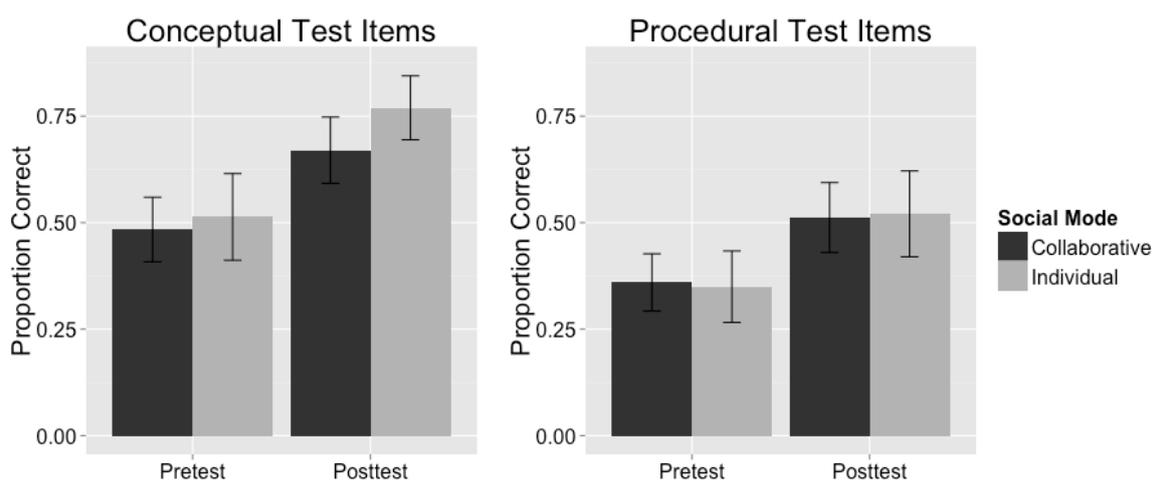


Figure 3. Learning gains on the conceptually-oriented tutors. Conceptual test items on the left and procedural test items on the right. Collaborative students are black and individual students are light gray.

First, we analyzed data from students who worked with the *conceptually-oriented tutors*. We tested whether there were statistically significant pre/post learning gains and we compared the learning gains between students working individually and students working collaboratively. We did so separately (with separate HLMs) for gains on the conceptual test items and gains on procedural test items. For the conceptual test items, a significant effect between pre and posttests was found with posttest having a higher score, $t(74) = -9.29, p < .001, r = .73$, while no significant effect for individual/collaborative learning, $t(70) = -1.25, p = .22, r = .15$, nor interaction, $t(74) = 1.47, p = .15, r = .17$, was found (see Figure 3). For the procedural test items, a significant effect between pre and posttests was found with posttest having a higher score, $t(74) = -6.80, p < .001, r = .62$, while no significant effect for individual/collaborative learning, $t(70) = 0.35, p = .72, r = .04$, nor interaction, $t(74) = 0.39, p = .70, r = .05$, was found (see Figure 3). Thus, although there was learning from pretest to posttest on both conceptual and procedural test items, there was no difference in learning between students working individually and students working collaboratively. To investigate whether there was a difference in the time it took students to complete the posttest, we ran a t-test. No significant difference was found for the time it took students to complete the posttest between the students working individually and students working collaboratively, $t(74) = -1.16, p = .25$.

Second, we analyzed data from students who worked with the *procedurally-oriented tutors*. We tested whether there were statistically significant pre/post learning gains and we compared the learning gains between students working individually and collaboratively. We did so separately (with separate HLMs) for gains on the conceptual test items and gains on procedural test items. For the conceptual test items, a significant effect between pre and posttests was found with posttest having a higher score, $t(68) = -4.49, p < .001, r = .48$, while no significant effect for individual/collaborative learning, $t(64) = 0.68, p = .5, r = .08$, nor interaction, $t(68) = -0.56, p = .58, r = .07$, was found (see Figure 4). For the procedural test items, a significant effect between pre and posttests was

found with posttest having a higher score, $t(68) = -8.40, p < .001, r = .71$, while no significant effect for individual/collaborative learning, $t(64) = 0.04, p = .97, r = .005$, nor interaction, $t(68) = -0.57, p = .57, r = .07$, was found (see Figure 4). Similar to what we found with the conceptually-oriented tutor, there was learning from pretest to posttest with the procedurally-oriented tutor, on both conceptual and procedural test items, but there was no difference in learning between students working individually and students working collaboratively. Also, no significant difference was found for the time it took students to complete the posttest between the students working individually and students working collaboratively, $t(68) = -0.49, p = .62$.

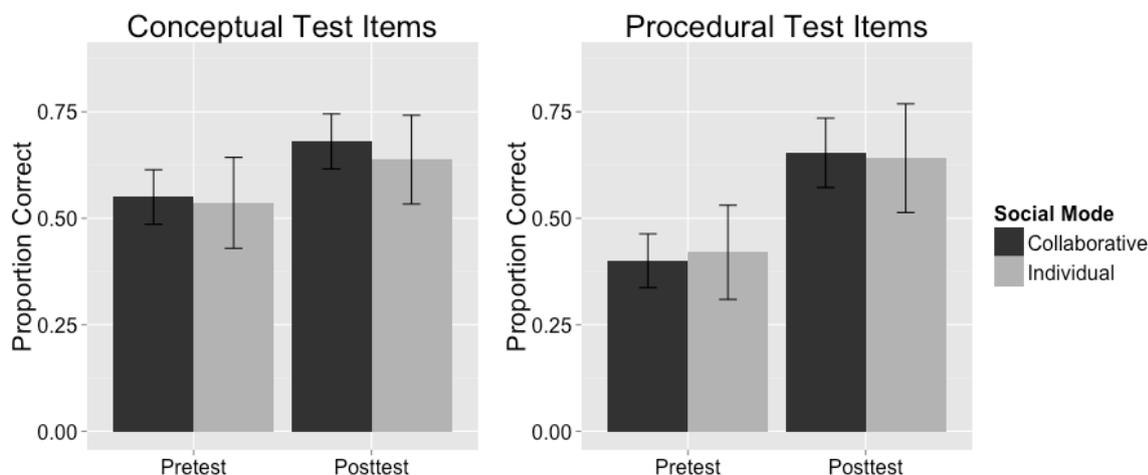


Figure 4. Learning gains on the procedurally-oriented tutors. Conceptual test items on the left and procedural test items on the right. Collaborative students are black and individual students are light gray.

Time on ITS

Table 1: Mean time and problems completed on the ITS for each condition and standard deviations

Tutor Materials	Social Mode	Time in Minutes: M (SD)	Problems Completed: M (SD)
Conceptually-oriented	Collaborative	73.32 (13.24)	44.08 (11.73)
	Individual	84.16 (18.11)	45.24 (14.50)
Procedurally-oriented	Collaborative	72.78 (13.34)	44.80 (10.29)
	Individual	84.11 (20.68)	45.30 (12.65)

For the study, ample time was given to students to work with the ITS with the expectation that all students would complete all of problems for the assigned units. Using two t-tests, we found no difference between students working individually or collaboratively in the number of problems completed for the conceptually-oriented tutors, $t(74) = -0.38, p = .7, r = .04$ or the procedurally-oriented tutors, $t(68) = -0.18, p = .86, r = .02$ (see Table 1). However, because the number of problems was fixed, students could work at their own pace and would finish the tutors at different times. Using two t-tests, there was a significant difference between students working individually/collaboratively on the time spent on the tutor for the conceptually-oriented tutors, $t(74) = -2.99, p < .01, r = .33$, and the procedurally-oriented tutors, $t(68) = -2.78, p < .01, r = .32$, where, surprisingly, students working collaboratively spent less time on the ITS (see Table 1). The students working collaboratively were able to complete the same number of problems as the students working individually but in less time.

Discussion and implications

The results of our study showed significant learning gains for students working collaboratively with both the procedurally-oriented and conceptually-oriented tutors. There was no difference in the learning gains of students working collaboratively versus students working individually. However, students working collaboratively completed the same number of tutor problems in less time on the tutor. These results confirm our hypothesis that young learners can be successfully supported in learning through the use of a collaborative ITS and indicate that a collaborative ITS is a viable option to use in the classroom.

The results did not support our hypothesis that students working on conceptually-oriented tasks would benefit more from collaboration and students working on procedurally-oriented tasks would benefit more from working individually. These results are consistent with the findings from our previous work where we also found no difference in learning gains (Olsen et al., 2014a). However, learning with an ITS individually has been shown to be very successful, especially within mathematics (Ritter et al., 2007; Rau et al., 2012), and it may be very difficult to design an intervention that can be added to an ITS to increase learning above what can be achieved working individually. Collaboration adds an extra layer of complexity that might be expected to inhibit the learning process, even within an ITS. Yet both in our previous work and in the current study, we found evidence that collaborative learning with an ITS can be more efficient than learning individually with an ITS. In our prior study, students had the same learning gains when working collaboratively as when working individually, but they had practice with fewer problems (Olsen et al., 2014a). In our current study, all students solved the same number of problems (i.e. had the same amount of practice), but the students working in the collaborative condition spent less time on the tutor. This was surprising; we expected to find the opposite (Lou et al., 2001). Collaboration may increase the time spent on each problem if students are discussing the solution. This is time that would not be spent in the individual condition. Our findings that students collaborating spent less time on the tutor could be due to students making fewer errors when collaborating, so they spent less time fixing those errors. To investigate this question, future work would need to analyze the learning that happened within the tutor and how it may have changed over time. By analyzing the process data, we could better understand what actions the students were taking with the tutor and where the efficiency gains were made.

While the collaborating students learned as much as their classmates who worked individually with regard to domain knowledge, they may have had more of an opportunity to develop their math reasoning skills and social skills by working with a partner. In a collaborative setting, students need to be able to construct their arguments well enough for their partner to understand their reasoning and are given the opportunity to ask questions (Chi & Wylie, 2014). This provides them with the opportunity to develop their math reasoning and to critique the reasoning of others. A next step would be to analyze the process data from the collaborative conditions to assess what opportunities the students had to explain their mathematical reasoning and to understand what the collaboration processes looked like that led to learning. A limitation of this study is that we did not assess social skills or mathematical reasoning skills for the students, which is where we would expect students collaborating to benefit beyond those working individually, as Rummel and Spada (2005) found that students who had an opportunity to collaborate had better knowledge about collaboration skills.

In summary, our findings show that both collaborative and individual learning with an ITS can be beneficial to young students. However, the process that the students go through for the learning may be different when working collaboratively compared to individually as indicated by less time being spent working with the tutor when working collaboratively. Students may benefit from having the opportunity to work both individually and collaboratively in a domain. Future work is needed to investigate how to best combine individual and collaborative learning and where in the process each would be appropriate. Individual and collaborative learning may lead to different learning processes, which raises the intriguing possibility that a combination of these modes might support more robust learning.

References

- Baghaei, N., Mitrovic, A., & Irwin, W. (2007). Supporting collaborative learning and problem-solving in a constraint-based CSCL environment for UML class diagrams. *International Journal of Computer-Supported Collaborative Learning*, 2(2-3), 159-190. doi:10.1007/s11412-007-9018-0
- Chen, W., & Looi, C. K. (2013). Group scribbles-supported collaborative learning in a primary grade 5 science class. In D. D. Suthers, K. Lund, C P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 257-263). Springer US.
- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219-243.
- Common Core State Standards Initiative (2015). Standards for Mathematical Practice: Construct viable arguments and critique the reasoning of others. Retrieved November 10, 2015 from <http://www.corestandards.org/Math/Practice/MP3/>
- Cress, U. (2008). The need for considering multilevel analysis in CSCL research—An appeal for the use of more advanced statistical methods. *International Journal of Computer-Supported Collaborative Learning*, 3(1), 69–84.
- Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R. (2010). Using intelligent tutor technology to implement adaptive support for student collaboration. *Educational Psychology Review*, 22(1), 89-102.

- Hausmann, R. G., Chi, M. T., & Roy, M. (2004). Learning from collaborative problem solving: An analysis of three hypothesized mechanisms. In K.D. Rorbus, D. Gentner, & T. Regier (Eds.), *Proceedings for 26th Annual Conference of the Cognitive Science Society* (pp. 547-552). Mahwah, NJ: Erlbaum.
- Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts—a conceptual analysis. *Educational Psychology Review*, *18*(2), 159-185.
- Lazakidou, G., & Retalis, S. (2010). Using computer supported collaborative learning strategies for helping students acquire self-regulated problem-solving skills in mathematics. *Computers & Education*, *54*(1), 3-13.
- Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, *71*(3), 449-521.
- Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education*, *20*(6), 507-528.
- Moss, J. (2005). Pipes, Tubes, and Beakers: New approaches to teaching the rational-number system. In J. Bransford & S. Donovan (Eds.), *How people learn: A targeted report for teachers* (pp. 309-349). Washington DC: National Academy Press.
- Mullins, D., Rummel, N., & Spada, H. (2011). Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *International Journal of Computer-Supported Collaborative Learning*, *6*(3), 421-443.
- Murray, T. (2003). An overview of intelligent tutoring system authoring tools: Updated analysis of the state of the art. In T. Murray, S. B. Blessing, & S. Ainsworth (Eds.), *Authoring tools for advanced technology learning environments* (pp. 491-544). Springer Netherlands.
- Olsen, J. K., Belenky, D. M., Aleven, V., & Rummel, N. (2014a). Using an intelligent tutoring system to support collaborative as well as individual learning. In S. Trausan-Matu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), *Proceedings for 12th International Conference on Intelligent Tutoring Systems* (pp. 134-144). Berlin: Springer.
- Olsen, J. K., Belenky, D. M., Aleven, V., Rummel, N., Sewall, J., & Ringenber, M. (2014b). Authoring tools for collaborative intelligent tutoring system environments. In S. Trausan-Matu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), *Proceedings for 12th International Conference on Intelligent Tutoring Systems* (pp. 523-528). Berlin: Springer.
- Rau, M. A., Aleven, V., Rummel, N., & Rohrbach, S. (2012). Sense making alone doesn't do it: Fluency matters too! ITS Support for robust learning with multiple representations. In S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings for 11th International Conference on Intelligent Tutoring Systems* (pp. 174-184). Berlin: Springer.
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346-362.
- Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive Tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review*, *14*(2), 249-255.
- Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *The Journal of the Learning Sciences*, *14*(2), 201-241.
- Slavin, R. E. (1989). Cooperative learning and student achievement: Six theoretical perspectives. *Advances in Motivation and Achievement*, *6*, 161-177. Greenwich, CT: JAI Press, Inc.
- Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology*, *31*(2), 207-220.
- Tsuei, M. (2011). Development of a peer-assisted learning strategy in computer-supported collaborative learning environments for elementary school students. *British Journal of Educational Technology*, *42*(2), 214-232.
- Walker, E., Rummel, N., & Koedinger, K. R. (2009). CTRL: A research framework for providing adaptive collaborative learning support. *User Modeling and User-Adapted Interaction*, *19*(5), 387-431.
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197-221.

Acknowledgments

We thank Amos Glenn, Ran Liu, and the CTAT team for their help. This work was supported by Graduate Training Grant # R305B090023 and by Award # R305A120734 from the US Department of Education (IES).

Integrating Science and Writing in Multimedia Science Fictions: Investigating Student Interactions in Role-taking

Shiyan Jiang, Ji Shen, and Blaine E. Smith
s.jiang@umiami.edu, j.shen@miami.edu, blaine.smith@miami.edu
University of Miami

Abstract: We explore characteristics of group interactions in an afterschool program designed for middle school students to practice integrated STEM learning and digital literacies. In the program, students self-selected roles (e.g., writers, scientists, artists, and engineers) and worked in small groups to create multimedia science fictions. In order to understand the high integration level between science and writing in one focus group's final product, this study examines students' collaborative learning processes, the characteristics of their interactions, and their role-changing patterns. We found that the high integration level between science and writing may be explained by (a) the interactions between the writer and the scientist during discussion and multimodal composing, and (b) the catalyst role of the artist in the group discussion.

Keywords: interdisciplinary learning, role-taking, multimodality, science literacy, digital literacies, collaborative learning

Introduction

Scientific literacy lies in the intersection between the literacy skills of speaking, listening, reading, and writing, and scientific practices such as evaluating science ideas and communicating science ideas to others (Glynn & Muth, 1994; Norris & Phillips, 2003). While there is still no universal agreement on the definition of scientific literacy, the literature explains that writing is one critical tool for promoting scientific literacy. Along with the need for developing literacy skills on traditional print texts, many have called for an expanded view of literacy where the learner is supported as both a critical consumer and skillful producer of digital multimodal texts (NCTE/IRA, 2012). Consuming, integrating, and creating multimodal texts in digital environments involves new ways of thinking (Leu, Kinzer, Coiro, & Cammack, 2004; Mayer, 2008), including the task of making meaning between the interactions of multiple media and modes (e.g., visuals, sound, text, and movement) and moving beyond disciplinary constraints (Jewitt, 2009).

Deep learning requires an environment in which students take on identities they value and become heavily invested (Gee, 2005). This perspective on identity has been broadly studied in the scientific literacy field. For instance, Reveles, Cordova, and Kelly (2004) found that students could be supported in scientific literacy when they were provided with opportunities to enact the identity of scientist in reading, writing, speaking, and thinking about science. The same perspective on identity was applied to digital literacies education. For example, Vasudevan, Schultz, and Bateman (2010) analyzed the multimodal storytelling composing process of fifth graders and found that the students experienced meaningful learning by taking the role as writers and building literate identities on authorship and authority towards stories they wrote. However, there has been much less work devoted to support developing students' discipline-specific identities in interdisciplinary environments. In particular, there is a paucity of research examining the collaborative learning processes of adolescents embodying different real-life identities (e.g., writer, scientist, and designer) while working in small groups.

Integrating perspectives from science literacy and digital literacies, we developed an afterschool program aimed at engaging middle school students in collaboratively learning through making multimedia science fictions. The learning objectives for participants included developing digital literacies and STEM practices, forming creative and collaborative habits of mind, and learning about different STEM careers. In an earlier analysis, we developed a matrix to examine if the student-generated fictions integrated science and literacy well (Jiang, Shen, & Smith, 2015). In this study, we focused on student collaboration processes in order to understand the mechanisms that contributed to the varied integration levels of their final products. The following research question guided this study: What are the key characteristics of group interactions related to student roles that contribute to high integration between science and writing?

Methods

Program design

Our afterschool program was first piloted in spring 2015 in a public middle school in an urban Southeastern city in the United States. The research team included two university professors (one focused on STEM education and the other on literacy education), three doctoral students, and one undergraduate student. The two professors taught the program and the other team members assisted with the program development and research. Sixteen students (grade 6-8) participated in the study initially. Students worked in small groups of three to four to create digital multimodal science fictions. Each team member could self-select one of the following roles (with the requirement that each team has to have a writer and a scientist): (1) *Writers* were in charge of developing the science fiction narrative based on brainstorming sessions and discussions with group members; (2) *Scientists* were responsible for monitoring the inclusion and accuracy of science vocabulary, concepts, and background knowledge, as well as linking their story to “science entries” where they elaborated on science concepts; (3) *Artists* led the creation of visual and audio representations for all main characters and/or scenes in the story; (4) *Engineers* were accountable for designing buildings, vehicles, and settings for the story. Despite the differentiated roles, the team members were asked to collaborate with each other on their individual tasks. For instance, the artist may create a drawing of a character and the writer can write a paragraph to describe the character. The pilot study included nine weekly sessions (approximately 1 hour each). Table 1 outlines the session plans. Throughout the sessions, students were introduced to several technological tools, including Scratch, Bitstrips, and other tools for creating multimodal artifacts.

Table 1: Session plans

Date	Research Team	Student Activity	Expected outcome
Jan 15	-Introduce the project -Prepare some science fiction topics for discussion	-Understand the different roles for the project -Brainstorm/share fiction topics	-Students understand what participation in this project entails.
Jan 22 (Session two)	-Science topic/resources presentations -Tutorial on iKOS for writing science fictions -Scratch training session	-Create iKOS accounts -Learn Scratch for creating animations -Students form groups with guidance	-Student groups formed -Student roles decided -User accounts created -Students learn how to and Scratch
Jan 29 (Session three)	-Tutorial on story writing -Share ZEE’s story (and elaborate on the roles)	-Small group works on their plot and further discuss each person’s role and responsibility -Small group debrief	-Each group has an outline of a story
Feb 12	-Bitstrips Training session	-Small group discusses their story -Peer feedback session -Learn Bitstrips for creating characters	-Each group learns what their peer group works on -Students learn to use Bitstrips
Feb 19	-Tutorial on how to include different science resources	-Small group works on their story/final product -Science knowledge check, search and sharing -Peer feedback session	-Each group learns what their peer group works on -Students understand the science involved in their fictions
Feb 26	-Milestone checks with each group	-Small group works on their story/final product -Group progress check	-Students meet the milestone check (Need to provide specific criteria/guidelines)
Mar 5 (Session seven)	-Role debrief	-Small group works on their story/final product -Role debrief (with groups of students of the same roles)	-Students refine understanding of their roles in relationship to their team work
Mar 12	NA	-Small group works on their story/final product	-Finish final products
Mar 19	-Evaluation	Final Presentation	-Present final products

All students used iKOS (ikos.miami.edu), an online knowledge building platform (Shen, Jiang, Chen, & Namdar, 2014), to create knowledge entries and multimodal science fiction stories. Figure 1 shows some screen shots from student artifacts.

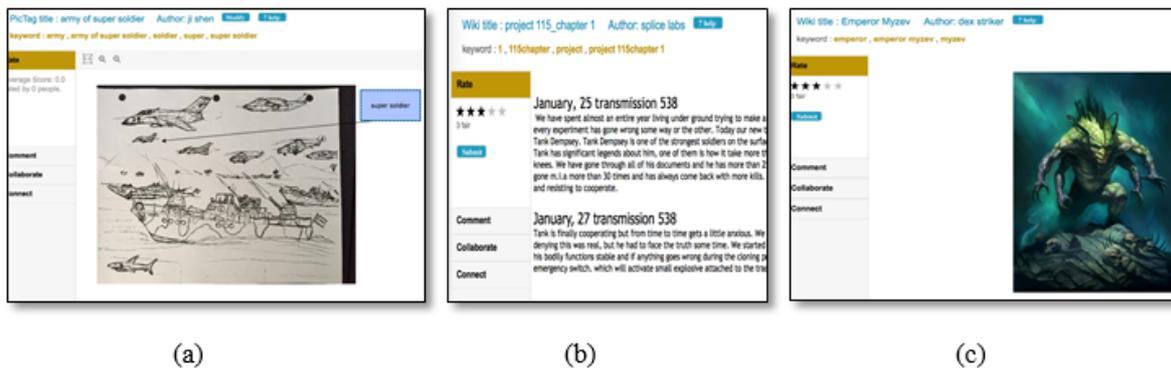


Figure 1. Sample screen shots from student artifacts including uploaded student drawing related to their fiction (a), story in a lab log style (b), and graphics included in the story (c).

Data collection and analysis

Three of the four groups completed their final stories, entitled “Teleportation”, “ReptilianKing”, and “SuperSoldier”. Our earlier analysis (Jiang et al. 2016) examined the integration level of students’ multimodal science fictions. One of the integration aspect indicated how well science concepts were interwoven into the science fiction story line. For example, here was a sentence that was coded as high integration in story “ReptilianKing”: “The rebels released a gas Exavier had produced called Antidotum that could render Vennenum useless by having it combine with the Anidotum producing a harmless gas (a new substance) instead, in a chemical reaction”. The sentence demonstrated that an imaginary chemical reaction was incorporated in its key plot: Exavier, the main character, developed a material called Antidotum to protect the rebels from chemical weapon attack. It could form chemical reaction with Vennenum (a substance mined and used by the army of the Reptilian King as a chemical weapon) and release harmless gas. In contrast, story “SuperSoldier” had only included science vocabulary such as “clone” but did not weave these scientific ideas in their story narrative.

Among the three stories, the story “ReptilianKing” had the highest integration between science and story based on the aforementioned analysis. In this study we examined students’ group interactions and team dynamics that could have contributed to high integration in their final fictions. Our analysis drew on the multiple sources of data we collected: Screen captures together with audio records of students’ conversations; video records of students as they collaborated in small groups; knowledge entries including the science fiction chapters created in iKOS; semi-structured individual interviews with students to better understand their learning experience in the program.

Table 2: Coding categories that focus on examining whether students were enacting presumed roles

Code	Description	Example
Expected role = In-action role	Students act as roles that are the same as their self-selected roles (i.e., writer, scientist, artist, and engineer) during their speaking turns.	Student A: “what is the setting? How about in an alternate dimension?” (Student A, the expected “writer”, was discussing the story setting.)
Expected role ≠ In-action role	Students act as roles that are different from their self-selected roles during their speaking turns.	Student B: “Patrik (the name of a character), how about Logan?” (Student B, the expected “scientist”, was discussing the story characters.)
In-action role is general	Roles that students act as are general or obscure.	Student C: “what’s your idea?” (This statement from student C does not entail any role-specific information.)

We examined program sessions when the students had an opportunity to discuss and work on their project in small groups. We transcribed the discussion and removed off-task turns. We then coded the transcripts turn-by-turn using Atlas.ti 6.2 following the coding categories emerged from the analysis (Table 2). We specifically examined if the students were enacting their presumed roles in the project. Meaningful episodes were then selected to illustrate issues pertaining to the research question on investigating key characteristics of group interactions. Each episode consisted of a small set of consecutive turns among the group members. We focused mainly on the interaction between the scientist and the writer since the integration between science and writing was the focus of the study. In this paper, we only report findings from the group that demonstrated the highest integration level (group “ReptilianKing”).

Findings

The analysis described in this section examined key characteristics of group interactions that related to student roles that contribute to high integration between science and writing in story “ReptilianKing”. The storyline of the science fiction was centered on two characters who travel through a wormhole to rescue a partner only to get caught in a war between the Reptilian King and his rebels. The story had 2938 words in 160 sentences, among which 21 sentences contained science ideas. Students represented three roles in the group while creating “ReptilianKing”. According to the first session survey, the scientist and the artist picked their first preference as their roles and the writer picked his second preference as his role.

The process: characteristics of group discussion

We selected three sessions to analyze student interactions because these sessions provided opportunities for them to work in small groups on their fiction writing and they also represented different phases into the project. In session two students formed small groups, picked preferred roles, and brainstormed fiction topics. In session three, students further brainstormed details in their story. In session seven, they worked toward finishing up their main story line.

Turn and time distribution across roles

We first looked at each member’s contribution through turns they took. A turn is defined as the time period from the speaker starts talking to either a new speaker stops the speaker, or the speaker is cut-off, or the speaker finishes his or her turn in an overlapped transition, talking simultaneously with a new speaker who takes the subsequent turn (Silverman, 2006). There were 51, 105, and 59 turns in sessions 2, 3, and 7 respectively (Table 3).

Table 3: Distribution of participation in terms of number of turns and time in seconds

Session #	Writer	Scientist	Artist	Total
S2	16 (46.1)	16 (73.3)	19 (82.5)	51 (201.9)
S3	41 (351.0)	33 (130.3)	31 (100.1)	105 (581.4)
S7	22 (63.0)	14 (47.6)	23 (76.0)	59 (186.6)

We also tallied the time each speaker took (see Figure 2). In session 2, there was no notable difference in both the time and turn distributions across different roles. In session 3, the writer dominated the group discussion according to the time distribution illustrated in Figure 2. He tried to describe the whole story (e.g., plots and characters), and the scientist contributed science ideas and some plots most of which were included in the final story. The artist also wanted to add some plot ideas, but most of them were not included in the final story. In session 7, most of the time the artist explained her designs of characters and the writer gave some suggestions on the designs. However, the scientist rarely engaged in the discussion of designs. Instead, he pointed out that the writer’s entry was less scientific and proposed to write another entry based on the writer's entry but in a more scientific way. That’s why the scientist took the least talking time in session 7.



Figure 2. Time distribution according to role among the team members across sessions

Interaction patterns between the scientist and the writer

While examining and reflecting on the interactions between the scientist and the writer, we derived three interrelated themes that point to why this group produced a high science integration story. One thing to note is that this group had 3 members only; so only writer, scientist, and artistic were represented.

Theme 1: Adding science vocabulary to the story by the scientist. The scientist helped to add scientific vocabulary into the story while the writer described the plot. In the following excerpt from session 3, the writer explained that the main character developed a new technology that had the ability to make traveling possible to an alternate universe and then the scientist brought up the term “wormhole technology” without further scientific explanation.

Writer: No, he is a human who is trying to use a new technology, a portal technology
Scientist: Wormhole technology.
Writer: Yeah, wormhole technology. To use it to transport, two porters, one in the northern part and one in the southern part...

In their final story, there were some sentences coded as low integration, such as the following one: “Rogue ... activated the device and created a wormhole of gigantic proportions”. These sentences contained science vocabularies, but they were not central to the story development.

Apparently, this kind of interaction between the scientist and the writer may not directly improve the integration level of the story. However, we want to stress that adding scientific vocabulary increases the potential for high integration since it can trigger the discussion about the science idea.

Theme 2: Building on each other’s ideas to co-develop the story line. In this group, the writer and the scientist also frequently built on each other’s ideas. For instance, the following excerpt occurred in session 3:

Writer: It is kind of like an alternate universe, where, let's see, where men did not evolve, its species dominated. It took its place. Like the reptilian.
Scientist: Yeah. The meteor did not hit the earth. And it did not die out.
Writer: The dinosaurs
Scientist: But that means that they have evolved eventually because it has been a long time.
Writer: Yeah, it's intelligent. So our characters should be intelligent but looks like dinosaurs.

The writer thought that the setting was an alternate universe where human did not evolve and the scientist explained that the alternate universe should be a place that was different from the earth because it was hit by the meteor. Enlightened by the scientist’s explanation, the writer set the species in the alternate universe to be species evolved from dinosaurs. This discussion was reflected in their final story with the following high integration sentence: “Yes, for you see the meteor you said was supposed to hit our planet as it did your Earth never quite hit us, therefore our species (species evolved from dinosaurs) didn’t become extinct and we managed to survive and evolve.’ said Muktav”. These sentences not only contained science vocabularies with explanations but also elaborated on key elements in the story (in this case, the origins of key characters in the story).

Theme 3: Rewriting the writer’s entry to elaborate on the science aspect. The scientist rewrote the writer’s entry that contained a brief explanation on science ideas from the perspective of story plots and the writer hyperlinked the scientist’s entry with their group final story. As shown in the following excerpt, the scientist proposed to rewrite the writer’s entry in a more scientific way. The writer agreed and promised to hyperlink his entry with the scientist’s entry. The rewriting behavior could also explain why the integration level is high because the scientist and the writer had a common understanding on the science idea in this context.

Scientist: How about making an entry about the science? The entry you created is more about the story. I will make it more scientific.
Writer: And when you finish everything, I will hyperlink to that one (I created which is less scientific) and it will be a hyperlink on a hyperlink.

The writer hyperlinked the scientist’s entry “particle collider” with their final story. The following high-integration sentence in the final story reflected their discussion: “[Now nicknamed MOMENTUM, due to it causing molecules to bump into each other at an accelerated rate, causing a large dense mass to be created similar

to [particle collider \(hyperlink to the scientist's entry\)](#) creating both black holes and wormholes] had reached its and further use would cause it heat up and shut down". In addition to providing an intensive description of the science idea, these sentences explained the functionality of the key device in the story: the "MOMENTUM".

Role-changing patterns

Figure 3 shows students' different role-changing patterns across sessions. Data points above (below) the x-axis represent that the student's in-action role is the same as (different from) his/her expected role. Data points on the x-axis represent that the student's in-action role is obscure or general.

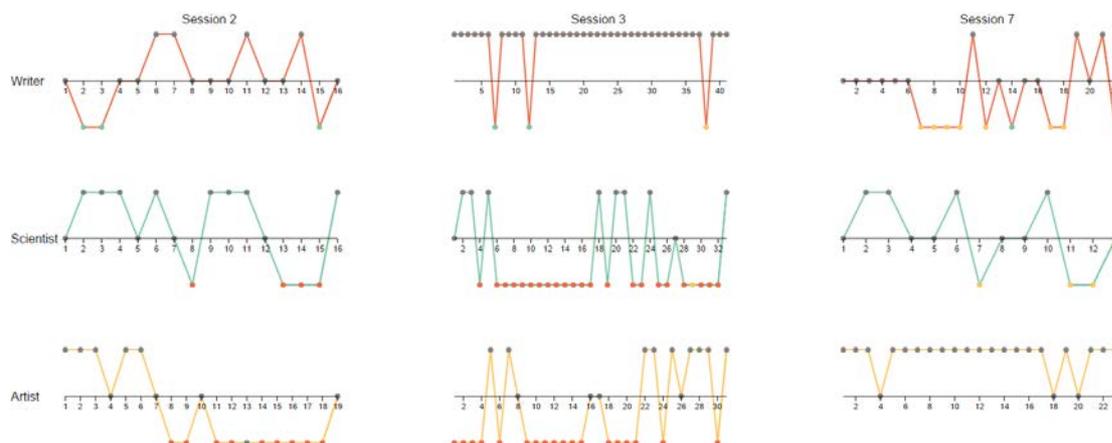


Figure 3. Role-changing pattern. Red represents writer; green represents scientist; yellow represents artist.

The writer. Overall, the writer played a leadership role in the group to coordinate with team members to push the story line for the group. He was also flexible at times to act as either a scientist or an artist. In session 2, the student mainly played the role of a coordinator. For instance, he invited other team members to express their ideas by stating sentences like "What type of Sci-Fi you want to write about?" He also started to take the role of scientist but gradually switched to be the writer. In session 3, he dominated the whole group discussion as a writer to discuss details of their story. For a couple of times he changed his role to be a scientist to respond to the scientist's suggestions. In session 7, he first worked as a coordinator and then he enacted as an artist to provide suggestions on the artist's designs. The writer's coordinating behavior might make the other group members feel that he was the dominator. He remarked in his post interview, "Although she [the artist] would probably say I am the leader, but I don't know, maybe I am".

The scientist. Overall, the scientist was faithful to his expected role to a great extent; he was also flexible to enact other roles. In session 2, the student started the group collaboration by taking his expected role to provide science ideas. He also enacted as a writer a few times later in the session when discussing with the team members about the story. In session 3, he constantly switched from scientist to writer to discuss the plot besides providing science concepts for the story. Recall on that session, a writing workshop was provided to all the students. In session 7, he worked more as a scientist and coordinator, but also contributed a few times as a designer. The scientist's role changing pattern was consistent with his interview response. He stated in the post interview, "I was supposed to be the engineer and scientist but turned out I did more work as a scientist or a writer sort of in between". Intuitively, the student's changing roles between scientist and writer helped contribute to the high integration of their final story.

The artist. Overall, the artist was very clear about her role in the group. At times, she enacted as a writer in order to join the conversation between the other two members. At the beginning of session 2, the student's in-action role was the same as her expected role, but later she changed her role into a writer to join the discussion between the writer and the scientist on brainstorming topics that they could write about. In session 3, she started as a writer to engage in the group discussion on describing plots and characters. Later, she disengaged herself from discussing the plots of the story since she was more interested in designs. Therefore, she changed back to be an artist to describe characters and search for pictures that match the descriptions online. In session 7, she showed her designs to the other two members and shared her insights about the designs. The observations also matched her interview responses: "My role is the artist and I animate everything. And I make sure the animations go smoothly and all of the visual effects. (Researcher: So did your role change?) Kind of, because I also did some of the other information and stuff, mostly I'm the artist".

Even though the artist's clear self-identification did not necessarily enhance an efficient interaction between the writer and the scientist, she served as a catalyst that initiated their interaction.

- Artist: It could be interesting if her age reversed. And her appearance will look differently.
- Scientist: Ok. The trick is due to wormholes, it works with time staff, and she has aged significantly than they should be. Significantly less, or
- Writer: Yeah, let's do it.
- Scientist: So Rogue reversed her age.
- Writer: She was originally 30 years old.

The artist brought up the idea of reversing age when discussing what would happen after one of the characters went through the wormhole. The scientist justified that it could be caused by the fact that the wormholes worked with time stuff, and the writer agreed on the idea and described that the character was originally 30 years old.

It is reasonable to argue that the high integration between science and writing for this group's final fiction was instantiated in their group interaction patterns. Specifically, the writer and the scientist enacted each other's roles frequently in order to carry out a fruitful conversation on both the story line and the science component. At times, the artist served as the catalyst for the interaction between the writer and the scientist.

Discussion

Our study examined how students' small group discussions and interactions were linked to their final products. We found that there were three different interaction themes between the writer and the scientist that possibly contributed to the high integration of their multimodal science fiction story: (1) The scientist adding scientific vocabulary into the story line; (2) team members building on each other's ideas to develop the story line; and (3) the scientist rewriting the writer's entry in a more scientific way. It is important to note that these themes were interrelated during their collaborations. The first theme is the base for the other two because the group would have science ideas to discuss and explore further only if someone brings in science vocabularies. The second and the third themes represented interactions from different angles: one from students' interaction on oral discussion and the other from students' interaction on composing artifacts.

We believe that these themes have practical implications in facilitating student collaboration beyond the specific context of our study. For instance, collaboration among learners can be facilitated at the surface level (e.g., bringing in different vocabularies) or at a deeper level (e.g., co-constructing artifacts). The key challenge is how to move the team members from the surface to the deeper level (e.g., using the surface level as a base for more fruitful collaboration). In addition, our themes also suggest that there are different objects students can co-construct (e.g., the conceptual storyline vs. the concrete knowledge entries in iKOS). A key research question is how to exploit the affordances and constraints of (and bridge accordingly) these different types of objects so that they mutually support each other.

Our research also examined the three students' role-changing patterns to explain the high integration level of the final story. Students self-selected their roles (accepted roles) as a writer, a scientist, or an artist at the beginning of the project, while their in-action roles were changing over time during their discussion. They tended to adopt and traverse their roles based on their interests, group dynamics, and the needs of the group. By comparing students' role-changing patterns across sessions, we see how the writer played a leadership role in addition to his expected role; the scientist was flexible in changing roles while fulfilling his expected role; and the artist was clear on her interest and utilized her advanced skills to be an artist. These observations pointed to the nuanced, dynamic, and complex processes in a collaboration environment. Besides team members understanding and enacting their specific roles in the collaboration process, our case hinted certain productive role-taking patterns: team members need to be flexible to flow between roles at times, (some) are willing to take on a leadership role, and (their participation) can elicit further interactions between other team members. Future empirical research should be devoted to understanding how we can design specific mechanisms to support these (or other) productive role-taking/changing patterns.

References

- Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-Learning and Digital Media*, 2(1), 5-16.
- Glynn, S. M., & Muth, K. D. (1994). Reading and writing to learn science: Achieving scientific literacy. *Journal of Research in Science Teaching*, 31(9), 1057-1073.

- Jewitt, C. (2009). *The Routledge handbook of multimodal analysis*. New York, NY: Routledge.
- Jiang, S., Shen, J., & Smith, B. (2016, April). Assessing students' scientific literacy in collaborative science fiction writing. Poster presented at the *2016 Annual Meeting of the American Educational Research Association (AERA)*, Washington, D.C.
- Leu, D.J., Jr., Kinzer, C.K., Coiro, J., & Cammack, D.W. (2004). Toward a theory of new literacies emerging from the Internet and other information and communication technologies. In R.B. Ruddell & N.J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1570–1613). Newark, DE: International Reading Association.
- Mayer, R. E. (2008). Multimedia literacy. In J. Coiro, M. Knobel, C. Lankshear & D. J. Leu (Eds.), *Handbook of research on new literacies* (pp. 359-376). New York, NY: Lawrence Erlbaum.
- National Council of Teachers of English/International Reading Association. (2012). *NCTE/IRA standards for the English language arts*. Retrieved from <http://www.ncte.org/standards/ncte-ira>
- Norris, S.P., & Phillips, L.M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87 (2), 224–240.
- Reveles, J. M., Cordova, R., & Kelly, G. J. (2004). Science literacy and academic identity formulation. *Journal of Research in Science Teaching*, 41(10), 1111-1144.
- Shen, J., Jiang, S., Chen, G., & Namdar, B. (2014). Designing the Innovative Knowledge Organization System (iKOS) for Science Learning. Poster presented at the *Annual International Convention of the Association for Educational Communications and Technology (AECT)*, Jacksonville, FL.
- Vasudevan, L., Schultz, K., & Bateman, J. (2010). Rethinking composing in a digital age: Authoring literate identities through multimodal storytelling. *Written Communication*, 27(4), 442-468.

Process and Output: Relation Between Transactivity, Temporal Synchronicity, and Quality of Group Work During CSCL

Vitaliy Popov, University of San Diego, vpopov@sandiego.edu
Anouschka van Leeuwen, Utrecht University, A.vanLeeuwen@uu.nl
Stan C.A. Buis, Vrije Universiteit Amsterdam, s.buis@vu.nl

Abstract: Do the simultaneous alignment of student activities (temporal synchronicity) and students successively building on each other's reasoning (transactivity) predict the quality of collaborative learning products? A mixed method approach was used to study 74 first year university students who were randomly assigned to work in dyads on an ill-defined problem of biodiversity collapse in tropical forests within a CSCL setting. The study results revealed that neither temporal synchronicity nor transactivity correlated with the quality of group products. The qualitative analysis of chat transcripts showed the variability between the groups can be explained by group dynamics, students' prior knowledge, confidence in managing the learning task, collaborative strategy, and communication skills.

Introduction

Recent advances in the Computer-Supported Collaborative Learning (CSCL) literature indicate that the efficacy of collaborative learning effort is thought to be influenced by the extent to which students can ensure the consistency of the joint work product by temporally synchronizing their collaborative activities (Erkens, Jaspers, Prangmsma, & Kanselaar, 2005), and by transacting on each other's ideas (Stahl, 2013). Building on previous research, this study investigates whether these two properties could predict the quality of group products.

Synchronizing collaborative activities across time and appropriate distributions of efforts and resources are of critical importance to the group's performance (Erkens, Jaspers, Prangmsma, & Kanselaar, 2005; Rummel & Spada, 2005). According to Baker (2002), "the degree of alignment refers to the extent to which participants are 'in phase' with respect to different aspects of the problem-solving activity, that is, to what extent they are genuinely working together" (cited from Arvaja, Häkkinen, & Kankaanranta, 2008, p. 268). Collaborative problem solving is non-aligned or non-synchronized when, for instance, one partner focuses on individual achievement over collective teamwork or there is no mutual agreement on a chronological order of activities. Failure to maintain continuous attention and reflection on one's own understanding as well as fellow group members can negatively affect temporal synchronization and lead to process losses (Baker, 2002; Schneider & Pea, 2013; Rummel & Spada, 2005). The type of CSCL environment in terms of synchronous or asynchronous forms of working and communicating has distinct variations of temporal synchronicity and its measurement - ranging from being in the same working space at the same time to coordinated effort over time (e.g., 24-hour knowledge factory well known in CSCW).

It is assumed that students working in groups adopt shared understanding and negotiate the meaning about a topic by asking questions, discussing, explaining, and providing extra information to support their viewpoints (De Lisi & Goldbeck, 1999). This type of group discussion is known as transactive discussion, i.e. students successively build on each other's reasoning by interpreting the meaning of their logical statements on the task at hand (Teasley, 1997). The way collaborating students build on each other's contributions can be carried out at low and high levels of transactivity (Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2013). At the lowest level, simple consensus occurs when group members accept what is said or done without further discussion. At the highest level of transactivity, a joint decision is made as a result of a dynamic incorporation of both agreements and disagreements between partners (Molinari, Sangin, Nüssli, & Dillenbourg, 2008). Previous research has shown that the highest level of transactive discourse increases the probability that learners trigger cognitive activity fostering individual and group performance (Stahl, 2013).

The present study was undertaken to combine the lines of research explained above into one set-up. The following research question was addressed: *To what extent do temporal synchronization of collaborative activities and the use of high levels of transactive discourse affect the quality of group products during synchronous CSCL?*

Methods

Participants

The participants in this study were 74 first year university students enrolled in an Environmental Sciences MSc program in the Netherlands. The sample comprised of 18 Dutch and 56 international students; 53% were women. The age of the respondents ranged from 19 to 37 years, with a mean age of 24 years ($SD = 3.2$); 96% of the participants were under the age of 30. The participants were randomly assigned to work in dyads. One dyad was excluded from the study as they did not use the CSCL-environment for communication, but used face-to-face communication.

Assignment and procedure

This study used an assignment that was part of an introductory course for MSc students, called Principles of Environmental Sciences. To fulfill this assignment, students had to analyze the problem of biodiversity collapse in tropical forests. While collaborating in dyads within a synchronous CSCL environment students were expected to inductively solve an environmental problem, by following three consecutive steps: (1) analyzing the problem of biodiversity loss by identifying causes and effects, (2) proposing possible Responses (solutions) to avert the biodiversity loss, and (3) selecting the most viable ways to tackle the problem of biodiversity loss by prioritizing the Responses. Students were expected to fill in a Driver–Pressure–State–Impact–Response (DPSIR) model for an ill-defined environmental problem to which several solutions could be proposed (i.e. DPSIR is a framework that helps to identify and describe processes and interactions in human–environmental systems, Fortuin, Koppen, & Leemans, 2011).

In the study, there were one introductory (plenary) and two online group work sessions over three days. The overall time required for completion of the assignment was about 8.5 hours. In the plenary meeting the DPSIR model was introduced. In the first group working session on day 2, students started by spending a few minutes getting to know each other and then they continued by studying the task materials at an individual computer. Students were then given time to post their individual thoughts, and to exchange ideas with their peers afterwards. On day 3, students continued working on the collaborative problem-solving task and a solution evaluation phase took place, which consisted of three subtasks. Namely, making a DPSIR-model, making a list of possible Responses, and reporting the overall prioritization of the Responses. The assignment ended in a finished DPSIR model, which was assessed. All online activity was automatically captured in log files, which were further analyzed for temporal synchronicity and high level transactivity (see below).

The dyads collaborated in a digital learning environment called Virtual Collaborative Research Institute (VCRI; Jaspers, Broeken, & Erkens, 2004). The VCRI groupware program incorporated both personal tools (Sources-tool and Notes-tool) as well as shared tools (Chat-tool, Cowriter-tool, and Diagrammer-tool) as shown in Figure 1. The Chat-tool allowed students to communicate with their collaborative partner. In the Cowriter-tool students wrote their Responses. In the Diagrammer-tool students edited elements of the DPSIR model.

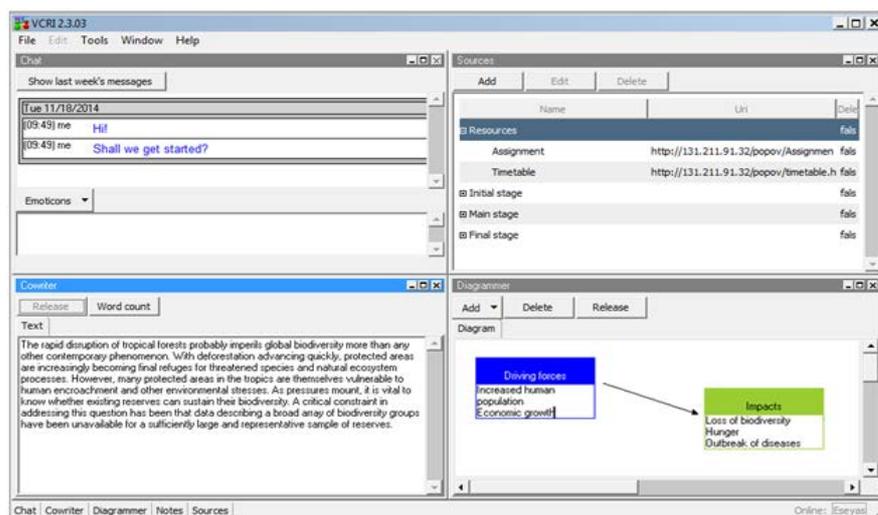


Figure 1. Screenshot of the VCRI environment, with Chat-tool (upper left), Sources-tool (upper right), Cowriter-tool (bottom left) and Diagrammer-tool (bottom right).

Instruments

Temporal synchronicity

Log files were automatically generated of each dyad's activities in VCRI, including which tools were opened and what actions were performed within the tools. It was quantified whether the members of a dyad mirrored each other's activity by opening and working in the same tool. An automatic coding mechanism, using Multi Episode Protocol Analysis (Erkens, 2005), was developed which coded at each time point whether or not the two members of the dyad were temporally synchronized, and if so, in which tool. The following two conditions had to be met in order for a dyad to be coded as temporally synchronized in a particular tool:

1) The dyad performed at least five consecutive actions within one specific tool, with the inclusion of Chat-messages. Because the Chat-tool was the primary medium for communication, the consecutive actions in a tool still count if an action in the Chat-tool occurs within this row of five activities.

2) The row of (at least five) consecutive activities includes activities produced by both members of the dyad. For example, five activities by member 1 in the diagrammer do not count as the dyad being temporally synchronized.

Thus, when a sequence of activities met the criteria, all activities in those sequences were coded as temporally synchronized. The percentage of the total number of activities that was coded temporally synchronized was used as the measure of temporal synchronicity for each dyad.

Transactivity

Student chats were analyzed for the occurrence of high levels of transactivity, namely *Integration* according to Noroozi's hierarchy (Noroozi et al., 2013; table 2), which means that learners adopt the perspective of their peers and build syntheses of the (counter) arguments uttered by their peers. A randomly selected third of all chats were coded by the three researchers and the outcomes were discussed until consensus about the coding was reached. After that the rest of the chats were divided between the researchers for coding all the remaining chats. In chat episodes where disagreement about the correctness of coding as high level of transactivity in the integration category occurred, consensus was reached through discussion by all three researchers. For all dyads the percentage of chat utterances that were coded as integration was then calculated by dividing the number of utterances in all integrative episodes by the total number of utterances in the chat conversation. Also, for each episode of high level transactivity discourse in the integrative category, it was checked whether the students transferred the discussed concept to the DPSIR diagrammer. The percentage of concepts in the diagrammer that was discussed on a high level of transactivity was then calculated.

Quality of students' group work

The assessment of the quality of the students' constructed DPSIR diagrams (obtained from the Diagrammer-tool) was made on a 5-step rating scale (5 being the best score) for three assessment criteria: width (the number of concepts in the diagrammer), correctness (the amount of concepts that is correct), and structure (the way concepts are grouped and related within the diagrammer). Two teachers coded all students' DPSIR diagrams independently and all disagreement and discrepancies were discussed until they reached an agreement. Both inter-rater agreement between two expert coders (Cohen's $k = 0.82$) (Landis & Koch, 1977) and intra-coder test-retest reliability for each coder for 15% of the data (90 % identical scores) were sufficiently high. Subsequently, all points assigned to each student dyad per criterion were added together and then divided by 3 (i.e. the total number of criteria). Each group of students could get a mean quality score of between one and five, which was converted to the Dutch 10 point grading scale. Scores below 5.5 were regarded insufficient to pass the course.

Analyses

A multiple regression analysis was performed to investigate whether temporal synchronicity and high level of transactivity predicted the grade that the dyads received for their diagrammer final product. The quantitative results were extended by adding qualitative descriptions of the dyads' collaborative processes.

For the qualitative analysis all dyads were categorized and placed in a table based on their qualitative characteristics. Dyads were categorized as having either high or low occurrences of high level transactivity (more or less than 33.3%), high or low temporal synchronicity (more or less than 33.3%), or an sufficient or insufficient grade (higher than or equal to or lower than 5.5). The collaboration of all dyads within each cell of the table was thematically analyzed by looking at the chat conversation as well as the ordering of events within the log files.

To analyze the chat conversations, an open coding approach was used to identify meaningful events such as the occurrence of disagreement between group members or reaching a shared solution to a question (cf., Barron, 2003; Rummel and Spada, 2005). Each event consisted of an episode in which each student had more than one

chat utterance. Using an iterative process, all chat transcripts were analyzed based on the emerged themes. Then, the three researchers wrote brief summaries of the way students collaborated. When there was doubt about the accuracy of a summary, a second researcher analyzed that dyad. From the summaries of all dyads per type, a characterization for each dyad type was composed.

Findings

Relation between high level transactivity, temporal synchronicity, and quality of group product

Table 1 displays the descriptive statistics for percentage of activities that was temporally synchronized, percentage of high level transactive utterances in the chat conversation, number of concepts in the diagrammer that was discussed transactively on a high level, and final grade for the diagrammer. As can be seen, on average the dyads achieved a relatively high level of temporal synchronicity, meaning that almost half of the time collaborating partners were using the same tools at the same time, including direct communication through the chat. The percentage of high level transactivity in the chat was relatively low (on average 15.09%), while about a third of the concepts in the final diagrammer product were discussed on a high level of transactivity. The average grade was 6.42, and out of 36 dyads, only five dyads did not achieve a sufficient grade.

Multiple regression analysis showed that neither temporal synchronicity ($t(35) = -.451, p = .655$) nor high level transactivity ($t(35) = .177, p = .861$) was a significant predictor of the grade that the dyads received for their diagrammer final product.

Table 1: Descriptive statistics for main variables

	M (%)	SD (%)	Min	Max
Temporal synchronicity	40.81	8.71	24.99	62.28
High level transactivity in chat	15.09	10.87	0.0	45.95
Number of Diagrammer concepts discussed transactively in chat	39.50	23.07	0.0	85.71
Grade	6.42	0.97	3.82	8.36

Qualitative description of dyads

No significant relationship between grades, high level transactivity and temporal synchronicity was found. In fact, both relatively low and high scores on high level transactivity lead to sufficient grades in this study. To explain these findings, qualitative descriptions of collaboration types that occurred in the dyads are provided in Table 2. These qualitative descriptions shed some light on the variability in outcomes between dyads. It also shows contrasts between groups with different characteristics and similar outcomes in grade, as well as groups with similar characteristics but different outcomes in grade.

A categorization of dyads was made based on three dimensions: synchronicity (low versus high), high level transactivity in discourse (low versus high), and grade (insufficient versus sufficient), leading to eight categories. Table 2 displays the distribution of dyads among these categories. As can be seen, one type of dyad did not exist in our dataset, namely the combination of low synchronicity, low occurrence of high level transactivity and an insufficient grade (Cell A).

In the sections below, we illustrate and contrast the types of dyads that showed similar process characteristics, yet differed in the grade they received. For example, dyad types A and C both show low temporal synchronicity and a low number of concepts discussed on high level transactivity, yet type A achieves an insufficient grade and type C a sufficient grade.

Dyad type A and C

Contrary to our expectations all dyads with lack of temporal synchronicity and a low number of concepts discussed on high level transactivity fell into dyad type C (sufficient grade) in our study. Further analyses of dyads type C revealed that in some cases the dyads consisted of two strong students that could manage their own tasks by distributing the workload and coordinating their actions, instead of trying to collaborate and learn from each other. In other cases, some of the students took almost full control of the task and finished the whole assignment on their

own. As an example, we consider one dyad from type C (see excerpt below), where one student is responsible for 98% for all activity in the Diagrammer-tool and as a group they had only 2.8% of the high level transactive discussions. Still, in contrast to dyad type A, this dyad achieved a high grade (8.4 out of 10 points). Student 2 gives Student 1 barely a chance to participate, as evidenced by the use of the phrase “my diagram”. As a result of this disproportional cooperation dynamic, the dyad achieves little high level transactivity and little temporal synchronicity.

- Student 1: Wow, you have finished the whole assignment.
 Student 2: I already made something but don't know if its correct
 Student 2: Do we need to put all driving forces, for instance, in 1 box or in seperate [separate] boxes?
 Student 1: I'm thinking about it now...I think we should ask someone else or the teather [teacher]
 Student 2: You're messing up my diagram :P
 Student 1: I'm trying to find out if I can add a box..
 Student 2: I can do it if you want
 Student 2: and you can chat me if you agree with what I add

Table 2. Summary of types of dyads, sorted according to temporal synchronicity, number of Diagrammer concepts discussed on high level transactivity in chat, and grade.

		Low occurrence of high level transactivity (number of concepts discussed on high level transactivity < 33.3%)	High occurrence of high level transactivity (number of concepts discussed on high level transactivity > 33.3%)
Synchronicity 0.0% - 33.3%	Grade < 5.5 Insufficient	A n = 0	B n = 1 SUMMARY: This particular dyad collaborated but was not critical of each other's work. They regularly wanted to show their work to the teacher, possibly because they were insecure.
	Grade > 5.5 Sufficient	C n = 5 SUMMARY: The dyads all divide the tasks between them, leading to cooperation instead of collaboration, and little temporal synchronization.	D n = 3 SUMMARY: The dyads all divide the tasks between them, leading to cooperation instead of collaboration, and no temporal synchronization. The dyads have some moments where their activities intersect and at those points they discuss the task material on high level transactivity.
Synchronicity 33.3% - 66.6%	Grade < 5.5 Insufficient	E n = 2 SUMMARY: The dyads have a lack of high level transactivity, in one case because the dyads divide the tasks and cooperate; in the second case because the students are too insecure to move beyond merely exchanging information.	F n = 2 SUMMARY: The dyads showed collaboration, including high level transactivity in discourse, and often challenged each other. The low grade may be explained in one case because the dyad based their discussion on incorrect information, and in the other case because the dyad spent too much time on figuring out the DPSIR model they had to construct.
	Grade > 5.5 Sufficient	G n = 10 SUMMARY: The dyads collaborated efficiently but showed no high levels of transactivity. There was a lack of challenging each other, instead quickly agreeing when the other proposed a solution. The dyads consist of motivated students with a focus on completing the assignment	H n = 13 SUMMARY: The dyads collaborated efficiently and with regular occurrence of high level transactivity. Ideas are constantly challenged in a friendly way, and the students regularly check whether they are still on the same page. Input from both students is combined into a co-constructed solution. The dyads consist of motivated students.

Dyad type B and D

Dyad type B occurred only once in our data set, whereas dyad type D had three instances. Both types of dyads, B and D, demonstrated similar style of working together. Namely, they discussed most subject matters in an open and constructive way. These dyads discussed their plans/activities verbally, agreed on what needed to be done, and then proceeded individually, without temporal coordination. Qualitative analysis of chat transcripts revealed that dyad type B had to ask the teacher for help over 10 times, partly because they thought they had to, and partly because they felt insecure about what they had produced (see excerpt below). In contrast, dyads type D managed their collaborative work without teacher's help and all content-related questions were discussed between the participants themselves.

- Student 1: so the diagrammer doesn't need to be modified .every item of response should be more specific
- Student 2: i don't know
- Student 2: maybe we could ask to a teacher isn't it?
- Student 2: did you saw what I posted about the 2 last points/
- Student 1: Yes
- Student 1: and i post the 1
- Student 2: it seems to be good
- Student 2: we should show the complete work to the teachers

Dyad type E and G

From the dyads that achieved a high level of temporal synchronicity in their collaboration and discussed a low percentage of their concepts of high level transactivity, 2 dyads scored an insufficient grade (dyad type E) and 10 dyads scored a sufficient grade (dyad type G). Overall, the dynamic of these two types of dyads was set on completing the DPSRI assignment, but not on learning from each other and understanding more about the subject matter. In situations when one student would pose a question or a suggestion, the other student would usually respond very briefly or agree very easily. The difference in grades may be explained by the dynamics of collaboration that these two types of dyads demonstrated. Collaborating students in both types of dyads, E and G, worked very closely with each other in terms of space/tools use, but dyads in type E spent more time figuring out the assignment and doubting actions (i.e., one dyad in type E used the word "Maybe" 42 times throughout their entire chat conversation) compared to dyads in type G. This is shown in the excerpt below.

- Student 1: but the problem is that I don't know if we have to focus on the case described in the article or if we can find another impacts that are not in the text
- Student 1: but I think you are right actually
- Student 1: and you speak about 50% forest loss but maybe we should put that in the pressues [pressures], as an explanation and proof of deforestation
- Student 2: maybe you are right
- Student 1: and the other ones, maybe we can develop a few by sentences from the text
- Student 2: I don't know if we can find some impacts that are not in the text, too. so, maybe now we can just use the things given in the article or in the film

Dyad type F and H

From the dyads with both high occurrence of high level transactivity and temporal synchronicity, only two dyads scored an insufficient grade (dyad type F), and 13 dyads scored a sufficient grade (dyad type H). This supports our expectation that high occurrence of high levels of transactivity and temporal synchronicity may lead to quality of group work. Further analyses of dyad type F shows that both dyad types are very polite, constructive and friendly in their chat. However, one of these two dyads had a low score (2 out of 5) on the correctness criterion, possibly indicating that they based the concepts in their diagrammer on misconceptions. The other dyad spent a lot of time figuring out what the assignment was about and what the idea behind a DPSIR model was, these difficulties in understanding the assignment seemed to be increased by language barriers between the two students in the dyad.

Twelve of the 13 dyads in dyad type H are characterized by challenging each other in a friendly and motivating way, showing patience and checking each other's ideas to come to co-constructed refined solutions

(see excerpt below). One dyad however, started working on the diagrammer very early on in the assignment, but then had an argument the students themselves called a ‘serious situation’. When the students realized the argument lasted for too long and they had to hurry to complete the assignment, they stopped arguing in the chat and were able to finish the assignment.

Student 1: Ah it belongs to air pollution

Student 1: or climate change

Student 2: true, but emissions is a pressure and it pollutes the air. Shall I add it to the pressures?

Student 2: and air pollutions keeps standing there as a state

Student 1: yes! perfect!

To summarize, the qualitative descriptions showed the variability between the groups in terms of characteristics such as group dynamic, students’ content knowledge, confidence in managing the learning task, collaborative strategy (cooperation versus collaboration), and communication skills.

Discussion

Findings and implications

In contrast to expectations, neither temporal synchronicity nor high level transactivity correlated with the quality of group products. Sufficient grades were achieved with only a high occurrence of high level transactivity, only a high level of temporal synchronicity, and even with low levels of both high level transactivity and temporal synchronicity. This suggests that none of these variables is vital for collaboration to lead to sufficient final product. The qualitative analysis of collaborating dyads showed that there were a number of variables, besides the two under direct investigation, that influenced whether a dyad would succeed or not. The clearest finding was that by discussing their activities and agreeing on a clear task division, two students with strong content knowledge can work independently from each other and still achieve a good grade.

The most surprising type of dyad we encountered was the type that scored an insufficient grade even though they had high levels of temporal synchronicity and high level transactivity. A possible explanation is that they reasoned on false beliefs, meaning that their misconceptions about the task material remained undetected by the students themselves. This finding shows the importance of both social aspects as well as cognitive/task-related aspects of collaboration. Perfect socio-collaborative skills are only beneficial when students challenge each other’s statements continuously and keep checking their output for correctness. Since it is a demanding task for teachers to monitor and regulate multiple collaborating groups at the same time (Van Leeuwen, Janssen, Erkens, & Brekelmans, 2015), automated moderation of discussions on both socio-collaborative and cognitive dimensions of collaboration could be beneficial. Some of the processes we observed in the dyads, such as the ‘overruling’ behavior when a dominant student was coupled to an insecure student, could be prevented with adequate moderation of discussions. Thus, there is a challenge here to take into account not only differences *between* dyads, but also differences *within* dyads.

Limitations and directions for future research

This study was carried out in an authentic learning environment, but the results of this study should be interpreted in light of some limitations. First of all, the task students worked on did not include an incentive to share knowledge - there was no dependency between the students within dyads. For example, if students had been given differing task materials, they would have had to share and discuss, and high level transactive discourse may have been more likely to occur. Another limitation of the study design is that we did not measure individual knowledge gains. It could be that although the investigated variables did not predict quality of the group product, the students in various types of dyads may have differed on individual knowledge gains.

Concerning temporal synchronicity, we checked whether both students had the same workspace activated, but not whether both students were actively contributing nor whether they actually acknowledged each other’s presence. It could be argued that the two students need to demonstrate awareness that “they are attending to something in common” (Tomasello, 1995, pp. 106). However, research on temporal synchronicity in the field of CSCL is relatively scarce (Rummel & Spada, 2005). It is not yet clear to what extent implicit coordination such as temporal synchronicity differs or correlates to explicit types of coordination (in which students openly discuss

coordination of activities), which is a much more common area of research (Järvelä & Hadwin, 2013). In this respect, we hope to have given input for future research.

Finally, in our analysis we did not take into account the role of students' cultural background. Students may have had communicative barriers that hindered high level transactive and synchronized collaboration (Popov, Biemans, Brinkman, Kuznetsov & Mulder, 2013). Indeed, our results pointed out the importance of group dynamics for dyads to succeed on the task. This could be a relevant direction for future research, also given the increasing numbers of massive open online courses (MOOCs) in which students from different countries and cultures participate. The observed differences, and different combinations of dyad characteristics, show the need for personalized support.

References

- Arvaja, M., Häkkinen, P. & Kankaanranta, M. (2008). Collaborative learning and computer-supported collaborative learning environments. In J. Voogt & G. Knezek (Eds.), *International Handbook of Information Technology in Primary and Secondary Education* (pp. 267-279). New York: Springer.
- Baker, M.J. (2002). Forms of cooperation in dyadic problem-solving. *Revue d'Intelligence Artificielle*, 16(4-5), 587-620.
- Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12, 307-359.
- De Lisi, R., & Goldbeck, S. L. (1999). Implication of Piagetian theory for peer learning. In A.M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 3-37). Mahwah, NJ: Erlbaum.
- Erkens, G. (2005). Multiple Episode Protocol Analysis (MEPA). Version 4.10 software. Utrecht University, The Netherlands.
- Erkens, G., Jaspers, J., Prangma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior*, 21, 463-486.
- Fortuin, K. P. J., Van Koppen, C. S. A., & Leemans, R. (2011). The Value of Conceptual Models in Coping with Complexity and Interdisciplinarity in Environmental Sciences Education. *Bioscience*, 61(10), 802-814.
- Jaspers, J., Broeken, M., & Erkens, G. (2004). Virtual Collaborative Research Institute (VCRI) (Version 2.0). Utrecht: Onderwijskunde Utrecht, ICO/ISOR.
- Järvelä, S., & Hadwin, A. (2013). New Frontiers: Regulating learning in CSCL. *Educational Psychologist*, 48(1), 25-39.
- Landis, J. R., & Koch G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174.
- Molinari, G., Sangin, M., Nüssli, M.-A., & Dillenbourg, P. (2008). Effects of knowledge interdependence with the partner on visual and action transactivity in collaborative concept mapping. In G. Kanselaar, J. van Merriënboer, P. Kirschner, & T. de Jong, (Eds.), *Proceedings of the Eight International Conference for the Learning Sciences (ICLS 2008): Part 2* (pp. 91-98). Utrecht, The Netherlands: ISLS.
- Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers and Education*, 61(2), 59-76.
- Popov, V., Biemans, H. J. A., Brinkman, D., Kuznetsov, A., & Mulder, M. (2013). Facilitation of Computer-Supported Collaborative Learning in Mixed- versus Same-Culture Dyads: Does a Collaboration Script Help? *The Internet and Higher Education*, 19, 36-48. doi: 10.1016/j.iheduc.2013.08.002
- Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting problem-solving in computer-mediated settings. *Journal of the Learning Sciences*, 14(2), 201-241.
- Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning*, 8(4), 375-397.
- Stahl, G. (2013b). Transactive discourse in CSCL. *International Journal of Computer-Supported Collaborative Learning*, 8(2), 145-147.
- Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaboration? In L. B. Resnick, R. Säljö, C. Pontecorvo & B. Burge (Eds.), *Discourse, tools and reasoning: Essays on situated cognition* (pp. 361-384). Berlin: Springer.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103-130). Hillsdale, NJ: Lawrence Erlbaum.
- Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of multiple computer-supported collaborating groups. *Computers in Human Behavior*, 52, 233-242.

The Learning Experiences of Youth Online Information Brokers

Jason C. Yip, University of Washington – Seattle, jcyip@uw.edu
Carmen Gonzalez, University of Washington – Seattle, cmgonzal@uw.edu
Vikki Katz, Rutgers University, vkatz@rutgers.edu

Abstract: In the U.S., there is a large proportion of families where one or both parents are English language learners (ELL). Children in these families often serve important roles as *brokers*, by engaging their linguistic capabilities, cultural familiarity, and technological skills to bridge their families' access to information resources. Despite the central role that child brokers play, scholars know little about how they search for, interpret, and translate online information. Using data from an exploratory study with Latino youth (ages 11-14) that involved interviews, online search tasks, and group discussions, we investigate the learning processes, challenges, and strategies that youth employ as they broker online information for their ELL parents.

Introduction

In multiethnic and multinational countries like the U.S., many immigrants face linguistic, cultural, and technological barriers to accessing information. To cope with these obstacles, parents as English language learners (ELL) often depend on their bilingual children to act as intermediaries to interpret and translate information (Katz, 2014B; Orellana, 2001). These youth often act as *brokers* through facilitating their parents' connections with technology and information (Katz, 2014B). Currently communities in the U.S. are experiencing demographic shifts in which youth brokering is becoming more prominent. In 2012, there were 25 million ELL individuals (9% of the population) ages 5 and older in the U.S. (Pandya, Batalova, & McHugh, 2011). This reflects an 80% growth in the number of ELL individuals between 1990 and 2010 (13 million in 1990 to 25 million in 2010). While prior studies document how youth brokers interpret a wide range of information (e.g., cultural, linguistic) for their ELL parents (e.g., Katz, 2014B; Orellana, 2001), we have less information about how these youths utilize information communication technologies (ICTs) and digital literacy skills and knowledge for brokering practices.

Without a deeper understanding of this phenomenon, we risk overlooking ways in which youth brokers develop creative learning strategies to negotiate high-pressure situations as they access vital information for their families. Connected learning is a learning process that is driven by personal interests, encouraged with peer and family support, and helps learners connect their out-of-school pursuits to formal academic and career possibilities (Ito et al., 2012). Youth brokers often facilitate information for ELL parents and build bridges between home and school (Orellana, 2001). By becoming more familiar with the challenges that youth brokers face during online searching for families, and what strategies they develop to support their families' information needs, we are seeking to develop more informed strategies to promote connected learning as part of supporting family learning and wellbeing. Our investigation expands work in the fields of information and communication by examining youth online brokers' processes and experiences as they navigate complex online information. Our focus on brokering moves away from a deficit perspective, which presumes families lack knowledge resources, and instead highlights the strategies that online brokers employ to broaden their information access. Therefore, our research questions are:

- RQ1. What learning roles do online information brokers play with their parents?
- RQ2. What strategies and challenges shape learning roles in online information brokers?
- RQ3. What are the affordances and limitations of digital technologies in brokering and learning?

First, we situate our investigation within prior research on youth brokering. Next, we highlight our theoretical framework using Vygotsky's (1978) social development theory and its affordance to understanding learning in youth online brokering. Finally, we present our findings and conclude with an examination of the role of digital literacy in online brokering, including and what opportunities are available for connected learning.

Background

Youth brokering for limited English proficient families

Lower-socioeconomic (SES) parents with limited English-language proficiency in the U.S. often depend on their children to search and interpret online information, since important resources (e.g., healthcare, education, social services) are predominately available in English (Katz, 2014A; 2014B). However, when online information is

translated into non-English languages, ELL parents may face accessibility issues related to limited literacy and education, cultural knowledge, and technical skills and access. For instance, websites translated into Spanish typically use one dialect; however, there are 20 distinct dialects of limited mutual intelligibility (Thompson, 1992). This makes reading and understanding translations from English difficult. As such, the responsibility of translating web content falls into the hands of the children of these ELL families. Although community resources and institutions can provide help with translation, information brokering is primarily an in-home activity. For instance, public libraries can offer help for information search, but lower-SES immigrant families may not access these resources due to busy work schedules, or fear of institutions (Gehner, 2010).

The role of learning in youth brokering

Although brokering can be quite a difficult task for youth, researchers from education, sociology, and communication have documented important opportunities for positive learning outcomes. Children often must take on “adult roles” as they become experts on behalf of their family; this reframing of roles often provides new learning opportunities (Eksner & Orellana, 2012). For youth, the act of brokering depends on high-level development of cognitive and social skill sets (Hall & Sham, 2007). Language brokering and biculturalism in youth has shown a positive role in Latino adolescent students’ academic performance (Buriel, Perez, Terri, Chavez, & Moran, 1998).

Orellana’s (2001) use of sociocultural theory to examine longitudinal studies of children’s brokering has contributed towards the understanding of skills development in both youth brokers and their parents. Youth participation with their families’ needs can help them engage in civic opportunities, character formation, and build stronger ties with their families (Orellana, 2001). For instance, Katz (2014A) notes that youth brokers often needed to translate complex health and medical information from doctors. Children framed these experiences with doctors as educational and informative of health issues. Treating these interactions as learning experiences helped to mitigate the difficulty of youth managing their parents and health providers’ expectations. Eksner and Orellana (2012) demonstrate the ways that language brokering with parents contributes to youth’s knowledge of the social world, skills development in home language and English, problem-solving skills towards meaning-making, and other learning opportunities. Katz (2014A; 2014B) examines media brokering and how children facilitate their parents’ understanding through media artifacts (e.g., print, online media) and mediated communication (e.g., phone calls, Internet). While current studies frame cultural, language, and media brokering as important to youth brokers’ development and learning, less is known the role of online information search and brokering for youth and their ELL families.

Theoretical framework

The analyses presented in this study are guided by Vygotsky’s (1978) social development theory, which posits that children’s learning and development cannot be separated from the early interactions with parents and other significant people in their lives. Development does not occur in universal stages; cultural context and social factors play a large role in shaping learning. Vygotsky’s “zone of proximal development” is the opportunity for children to engage in experimentation and social interaction to support learning beyond their individual capabilities. Therefore, in order for children to develop, they need access to more knowledgeable others, mainly parents, peers, and other significant people in their lives. When youth broker information for their families, such close and focused interactions can present learning opportunities for both the children and their parents.

Vygotskian theory (Vygotsky, 1978) suggests that these brokering moments can be quite powerful for learning. Youth brokers are guided by more expert others in language and home culture (i.e., parents, extended family). However, youth brokers must also act in an “expert” position, in which they articulate their understanding of complex online information to their parents, who are learning from their children (Dorner, Orellana, & Li-Grining, 2007). In this study, we specifically examine these learning interactions around digital technology, a phenomenon known as joint media engagement (Takeuchi & Stevens, 2011). Today, new ICTs have made collaborative learning between adults and children more two-way and less defined by parental authority. While we know more about learning together with children and families in mainstream populations (Barron, Martin, Takeuchi, & Fithian, 2009), we know that youth brokers are both teaching and learning about technology, developing fluency in their language, and code switching (Katz, 2014B). Therefore, joint media engagement in ELL immigrant families are quite diverse with families’ different priorities, challenges, and motivations (Katz & Gonzalez, 2015).

Methodology

This study focuses on how youth brokers engage in online information problem-solving tasks. This is a qualitative, exploratory research study that utilized semi-structured interviews, information search tasks for youth, and group discussions with youth.

Context and participants

We conducted this study with 10 youth participants (Table 1, 9 different families, all names are pseudonyms) who met the following criteria: 1) Self-identifies as the child of a Latin American immigrant; 2) Has at least one parent with limited facility in verbal or written English; 3) Between ages 11 to 14; and 4) Reports helping his/her parent(s) understand and navigate online content and ICTs.

Table 1: Demographic information of the youth brokers

Name (Pseudonym) and age	Country of origin	Family characteristics
Amanda (age 13)	Spain (mother)*	2 sisters (ages 5 and 9) and mom
Bella (age 14)	Dominican Republic	3 older sisters, 1 older brother (siblings not home) and mom
Betty (age 12)	Dominican Republic	1 brother (age 23), 1 sister (age 21), mom and dad
Don (age 11)	Dominican Republic	1 brother (age 23), 3 sisters (ages 23, 20, and 18), mom and dad
Fernando (age 13)	Dominican Republic	1 sister (age 18, not home), mom and dad
Jennifer (age 11)	Ecuador (mother) and Mexico (father)	1 brother (age 8), 1 sister (age 6), mom and dad
Marina (age 12)	Dominican Republic	1 sister (age 18), mom and dad
Nina (age 13)	Ecuador	1 brother (age 17), mom
Ricardo (age 13) and Raphael (age 13) (twins)	Mexico	3 brothers (ages 7, 14 and 18), mom and dad, aunt and older cousin (18)
*We included this participant because her family identified themselves as Latino. While her mother is from Spain, she did not report where her father is from. She has strong family connections to the Dominican Republic.		

This study occurred in a major metropolitan area in the Northeast U.S. In the middle school where we recruited participants, 72% of children were receiving free- or reduced-cost meals and 59% were Latin American. We chose Latin Americans as the focal group for this study because they are the fastest growing “minority” group in the U.S. (US Census Bureau, 2012), have the highest rates of ELL individuals, and report youth brokering as a common phenomenon (e.g., Orellana, 2001). In 2010, the Latino population in the U.S. accounted for 66% of the overall ELL population (16 million ELL Latinos) (Pandya et al., 2011).

Prior studies on youth Internet search studies have ranged between 30 and 90 interviews (e.g., Druin et al., 2009; Foss et al., 2013). Our sample size was smaller because (a) the study was meant to be exploratory and lay the foundation for a future study; and (b) brokering is a phenomenon that is not evenly distributed across a population, and that children do not necessarily admit to brokering for family information (Katz, 2014B). Furthermore, given that we were looking to identify children of immigrants, anxieties related to residency status prevented a broad sampling of the school’s students. However, prior qualitative research establishes that small samples (5 – 12 interviews and participants) have utility for theory building in new research areas (Baker & Edwards, 2012). We limited our investigation to youth between 11 and 14 years old because prior research indicated that middle school marks the beginning of parents’ intensive dependence on children’s brokering skills (Hall & Sham, 2007).

Research design and data collection

The instrumentation for this study is an adapted information search protocol from Bilal (2002) and Foss *et al.* (2013). We conducted 10 one-on-one interviews and search tasks with youth brokers in four sessions in an afterschool setting. After individual interviews, respondents participated in a moderated group discussion about their brokering experiences. Four of the five researchers on this project spoke Spanish proficiently or fluently and were able to engage in Spanish dialogue with the youth.

Interviews and search tasks: We used a semi-structured interview protocol (Merriam, 2009) to allow us to focus on our research questions, but with enough flexibility to identify emergent issues. We asked questions about family demographics, ICT usage, and general search behavior. Next, for 40 minutes, we asked youth to engage in a series of search tasks that were self-generated, imposed, simple, and complex. *Self-generated* tasks focus on searches with fewer constraints (Bilal, 2002; Gross, 1999). *Imposed* tasks provide limitations for children to search for information they would not normally search for (Gross, 2006). *Simple* tasks are baseline search tasks

to make sure children can do basic searches (Bilal, 2002). Finally, *complex* tasks provide a challenge and allow us to see how persistent and resourceful children are at searches (Bilal, 2002). Based on these criteria, we asked respondents to engage in the following tasks using school laptops:

1. If your parent needed information on the public schools in this area, what information would you look up for them? (*imposed and simple task*)
2. Can you show me how you find information for something your parents have asked you to search before? (*imposed task*)
3. Can you show me how you find information for a problem you think your parents might have today? (*self-generated task*)
4. If one of your parents wasn't feeling well or had been sick for a few days and wanted to know why, what information would you look up for them? (*imposed and complex task*)

During these search tasks, we had the respondent narrate how they were conducting the search and their opinions on the task. We video recorded screen interactions as the children explained their search tasks. During this time, we asked children to search as if their parents were present.

Group interviews and field notes: After the search tasks, we had the youth engage in a 20-minute group discussion about how they search for information for their ELL parents. Each group discussion (four total) had two to three youth from the interview sessions and two to three researchers. We conducted these group discussions to better understand what aspects of these experiences appear to be common among youth online brokers. After each session, we generated individual summative memos of our experiences and contributed to efforts to interpret the collected data, to varying degrees.

Data analysis

The data for analysis consisted of verbatim individual and group interview transcripts, transcripts of the video-recordings, and detailed field notes constituted. We adhered to the inductive analytical approach developed by Strauss and Corbin (2007), to develop themes and categories to capture children's search and learning experiences for their ELL parents. Two researchers open-coded the data independently for themes, such as parental interaction, affect, search challenges and frustrations, and strategies for brokering. We categorized, sorted, and compared the themes to further develop categories for analysis. We then systematically compared and contrasted the themes between the researchers. Following the open-coding analysis, we used axial coding to make connections between a category and its subcategories more explicit. We used selective coding to see if additional categories were needed. This sorting, comparing and contrasting was performed until we reached theoretical saturation and that no new codes or categories were generated.

Findings

We present our findings based on our research questions. First, we outline the learning roles children and their families take on in the online brokering process. Second, we highlight the different challenges and creative strategies that online youth brokers engaged in. Finally, we examine the affordances and limitations of ICTs for learning in the process of online youth brokering.

Child and family roles in online brokering and learning

Youth brokers in this study indicated that their parents' requests for information were a product of the specific kinds of challenges each family faced. Although we did not interview parents, we were able to ask youth brokers what kind of online searches their parents commonly assigned to them. We characterized one set of tasks as "non-urgent searches"; that is, searches that are important, but low in priority. These searches include maps and directions, online shopping, access to entertainment (e.g., videos, music), news, and recipes. Youth also indicated "urgent searches"; these online searches are high priority and relate to complex life issues such as health and medicine, school choice, and immigration issues. For all these searches, children take on a myriad of roles in learning and teaching with their parents. For example, youth brokers act as *synthesizers* and must quickly glean and explain information, their level of understanding operates across a spectrum of varying information. Some of these summaries of information are conducted in Spanish, even though the information is in English. For instance, we asked Betty to describe how she would explain information about school rankings to her mother; she did so in Spanish as the interviewer played the role of her mother:

Betty: *Mira mami, estos son los mejores escuelas que hay en la ciudad.* [Look mom, these are the best schools that there are in the city]. *Mira, estos son las Matemáticas y el Ingles. Y mira,*

como están haciendo entre las escuelas. Ellos hacen mejor que los otros en una escuela como esta, 4.0 y este 3.7. Quiere decir que esta es mejor que esta. [Look this is Math and English. And look at how the schools are doing. They are doing better than others in a school like this one, 4.0 and this one 3.7. That means that this one is better than this one.]

Although all ten youths indicated they conducted online brokering, not all youth brokers had the same skills for translation. Some youth brokers (Don and Marina) had developing Spanish skills and mostly relied on a combination of online consumer machine translation tools (i.e., Google Translate™) and their limited Spanish fluency. Other youth brokers were so fluent in Spanish that they acted not only as translators, but more engaged **editors**. For example, Fernando would spend time correcting longer Google Translate™ outputs for his parents to read (which could take up to an hour). We observed in this study that all the youth who were brokering urgent online information for their ELL families needed strong literacy skills in both English and Spanish. For some families, there is a high expectation that their children can translate any information. Bella noted, *“We don't get recognition because it's something we're supposed to do, if you're in a Dominican family. No, I'm serious; you're supposed to do it, you have to be the kid and translate.”*

Parents were not only depending on their children for online information brokering, they needed their children as **technology specialists**. Many parents were learning about technology from the youth brokers, as they were searching online together. For instance, Oscar explained how he was helping his mother understand how to access her paycheck online: *“I was helping her log on and then...As I was helping her, I didn't know what her password was and stuff...So I told her where to put it, then. So my mom didn't know where to put it anywhere.”* Other youth brokers explained that while they were searching for online information, they found themselves often teaching their parents how to address technical issues with the devices they were using. Interestingly, Betty indicated it was her mom that first taught her how to use the computer to search online. In this case, some ELL parents do have technology literacy to work with children on basic understandings of digital tools. As their children progress with the technology, parents may be asking for help with more advanced functionalities.

To be as resourceful as possible, youth needed to take on **social mediation** roles to find the help they needed as they searched for and interpreted information for their families. The online brokers had high reliance on their social networks to support such tasks. Respondents received help from the other parent, older siblings, extended family members, and even teachers and neighbors. The family unit itself takes on a **inquirer** role in social learning; that is, parents and other family members provide imposed queries (Gross, 1999) to the youth brokers, which often forces the youth out of their comfort zone. For example, the youth brokers in this study needed to search for things like *Remicade* (a drug) that may cause cancer, phyllodes tumors, immigration policies in online news sources, and school choice issue. These high-priority imposed queries, while incredibly difficult, push youth to be exposed to new online information they may not normally search for on their own personal interests. Unlike homework assignments from teachers, which can be ignored or forgotten, these online brokering tasks are obligations for youth and cannot be easily dismissed. This is a unique opportunity for learning to occur in a less structured way around technology. Both the ELL parent and the child interact in joint media engagement and can learn from the online search and brokering experience.

Challenges and strategies in online brokering for learning

Although learning together through online brokering processes can help strengthen family bonds (Katz, 2014), there are many challenges to learning that youth brokers encounter. The first challenge is dealing with **family pressures**; youth brokers indicated stress and mistrust in working together with ELL parents. Although technology for search is becoming more mobile (i.e., smartphones, tablets), the youth in this study indicated that their ELL parents preferred large displays from laptops and desktops so that they could sit together with their children in the brokering process. However, Nina noted the frustration of having her mother next to her: *“For me, sometimes I want to, like, find the right thing [online], but...I take a long time reading and then [my mom is] like, ‘Okay, can you please hurry up?’ or something. I'm like, ‘I'm trying to find the correct information for you, so relax!’”* Some youth indicated their ELL parents did not trust their children's online activity, which lead the parents to sit and monitor their children's search practices. Amanda explained, *“If their instincts tell them we're doing something wrong, they go crazy; like, ‘No, we're doing something wrong. That's not the right way.’ But you don't know what it means, so why are you saying that?”* Despite the tensions that the respondents noted, it is important to emphasize the shared commitment, persistence, and collaboration around these family needs.

A second challenge towards learning we observed was the developing **digital literacy skills** in the youth brokers. Digital literacy refers to the cognitive, motor, emotional, and sociological skills necessary to interpret and synthesize a wide range of digital information from multiple sources (e.g., Eisenberg, 2010). One of the pressing issues the youth faced in online brokering tasks was comprehending technical words and phrases.

Medical terminology, immigration policy terms, school choice, and even common everyday terms (e.g., “preservative”, “inseam”) were difficult to understand and translate for the youth brokers. While even fluent speakers of English could have trouble deciphering complex terms, ELL parents depend greatly on these youth brokers developing search skills. For challenging online information searches, youth had difficulty with query formation, filtering pertinent information, determining information reliability, and other known problems with searching and youth (Druin et al., 2009; Foss et al., 2013). While youth relied greatly on social networks for search help for searches, none of the youth brokers in this study mentioned local and community resources, such as schools, libraries, and community organizations, that could support their efforts. For instance, Fernando recalled that he needed to find information on a particular medicine for his mother: “*Well, my mom told me to search it [Remicade, a drug] up in the computer ‘cause that’s the injection that they’re gonna, that they’re telling her that she should have. But like, she doesn’t know if she should, ‘cause they said that [the drug] might give her cancer.*” To find more information, Fernando searched using the query, “Remicade effects on the immune system.” When he examined the search results, Fernando went directly to Google Scholar™, thinking the scholarly information there would be more reliable. He arrived at an esoteric scientific journal as the first result on the page. Although he believed that the information was reliable, he did not understand the information or was able to translate it. Fernando’s example is a case in which a librarian, teacher, or community member could be helpful and scaffold the search process. However, the sensitivity of the information and limited community can amplify the difficulty of youth online brokering and learning.

We would be regretful to only highlight the challenges to learning in brokering. On the contrary, under pressure and constraints, youth brokers came up with creative strategies to help synthesize and convey their knowledge. Funds of knowledge (e.g., Moje et al., 2004) are the everyday diverse learning experience that youth rely on. Here, joint media engagement around the technology, search tasks, and translation push children to come up with creative ways to convey complex information. One set of strategies involved using *gestures*. For instance, during the search task, Amanda described using hand gestures to describe the information she was explaining. Jennifer called what she was doing, using “charades” to demonstrate the importance of the online information she found. Other children *drew pictures*. Bella noted the difficulty in explaining what “preservatives” in food were and needed to draw out pictures of this concept. Finally, some children used *physical objects* to help their families out in online search. Marina, when she found a recipe online, translated the ingredients to her mother by finding and gathering the physical ingredients that were available. In order to describe what an “inseam” was for a school uniform website, Amanda needed to get a real pair of pants from the closet to show her mother. In these examples, we find that even though online brokering occurs in the digital space, learning and communicating digital information is embedded in physical spaces and is context-based (Lave & Wenger, 1991).

The affordances and limitations of digital technologies in brokering and learning

ICTs for online brokering gives many affordances for learning. The youth brokers in this study all indicated they had access to wireless Internet at home and multiple ICTs (e.g., desktops, laptops, smartphones, tablets, etc.). In this study, youth brokering occurred mostly in the home space, where privacy, comfort, and access occur together. In the home space, ELL parents and their children can come together to search together, particularly around larger screens (i.e., desktops, laptops) and can engage in joint media engagement towards problem solving. Youth brokers also had access to online consumer machine translation tools (i.e., Google Translate) and online dictionaries (i.e., Dictionary.com). This helped learners to access the Spanish language for quick translations and understandings.

However, access to technologies for brokering also have their limitations for learning. First, even though all ten online brokers indicated they have wireless Internet at home, digital technologies do not directly teach digital literacy skills. Youth brokers in this study expressed being overwhelmed by the online information they encountered. Learning about online technologies and digital literacy is still a social process. While the larger screens helped to bring the families together, youth brokers complained that dealing with their parents’ frequent requests and pressure was not necessarily a positive attribute for online brokering. For instance, Bella described that she and her mother could be at the computer for an hour going from site to site: “*We end up in stupid places because she wants to keep branching out somewhere.*” The user interface design of the technologies also did not support information accessibility for the youth brokers. For example, we noted that Google Translate™ and Dictionary.com™ do not use visual representations as a way to support understanding of the text, which can make multimedia learning more challenging (Mayer & Moreno, 2002). Finally, the youth in this study expressed fears and concerns about technology breaking. They explained about the anxiousness over computer viruses and explicit popup ads. These fears have been documented extensively in lower-socioeconomic Latino families (Katz & Gonzalez, 2015), and often prevent families from trusting their children to access online information. For instance, Fernando described explicit popups as making online brokering more difficult, especially when his parents are

sitting with him: “And I didn't click that [explicit pop-up ad] and it just gets me mad. And then they [parents] just, then they get mad. And they sometimes just take the computer away. It's not my fault when the ad appears.”

Discussion and conclusion

Vygotsky (1978) theorized that through social play, children and youth learn to develop higher order skill sets beyond their individual capabilities. Digital media scholars argue for the need to pay more attention to youth's personal interests and play in order to understand more about how youth are learning (e.g., Ito et al., 2012; Jenkins, Clinton, Purushotma, Robinson, & Weigel, 2006). Scholars are also examining more closely the role of parents in co-learning with children through play and digital technologies (e.g., Barron et al., 2009; Takeuchi & Stevens, 2011). This exploratory study demonstrates that digital technology also contributes to youth learning through supporting family needs. New ICTs allow digital collaborations and learning between adults and children to become more two-way, child-centered, and less hierarchical (Clark, 2011). In the case of youth online brokering, joint media engagement is less about parents learning about their children's digital interests, and more about how ELL families can thrive in the face of contextual, linguistic, and cultural constraints. Youth participants in this study indicated close interactions with their parents through various technologies. They reported their parents sat down with their children to access information for their family's wellbeing. While these interactions can be frustrating, they are opportunities for youth to interact with their parents more directly. Youth are also co-learning in this process. In the face of hardships, youth online brokers must become creative to adapt to the situation. They needed to rely on their social networks, find non-digital solutions to communication issues (e.g., using gestures, developing analogies, drawing pictures, finding physical objects), and quickly synthesize and translate complex information (for which they may not fully understand) into their parents' language. Youth brokers in this study were motivated to demonstrate what they could accomplish for their ELL families. For instance, Betty emphasized in the group interview that she was motivated to show off her abilities in search and translation, especially when her parents did not think she knew Spanish well enough.

In conclusion, the findings of this exploratory study highlight the need to examine brokering and family responsibilities in the context of digital literacy and connected learning. Our findings suggest that as more information (both urgent and non-urgent) is digitized and accessible online, youth brokers will face a flood of complex information decisions (e.g., query formation, information reliability) that will require deeper learning and connections to different domains. Online brokering and learning is not just about the learning dyad between the ELL parent and youth broker. We believe schools, libraries, community organizations, and other local institutions need to partner together with families to support digital literacy help connect learning practices in youth. For instance, youth brokers search for health issues for their ELL family could be augmented with science learning from schools, digital literacy skills from libraries, and local community supports. We believe that youth brokering process will be become more dependent on the integration and fluency of cultural, linguistic, and digital literacy skillsets. As such, access to technology for brokering online is not enough. This study advocates for a view of technology for *meaningful connectivity* for learning. While all the youths in this study had access to technologies for online brokering, our findings suggest a need for a closer examination beyond a simplistic view of the digital divide. Meaningful connections emphasize both the access to the technologies and the support for engagement to develop skills for learning (Katz & Gonzalez, 2015). We recommend future studies examine the learning experiences of online youth brokers and their ELL families in the contexts of their homes and domestic settings over a period of time.

References

- Baker, S. E., & Edwards, R. (2012). How many qualitative interviews is enough: Expert voices and early career reflections on sampling and cases in qualitative research. National Centre for Research Methods. Retrieved from <http://eprints.ncrm.ac.uk/2273>
- Barron, B., Martin, C. K., Takeuchi, L., & Fithian, R. (2009). *Parents as learning partners in the development of technological fluency*. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/ijlm.2009.0021>
- Bilal, D. (2002). Children's use of the Yahoo!igans! Web search engine. III. Cognitive and physical behaviors on fully self-generated search tasks. *Journal of the American Society for Information Science and Technology*, 53(13), 1170–1183.
- Buriel, R., Perez, W., Terri, L., Chavez, D. V., & Moran, V. R. (1998). The relationship of language brokering to academic performance, biculturalism, and self-efficacy among Latino adolescents. *Hispanic Journal of Behavioral Sciences*, 20(3), 283–297.
- Clark, L. S. (2011). Parental mediation theory for the digital age. *Communication Theory*, 21(4), 323–343.

- Dorner, L. M., Orellana, M. F., & Li-Grining, C. P. (2007). "I helped my mom," and it helped me: Translating the skills of language brokers into improved standardized test scores. *American Journal of Education*, 113(3), 451–478.
- Druin, A., Foss, E., Hatley, L., Golub, E., Guha, M. L., Fails, J., & Hutchinson, H. (2009). How children search the internet with keyword interfaces (pp. 89–96). *Proceedings of SIGCHI*. ACM.
- Eisenberg, M. B. (2010). Information literacy: Essential skills for the information age. *Journal of Library & Information Technology*, 28(2), 39–47.
- Eksner, H. J., & Orellana, M. F. (2012). Shifting in the Zone: Latina/o child language brokers and the co-construction of knowledge. *Ethos*, 40(2), 196–220.
- Foss, E., Druin, A., Yip, J., Ford, W., Golub, E., & Hutchinson, H. (2013). Adolescent search roles. *Journal of the American Society for Information Science and Technology*, 64(1), 173–189.
- Gehner, J. (2010). Libraries, low-income people, and social exclusion. *Public Library Quarterly*, 29(1), 39–47.
- Gross, M. (1999). Imposed queries in the school library media center: A descriptive study. *Library & Information Science Research*, 21(4), 501–521.
- Hall, N., & Sham, S. (2007). Language brokering as young people's work: Evidence from Chinese adolescents in England. *Language and Education*, 21(1), 16–30.
- Ito, M., Gutiérrez, K., Livingstone, S., Penuel, W., Rhodes, J., Salen, K., et al. (2012). *Connected learning: An agenda for research and design*. Chicago, IL, USA: MacArthur Foundation.
- Jenkins, H., Clinton, K., Purushotma, R., Robinson, A., & Weigel, M. (2006). *Confronting the challenges of participatory culture: Media education for the 21st century*. Chicago, IL, USA: MacArthur Foundation.
- Katz, V. (2014). Children as brokers of their immigrant families' health-care connections. *Social Problems*, 61(2), 194–215.
- Katz, V. (2014). *Kids in the middle: How children of immigrants negotiate community interactions for their families*. Rutgers University Press.
- Katz, V. S., & Gonzalez, C. (2015). Community variations in low-income Latino families' technology adoption and integration. *American Behavioral Scientist*, 0002764215601712.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.
- Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12(1), 107–119.
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: John Wiley and Sons.
- Moje, E. B., Ciechanowski, K. M. I., Kramer, K., Ellis, L., Carrillo, R., & Collazo, T. (2004). Working toward third space in content area literacy: An examination of everyday funds of knowledge and discourse. *Reading Research Quarterly*, 39(1), 38–70.
- Orellana, M. F. (2001). The work kids do: Mexican and Central American immigrant children's contributions to households and schools in California. *Harvard Educational Review*, 71(3), 366–390.
- Pandya, C., Batalova, J., & McHugh, M. (2011). *Limited English proficient individuals in the United States: Number, share, growth, and linguistic diversity*. Washington, D.C.: Migration Policy Institute.
- Strauss, A. L., & Corbin, J. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory*, 3rd ed. SAGE Publications.
- Takeuchi, L., & Stevens, R. (2011). *The new coviewing: Designing for learning through joint media engagement*. The Joan Ganz Cooney Center at Sesame Workshop. Retrieved from <http://www.joanganzcooneycenter.org/Reports-asc.html>
- Thompson, R. W. (1992). Spanish as a pluricentric language. In M. G. Clyne and W. de Gruyter (eds.). *Pluricentric languages: Differing norms in different nations* (pp. 45–70).
- US Census Bureau. (2012). *The 2012 statistical abstract - U.S. Census Bureau*. Retrieved February 6, 2015, from <http://www.census.gov/compendia/statab/>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*.

Acknowledgments

We want to thank Michael Levine and Lori Takeuchi at the Joan Ganz Cooney Center for supporting this work. We thank Rocio Almanza and Alexia Raynal for volunteering on this project. We thank the youth, teachers, and staff at the local school that participated in this study. This study is a part of the Families and Media Project and supported by funding from the Heising Simons Foundation and the Bezos Family Foundation.

How Do We Assess Equity in Programming Pairs?

Elise Deitrick, Tufts University, Elise.Deitrick@tufts.edu
R. Benjamin Shapiro, University of Colorado Boulder, Ben.Shapiro@colorado.edu
Brian Gravel, Tufts University, Brian.Gravel@tufts.edu

Abstract: We comparatively apply methods for assessing equity to find ways that allow us to better interrogate, describe, and understand the construct's complexities. To do so, we replicate prior work assessing equity in programming pairs (Lewis & Shah, 2015; Shah, Lewis, & Caires, 2014) by creating distributions of coded (1) computer use, (2) turns of talk, and (3) questions and commands; furthermore, we build upon their application by (4) examining students' embodied interaction and social positioning. Using these four methods to analyze the discourse of a pair of students programming, we identified numerous examples of positioning shifts, and see how inequities could potentially emerge through discursive mechanisms. We then reflect on what each of these four different methods foreground, their strengths and shortcomings, including their sensitivity to time-scope and pedagogy, and triangulate the quantified and qualitative results.

Keywords: equity, collaborative learning, research methods, computer science

Introduction

Many researchers are investigating issues of equity in STEM contexts, in particular, equity within collaborative STEM learning environments, like group work. A place that has received particular attention recently is Computer Science (Margolis & Fisher, 2003; Margolis et. al, 2010). Computer science, and in particular programming, which typically uses a computer, presents a particular challenge in studying. The literature contains studies that use multiple methods encompassing multiple aspects to look at how and why equity can emerge in computer use (Cole, 1995; Lewis & Shah, 2015; Shah, Lewis, & Caires, 2014). In terms of methods, these studies account for distribution of computer use as well as types of talk (e.g. questions and commands). These papers all contrast two groups doing a computer task and discuss how their selected research measures show different facets of social dynamics being more or less equitable given the group. In addition, their conclusions are similar in they cite within group context as having the power to invalidate their methods: Cole (1995, p.73) concludes that “the way the mouse gets used in a group has much more to do with the interpersonal meanings brought to the group and developed during the interaction” and Shah, Lewis, and Caires (2014, p. 501) conclude that “it may be more useful to conceptualize equity as contextual.” These studies highlight the need for a re-evaluation of how we study equity in computer science.

The studies cited above as well as others focus on a singular, monolithic definition of equity: Chizhik (2001) looked at how task characteristics (e.g. the number of solutions) affected the distribution of forms of participation and achievement on a test; Sullivan and Wilson (2015) looked at how positioning via “playful talk” affected status within the group, which has been shown to affect opportunities to learn. This is problematic because there are many aspects affecting and constituting equity and “these relationships [between aspects and facets of equity] turn out to be much more complex than one would expect” (Cole, 1995, p.68). So, we ask “what does equity mean in programming pairs?” To address this question we employ the methods presented in Lewis and Shah (2015) and Shah, Lewis, and Caires, (2014) because they specifically address Computer Science and include multiple measures of equity including quantitatively coding turns of talk, questions, and commands. However, we want to further elaborate on what equity means.

In an attempt to better inform how these research methods can account for the multiple aspects that affect and constitute equity, we turn to a well-cited definition for equity, which accounts for the multi-faceted nature of equity. Esmonde (2009, p.249) defines “equity as the fair distribution of opportunities to learn.” She further breaks down this definition into two parts: (1) “opportunities to make sense of mathematical ideas and to participate in mathematical discourse practices” and (2) “students’ opportunities to develop positive positional identities that place them as authoritative and competent members of the classroom community.” We can map our selected methods onto Esmonde’s definition by slightly modifying it. First, we are going to look at actualized engagement as supposed to “opportunities to learn” due to the difficulties associated with measuring an opportunity. For example, if a student is working in a group, and she does not engage in the activity, how can we differentiate whether she had the opportunity and chose not to take it, or whether she never had the opportunity in the first place, perhaps due to factors beyond our view? This complexity makes it very challenging to assess

opportunity in classroom examples of group work. In addition, we divide Esmonde's first branch further into (1a) sense making and (1b) practicing disciplinary discourse to help differentiate these two observable facets of equity she is presenting. The resulting definition is mapped to our quantitative methods below. We first explain how we triangulate that this mapping is theoretically sound by using interaction level positioning to see what these quantitative indicators are capturing.

Cole (1995, p.71) states that "surface indicators alone, as convenient as they may be, cannot tell us at a glance whether a group is working well together." In order to get a better sense of group interactions, we extend Lewis and Shah's work (2015) by using the construct of positioning to examine discourse. Michaels et. al (2007, p.294) states that "all social relationships are in play in the accomplishment of deliberative discourse." Van Langenhove and Harre (1999, p.17) define "the act of positioning" as "the assignment of fluid 'parts' or 'roles' to speakers in the discursive construction of personal stories that make a person's actions intelligible and relatively determinate as social acts." Positioning in discourse could look like supporting an idea or critiquing an idea, which leads to the speaker being positioned as more knowledgeable or not. For example, a passing teacher may position a student towards the project by asking how it works. The student in turn could position him or her self away from the project by replying their partner built it so they do not know. In our study, we avoid presenting data about teacher and student interactions because of the intrinsic power dynamics between students and teachers in a classroom, making the positioning inherently inequitable. In addition to discursive positioning, we incorporate another facet to positioning: embodiment.

We apply the construct of embodied positioning used in math education (Leander, 2002; Dookie, 2014). Engle, Langer-Osuna, and McKinney de Royston (2014) present a framework that includes not only degree of intellectual authority and access to the conversational floor, but also degree of spatial privilege. Leander (2002, p. 193) also argues "that 2 key processes are active: narrating social 'scenes' through talk and producing embodied spaces." Embodiment is relevant to analyzing pair programming because pair programming typically involves a pair of people sharing a computer with a single mouse and keyboard; the person whose hands are on the keyboard and mouse has more power than the person who does not. By accounting for this embodied power via analysis of who has control of the computer at what points in time, we can better understand equity in the situation.

Using our two modifications of Esmonde's (2009) definition, we operationalize equity into three observable and quantifiable equity indicators that can provide an approximate measure of equity: Tool use, talk distribution, and question/command distribution. These indicators form the basis for a coding scheme (described below) that we use to analyze video recordings of students' discourse.

Our first equity indicator, tool use, ties back to the first part of our definition - sense making. Active computer use could indicate engagement with disciplinary ideas. While it is unlikely that computer use is the only way to engage with disciplinary ideas or that all computer use signals engagement in disciplinary ideas, the two are related. "Inequitable access to the computer becomes... a lack of exposure to the curriculum" (Cole, 1995, p.67) and computer use has been used previously as a metric in studies of equity in paired programming (Plonka et al, 2012). Theoretically, in an equitable pair, we would see a near equal distribution (50-50).

Our second equity indicator, talk distribution or "air time", ties back to the first part of our definition – practicing discourse. By measuring the distribution of student talk (Lewis & Shah, 2015; Shah, Lewis, & Caires, 2014) we can see indications of engagement with discourse practices. While it is unlikely that every turn of talk uttered by a student is engagement with discursive practices, we have a couple of reasons to believe that this metric, in our data, will still reflect an approximation of the ratio of discourse practices: (1) the students in our study have the same level of programming experience, so it is reasonable to assume that they engage in discursive practices at approximately the same rate and (2) "verbal matching within a group is a significant indicator of how well members of that group like one another" (Gonzales, Hancock, & Pennebaker, 2009, p.9), meaning that since the students presented below are very friendly towards one another, if one student starts or stops engaging in discursive practices then the other student will likely follow. If we then assume equitable collaboration, "the number of turns would be near equally distributed (i.e., 50-50)" (Shah, Lewis & Caires, 2014, p. 497).

Our third equity indicator, authority, ties back to the second part of our definition – authoritative and competent identities in the classroom. Commands and questions are of particular interest because they can indicate positioning. Questions generally give status to those being asked by positioning them as a relatively knowledgeable since "asking a partner a question can be seen as actively positioning that person as competent and, alternatively, not asking a partner a question can be seen as passively positioning that person as less competent" (Shah, Lewis, & Caires, 2014, p.499). Commands generally give status to the speaker, positioning them as relatively knowledgeable since "issuing many directive statements may indicate a lack of respect for the intellectual capacity of that individual to contribute" (Shah, Lewis, & Caires, 2014, p. 499). We use the word relative to indicate that when a speaker asks someone a question or commands the listener to do something, this positioning is not global, it is a relational shift between the speaker and listener. This position of knowledgeable

is relative to the other person and only exists within the pair at that moment. We use coding to give us a rough idea of equity within the pair since one would assume about a 50-50 distribution of questions to each other to keep positioning equitable.

Now we can address our research question: How do different nominal indicators of equity differentially describe peer interactions? In this paper, we present the results gathered by applying these methods, we then reflect on what each of these methods foregrounds, their strengths and shortcomings, including their sensitivity to time-scope and pedagogy, and triangulate the quantified and qualitative results. We conclude by suggesting ways in which future work could more sensitively theorize about and assess equity in students' collaborative activity.

Methods

The data in this paper were collected in a New England high school's recently established makerspace, where we taught a single-semester elective course in digital making. The high school is highly diverse with 70% minority students and 58% of students coming from low-income families. The school's diversity is reflected in the 21 elective students (8 Haitian-Americans, 6 Hispanic, 2 African-Americans, 2 Asian-Americans and 3 other non-White students). The project captured approximately 260 hours of video, copies of student notebooks, pictures of artifacts, and the staff's daily field notes.

The data in *this* paper (a 35 minute segment of video) were collected when students had already completed one project and were working on their second project. The objective of this project was to make the makerspace "smart," where students were asked to use technology to make the space interactive in similar ways to home automation. Students were using BlockyTalky, a tangible networked tool kit used to build custom devices consisting of motors, sensors and LEGO pieces for structure. BlockyTalky devices can send messages over wireless networks to other units, to computers with music synthesizers and to Android apps created on MIT App Inventor.

In this paper, we present an episode of a pair of two sophomore girls. This pair was selected because from the outside, it was hard to tell whether the pair was inequitable or equitable. One girl did the vast majority of programming but the two girls seemed open to talking to one another. The traits of unequal tool use and openness to talk were widespread throughout the groups, making this pair representative of the groups in the class at different points in time, but not necessarily representative of the pair for the entire duration of the project.

In order to quantitatively assess how equitable a pair is, we use Shah and Lewis's scheme (Shah, Lewis & Claires, 2014; Lewis & Shah 2015) to quantify the students' relative distributions of computer time, talk, commands and questions. Computer time was attributed to the student who was actively using the keyboard or mouse. We then computed the ratio of computer time between the girls to arrive at a measure of equality. We also created transcripts for the selected episode, dividing them into turns, and tagging all turns containing a question or direct command. We counted the number of turns and tags for each girl and created a distribution of talk and talk type.

We also employ a more qualitative method, positioning, to investigate how these indicated inequities are created. To find instances of positioning, we consider each utterance and identify any positioning. In interpreting these utterances we code intentional positioning moves in the data (i.e. "I am going to program"), and tacit positioning (i.e. you are not programming). In the presentation of this coding below, we talk about positioning as authoritative or knowledgeable as positioning *up* and positioning as less authoritative or knowledgeable as *down*. In a few instances we also code positioning *towards* or *away* from the project itself, which is different in that it is positioning one as an owner or doer of the project. We make this distinction because positioning *towards* or *away* from the project is more likely to be in conjunction with embodied positioning. We also take note of embodied positioning by noting for each turn of talk who has control over the computer, paying close attention to transfers of possession and considering how it affects the meaning of the verbal positioning or vice versa.

Results

The pair of girls chose to address the teacher's inability to hear the bell within the woodshop because this problem had been causing students to be late for class. These two students attempted to make this "smart" bell by programming a BlockyTalky device with a sound sensor in the hall to sense the bell ringing, and then send a message to a second BlockyTalky device to have a synthesizer make noise in the room. This episode starts on a Friday afternoon, only a few days into starting the project. Sarah and Alice are sitting at a laptop at a table programming two BlockyTalky devices named Bruce and Eve.

From the quantified perspective, the pair looks inequitable (Figure 1). Computer time was distributed 16% to 84% in Sarah's favor (N = 12:48 minutes). Overall talk distribution was 37% to 63% in Sarah's favor (N=316 turns). Commands were distributed 53% to 47% for Alice and Sarah respectively (N=17 turns). Questions

asked were distributed 58% to 42% for Alice and Sarah respectively (N=50 turns). Our coding of these data is marked using boxes around the transcript in the qualitative findings section below with the following notation: coded as command and coded as question. Tool use, including computer time, is described within the analysis following the transcript.

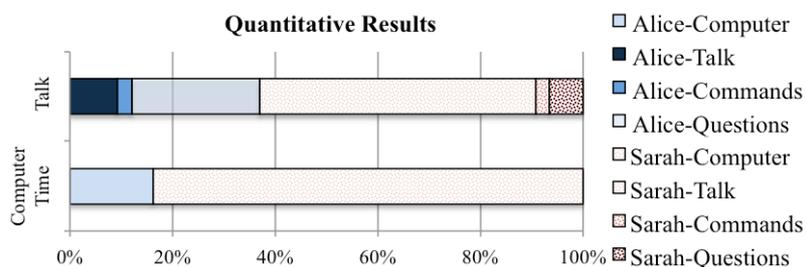


Figure 1: Computer and Talk Distribution by Student

We now have an indication of how equitable the pair was. The pair appears to be inequitable in the measures of computer use and turns of talk, but equitable in the measures of commands given and questions asked. In theory, the dynamics of positioning we see in the qualitative analysis should approximately match the proportions above. It then follows that if we see Sarah making moves to keep tool usage or a lack of Alice making bids to use the tools, a similar yet less extreme dynamic should be seen with air time, and an almost equal distribution of moves to position oneself as knowledgeable – these indicators are good approximations of actual positioning. If not, we would need to question whether these indicators are a reasonable construct to apply in this context.

We start the qualitative analysis about 24 minutes into the episode where we see a transition from talking to active programming on the computer. This piece of the episode was chosen due to the lack of programming in the first 20 minutes and the fact that teacher interactions were minimal in this piece of the episode. While most of our methods would work on student discourse without tool use, our aim is to present all of the indicators in conjunction with the qualitative analysis and we cannot study the effect of positioning on tool use when there is no tool use.

[23:39] S: Wait, stop.

[23:42] A: That was fun

[23:43] S: As you can tell, I didn't program it to do any kind of actual sound. I just did a bunch of random buttons.

[23:56] S: Hold on a second, stop stop.

The transcript above is presented based on time stamps, not turns of talk as defined in the methods section, so when trying to understand this coding, each line needs to be dissected into turns. For example, [23:43] has two turns of talk within the single time stamp.

Our equity indicators in this snippet show Sarah dominating through computer use, more turns of talk and multiple commands. This matches the qualitative analysis where we see Sarah positions Alice away by commanding her to stop pressing the buttons multiple times ([23:39], [23:43]). Sarah explains to Alice that she programmed the device to make random noises when triggered, emphasizing the randomness with a gesture. Looking at embodied positioning, Sarah is on the computer for most of it ([23:39]- [23:43.]). At the beginning,



Figure 2: (a)Left - Alice reaches across table to touch phone
(b) Right – Sarah motions to Alice to stop pressing button on phone

Alice reaches over to touch the phone in front of Sarah (Figure 2a) then Sarah relocates the phone in front of Alice ([23:39]) before making an embodied move in conjunction with her command ([23:56], Figure 2b). By reaching out to prevent Alice from touching the phone it emphasizes Sarah's authority. In isolation, these few turns could be misleading, but as part of the whole episode we see how this temporarily lop-sided interaction leads to skewing the equity indicators of commands.

[24:08] A: Hey, did you want me to make an actual beat?

[24:11] S: If you want to. I was just going to go ahead and [inaudible]

[24:14] A: Alright, well that sucks.

Our equity indicators in this snippet show almost equal distributions, as Sarah stops using the computer, allowing Alice to start using it, there are almost equal turns of talk and only a single question. This almost matches the qualitative analysis where we see Alice volunteers and then starts programming an actual beat ([24:08]) – though in terms of disciplinary content this is arguably a peripheral part of the programming project. Additionally, Sarah positions Alice's contribution as minimal ("If you want to"), positioning Alice down. Sarah starts this snippet on the computer but Alice's question ([24:08]) leads her to move away from the computer and leave the table while Alice gets a chance to program ([24:14]). Shortly after, Alice finishes programming and asks Sarah for help uploading her code to the BlocklyTalky. Notice how drastically different the equitability of the pair is between the two presented snippets – in a matter of seconds we go from Sarah dominated to almost equitable just because of where these turns of talk were grouped.

[26:06] A: How did you upload it?

[26:13] A: This is the old one still

[26:19] S: Upload

[26:28] S: Stop hitting it for a second. That was really weird. Okay. We'll leave it at that. Okay. Now we just need the sound sensor

[26:50] S: Oh you changed the sound

[26:53] S: Please stop because [inaudible] over load the program.

Our equity indicators in this snippet show Sarah dominating the pair again, as Sarah takes over the computer, dominates turns of talk and issues two commands – exacerbated by Alice asking a question. This almost matches the qualitative analysis where we see Sarah returns to the table and to the embodied position of computer user after Alice positions her as knowledgeable by asking a question ([26:06]). Alice presses the start button on the phone ([26:13]) and notes the lack of difference – a meaningful contribution not picked up by the indicators - and Sarah responds by showing Alice the upload button ([26:19]). While this enables Alice to contribute, Sarah continues to use the computer, not returning it to Alice. Sarah then proceeds to once again command Alice to stop pressing the button, emphasized by a less drastic gesture than illustrated in Figure 2b ([26:28]). Alice starts repeatedly pressing the button enthusiastically once her new tune starts coming out of the synthesizer ([26:50]). Sarah issues another command, gesturing by touching the phone ([26:53]). This interaction positions Alice further away from the project. Sarah's comment about Alice's contribution is a little hard to interpret since there is no value statement in it ([26:50]) – she is simply stating that she observed Alice's contribution. It could be argued that she accepts the modification and thus positions Alice up since Sarah does not change the modification.

[27:00] S: We need a sound sensor or a sensor that senses sound, okay

[27:03] A: Sound sensor? What do you mean? Something that picks up sound?

[27:05] S: Yeah. But I think it's already doing that.

[27:09] A: Go on the phone. The phone screen then get the music one and drag it over

[27:17] S: Hmm?

[27:17] A: Go on the phone

[27:19] S: No, this is for Bruce

[27:20] A: Oh

[27:21] S: Then Bruce to that

[27:27] A: Don't we need two phones then?

Our equity indicators in this snippet are hard to interpret by themselves - during this whole snippet, Sarah is using the computer, but neither girl dominates turns or types of talk. Turning to the qualitative analysis where we see Sarah maintains her position as knowledgeable by making a statement and having Alice ask clarifying questions ([27:00] - [27:03]). Sarah clarifies what she means by the sound sensor, resulting in Alice being positioned down. Alice, in return, positions herself up by starting to direct Sarah to program a sound sensor using the App Inventor ([27:09]), making gestures in the air of where to go on the screen. Sarah clarifies that the sound sensor is for the BlockyTalky Bruce, adding emphasis by touching the device with her left hand, not the phone ([27:19]). However, Alice does not seem to understand, positioning Sarah up yet again by asking if they need two phones ([27:27]). Sarah ignores the final question from this snippet so Alice asks her again.

Almost all of the commands and questions position the speaker and listener as theorized with exception of the final command in this snippet, which due to Sarah's rejection of Alice's positioning ([27:19]) results in an overall downward positioning for Alice. This error within the equity indicator construct, however, seems to be the exception, not the rule. When looking at other instances of commands in this episode, we notice this is the only instance of a command resulting in the speaker ultimately being positioned down.

[27:49] A: Don't we need two phones then to transmit it?!

[27:50] S: No. Because if this hears sound, it will set off Eve. If someone looks at the time and realizes the time they hit E-they hit the button and it sends a message to Eve

[28:01] A: Okay

[28:12] S: Sound sensor port 1, yes. But what do we do with the sound sensor?!

Our equity indicators in this snippet are also hard to interpret by themselves, distributed similarly to the previous snippet. Turning to the qualitative analysis where we see Alice reiterates her question ([27:49]) and Sarah replies ([27:50]). Sarah, who is on the computer with the exception of this ([27:50]), not only fully explains the flow of execution so Alice knows how the project works, but stops using the computer to touch each object she references as she walks through it. Alice replies "Okay" but sits back in her chair so she cannot see the computer screen and starts looking towards the back of the room. Unlike the last snippet where this mixture of indicators shook out to be about equitable, we see Alice unengaged and leave Sarah to dominate the project. This is an interesting case of where the indicators – despite having similar quantities – mean different things. Sarah continues by positioning Alice up by starting off the next step of the project saying "we" as supposed to "I" ([28:12]) and asking Alice about what to do with the sound sensor.

In Figure 3, the lines represent the ratio from the start of the episode to that turn of talk. The lines climb as measures are coded as Sarah, and the lines go down as measures are coded for Alice. As these are cumulative sums of the codes, they start to level off to indicate an average ratio. Looking at where these indicators level off the emerging averages even for the subset of the total coded set show the same trends discovered by coding the whole episode.

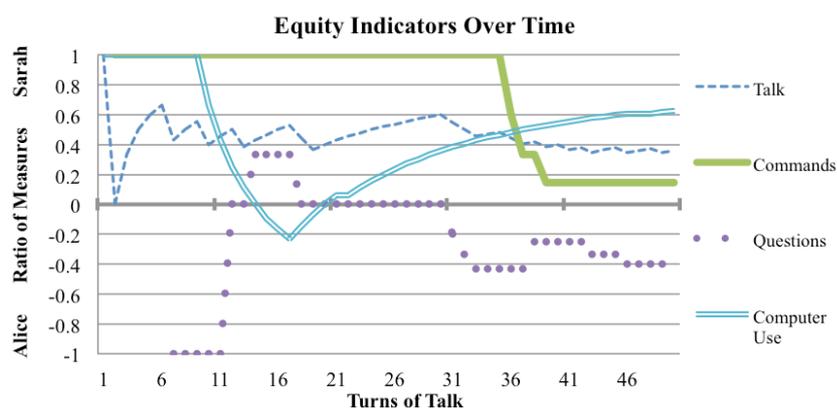


Figure 3: Equity indicators in favor of each student over the presented transcript

Discussion

We discuss the four measures of equity in terms of what each measure lets us see, the strengths and weaknesses of each measure, and agreement between each of the measures.

We investigated how the girls' *tool use* would help us quantify engagement with disciplinary ideas. Overall, we see a disparity between computer use, Sarah using it actively for 84% of the time. While one of this measure's strengths is that it can pick up embodied moves, like Alice and Sarah reaching for/starting to use unoccupied tools with no verbal indications, and that it aligns with the dialogue when present, we do notice that in this learning environment, there are more tools than just the computer. In addition, in learning environments such as in Shah and Lewis's study, where students are in a paradigm where keyboard use is forced to be equal, this measure would be an artificial indicator—meaning that this measure is sensitive to the structure of the activity. Finally, in contexts such as this where there are multiple tools and multiple ways to interact with computer science skills, this measure would struggle to mirror reality without being modified to account for this change. Most notably, every time the girls test the bell, they both can *hear* the test since sounds are being played – making it extremely difficult for one student to try to prevent the other from participating in testing. As researchers experiment with different pedagogical models or technological tools for learning computing, the structures embedded in these models could have large impacts on the quantitative measures used in this study, without necessarily impacting overall equity in a collaboration. Developing ways to triangulate between multiple ways of assessing equity, as we do in this paper, may help to ameliorate over- or under-estimation of equity that single methods might cause.

We investigated how the girls' *distribution of talk* would help us quantify practicing disciplinary discourse. The two girls enacted a dialogue with lots of back and forth in the form of typical conversation turn taking. This dynamic results in about equal turns of talk and could be held as standard and the deviations from this as being where the power shifts occur. While we do see engagement with disciplinary ideas happening discursively, this measure has no way of differentiating between an engagement with a disciplinary idea and a non-disciplinary utterance like “Okay, let's see.”

Despite the weaknesses in both this measure of talk and the measure of computer use, we see that on a larger scale, both of these measures are in agreement that Sarah engaged in more disciplinary knowledge building – both through activity and discourse. Looking at the transcript and video, we would agree with this characterization of the pair. Since both girls seemed to engage in disciplinary content when talking and using the computer at the same rate, these measures reflect researcher perceptions of the data. However, both of these measures are sensitive to differences. For example if someone is a much slower typist, or if the style which someone speaks, for example Sarah, is choppy than Alice, resulting in more but shorter turns of talk, these measures might be artificially inflated.

We claimed *commands issued and questions asked* would help us quantify authoritative and competent identities in the classroom. Overall both of these measures indicated equity of authority. This makes sense given that throughout the episode there were numerous examples of authority being shifted between the two girls. The salient positioning mechanisms that used commands and questions have been mentioned above – commands, asking for tool use, monologues and questions. This also makes sense given that causing inequities in the other measures requires a power shift. In terms of weaknesses, there were a couple of instances where Sarah used “I” or “we” to describe working on the project, which caused power shifts that are not accounted for by the indicator.

We claimed *positioning* would help us triangulate the other coding schemes. To get an idea for a larger scope with this method, we attempt to distill trends in the codes seen throughout the episode to compare them to the quantified indicators. First, a lot of the discourse around the tools, even the phone that was not included in the computer use code, was power moves by Sarah who was actively using the computer or commanding Alice to stop using the phone to test. This trend of inequity in tool use is reflected in the quantified measure. Second, looking at talk distribution, we see that Sarah occupies the conversational floor for the majority of the time, expressing more disciplinary thoughts than Alice. This trend of inequity is reflected in the quantified measure. Third, we see almost constant shifting in authority between Alice and Sarah, without any long periods of time being dominated by one girl's authority. This trend of relative equity is reflected in the quantified measure.

One major weakness of models from current literature, which look at change over time, is their drastic sensitivity to *scope* – or the amount of time the method is applied over. For example, one difference between the positioning analysis presented above and the one presented by Shah and Lewis (2015) is scope is instead of a dozen of turns of talk that are supposed to characterize the pair's interaction as a whole, we present multiple minutes of transcript. There are many ways to slice the data into a small series of turns, however, because of how fluid these measures are, the picture captured in those series looks drastically different depending on where you select them (Figure 3). The cumulative coding over time is rather messy though they start to level out over time. This variability within the measures means that applying these measures to short pieces of data will result in unreliable and not necessarily valid findings. In the same vein, analyses that only show small snippets of the data hide a lot of the complexity and overall trends. This indicates that we need a theory and methods that reflect the time-course of real-world phenomena.

Conclusion

Equity is a complex construct. This research illuminates the need for theories and analytical methods that better engage with these complexities. This analysis reflects results from methods from the literature, and adds support to the claim that the quantifiable metrics serve as a proxy for the different facets of equity we investigated. There are many other forces, including socio-cultural processes that can empower some students and disempower others, that while we did not wade into them, we hope future efforts to put this work in conversation with critical and structural frameworks for equity could further advance the field. We also highlight the pedagogical and technological sensitivities within the methods of the literature, meaning that quantitative methods need to be adapted to the context that they are applied in. Because of the variance in how measures can quantify interactions, researchers must be careful to select equity measures that are appropriate to the contexts that they are studying, as well as to explain their processes for selecting those measures. We argue that the scope of data used needs to not only be carefully considered to avoid selection bias, but justified in these studies, possibly through quantitative analyses, due to the sensitivity of qualitative methods to local maximums and minimums illustrated in selected clips. Lastly, we urge researchers to define and present the type of equity they are prioritizing when we talk about designing and implementing equitable learning environments to help increase clarity in the field. Ultimately we are left with the question: how would someone go about convincingly evaluating equity within a pair?

References

- Chizhik, A. W. (2001). Equity and status in group collaboration: Learning through explanations depends on task characteristics. *Social Psychology of Education*, 5(2), 179-200.
- Cole, K. A. (1995, October). Equity issues in computer-based collaboration: Looking beyond surface indicators. In *The first international conference on Computer support for collaborative learning* (pp. 67-74). L. Erlbaum Associates Inc..
- Dookie, L. (2014) A Case Study Examining the Microdynamics of Social Positioning within the Context of Collaborative Group Work. *Proceedings of the 2014 Annual Conference of the Learning Sciences*. Boulder, CO. Boulder, CO.
- Engle, R. A., Langer-Osuna, J. M., & McKinney de Royston, M. (2014). Toward a model of influence in persuasive discussions: Negotiating quality, authority, privilege, and access within a student-led argument. *Journal of the Learning Sciences*, 23(2), 245-268.
- Esmonde, I. (2009). Mathematics learning in groups: Analyzing equity in two cooperative activity structures. *Journal of the Learning Sciences*, 18 (2), 247-284.
- Gonzales, a. L., Hancock, J. T., & Pennebaker, J. W. (2009). Language Style Matching as a Predictor of Social Dynamics in Small Groups. *Communication Research*, 37(1), 3–19.
- Leander, K. M. (2002). Silencing in classroom interaction: Producing and relating social spaces. *Discourse Processes*, 34(2), 193-235.
- Lewis, C. M., & Shah, N. (2015, July). How Equity and Inequity Can Emerge in Pair Programming. In *Proceedings of the eleventh annual International Conference on International Computing Education Research* (pp. 41-50). ACM.
- Margolis, J., & Fisher, A. (2003). *Unlocking the clubhouse: Women in computing*. MIT press.
- Margolis, J., Estrella, R., Goode, J., Holme, J. J., & Nao, K. (2010). *Stuck in the shallow end: Education, race, and computing*. MIT Press.
- Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in philosophy and education*, 27(4), 283-297.
- Plonka, L., Segal, J., Sharp, H., & van der Linden, J. (2012). Investigating equity of participation in pair programming. In *AGILE India (AGILE INDIA), 2012* (pp. 20-29). IEEE.
- Shah, N., Lewis, C. M., & Caires, R. (2014) Analyzing Equity in Collaborative Learning Situations: A Comparative Case Study in Elementary Computer Science. *Proceedings of the 2014 Annual Conference of the Learning Sciences*. Boulder, CO. Boulder, CO.
- Sullivan, F. R., & Wilson, N. C. (2015). Playful talk: Negotiating opportunities to learn in collaborative groups. *Journal of the Learning Sciences*, 24(1), 5-52.
- Van Langenhove, L., & Harré, R. (1999). *Introducing positioning theory*.

Acknowledgments

We thank the teachers and students who participated in this research. This work was supported by the National Science Foundation (IIS-1450985 and CNS-1418463) and LEGO Education.

The Effect of Concrete Materials on Children's Subsequent Numerical Explanations: Metaphorical Priming

Andrew Manches, University of Edinburgh Moray House School of Education, a.manches@ed.ac.uk
Mihaela Dragomir, University of Edinburgh School of Informatics, m.dragomir@sms.ed.ac.uk

Abstract: This paper contributes evidence for the claim that gestures used to support numerical thinking can simulate prior concrete experiences. 114 children aged 6-9 years explained a numerical relationship (additive composition) three times consecutively. All children explained without materials for the 1st and 3rd explanation. For the 2nd explanation, children were randomly assigned to one of three conditions to use: physical objects; a number line; or no materials (control condition) to explain their thinking. Findings showed how using physical objects significantly influenced the particular types of gestures (e.g. splitting), hand morphology (e.g. pinching), and words (e.g. "take", "big") that children used in subsequent explanations without materials. Similar (although less pronounced) priming effects were found for the number line condition. The study provides support for conceptual metaphor theory (used to categorize gestures and language), and the potential for gesture research to address long-standing questions concerning the role of concrete materials in learning.

Introduction

As well as the importance of informing everyday pedagogy, substantial research has examined if and how children's interaction with physical materials supports learning in order to evaluate new digital forms of interaction and also to better understand the relationship between action and cognition. Gesture research has offered a unique window into this research area by suggesting that some gestures, which are used to explain particular concepts, simulate prior interactions with physical materials. The contribution of this paper is to support this claim by demonstrating how asking children to explain a particular numerical relationship *with* physical materials significantly influences the types of gestures and language children subsequently use to explain their thinking *without* materials. By providing evidence of this internationalization of action experience, the work reported provides empirical support for claims surrounding the embodied nature of cognition.

The role of physical materials in learning

From plastic blocks to a counting abacus, physical learning materials are pervasive in early learning classrooms, and have been designed and advocated by many educational pioneers, from Froebel to Dienes. Such use has attracted a wealth of research attempting to understand how, and if, they support children's learning, particularly in domains such as mathematics. Unfortunately, although a relatively recent meta-analysis suggests a slight positive effect (Carbonneau, Marley, & Selig, 2012), there remains a lack of consensus over their benefits (McNeil & Jarvin, 2007), and research remains limited in its capacity to inform everyday practice.

It is not simply their pedagogical value however that has attracted research into physical learning materials. These objects encapsulate fundamental questions concerning the relationship between action and cognition. In this regard, the last two decades has witnessed increasing interest in this field, owing to two main reasons. The first is that new technologies are changing children's interaction with the world, raising pertinent questions about the subsequent impact on conceptual development. As well as addressing concerns that devices (e.g. mouse/touchscreen) limit accessibility, researchers have investigated the potential for more interactive devices (e.g. tangible technology/gesture recognition devices) to leverage body based learning mechanisms (Manches, O'Malley, & Benford, 2010).

The second reason for renewed interest in physical learning materials concerns theoretical developments in cognitive science over the last twenty years regarding the relationship between body-based experience and cognition. Embodied Cognition is an umbrella term capturing various claims that cognitive processes are best understood when perceived as grounded in our body's interaction with the world. One of the more controversial claims of Embodied Cognition is that our 'offline' thinking (thinking in the absence of relevant stimuli) is body-based (Wilson, 2002).

Whilst the importance of action experiences is a core feature of prominent cognitive developmental theories, Embodied Cognition differs in its proposition that that these experiences are not simply precursors to more abstract thinking, but are encoded as an integral part of developing concepts in the form of sensor-motoric representations. One interpretation of this proposal is to emphasize the importance of particular action experiences in learning, thereby providing a novel rationale for the benefits of physical materials (Pouw, van Gog, & Paas,

2014), as well as new forms of digital designs that can elicit and dynamically represent particular physical actions (Abrahamson, 2009).

Action experiences and conceptual development

An important question raised by embodiment theories therefore concerns which types of actions might be more beneficial for developing particular concepts. Should we encourage children to move their hands together when learning the concept of addition for example? After all, most forms of interaction in learning tasks, from pens to computer mice, generate some form of action. This has led researchers to refer to the extent to which particular actions are ‘congruent’ with the concept at hand (Segal, 2011). According to Bakker, Antle and Van Den Hoven (2012), the relationship between particular actions and concepts depends upon the embodied metaphors underlying the concept. Metaphor in this context refers to the mapping of bodily originating schemata onto a conceptual domain (Lakoff & Johnson, 1980) (e.g. mapping body experience of balance with the ‘abstract’ concept of justice). Bakker et al. suggest that such metaphors might be identified from existing theoretical literature or investigated through empirical work examining the way individuals move their bodies when articulating their thinking about particular concepts.

Gesture research

The most prominent way we move our bodies when thinking is through our hands, and in support of embodied claims, research has demonstrated that the main function of gesture is to support the speaker’s thinking rather than just the listener’s comprehension (Tellier, 2009). The last two decades has seen increasing work investigating the way children and adults gesture when explaining their mathematical thinking (e.g. Alibali & Nathan, 2011; Goldin-Meadow, Kim, & Singer, 1999). Edwards (2009) for example, examined student teachers’ explanations of fraction concepts. Of the 251 gestures generated by the 12 participants, 32% referred to fractions. Of these 81 were categorized as primarily metaphoric, where the pictorial content presents an abstract idea rather than a concrete object or event, and 35% as primarily iconic, that “bear a close formal relationship to the semantic content of speech” (McNeill, 1992, p.14), interpreted here as visually resembling their concrete referents.

Whilst many researchers view the categories of iconic and metaphoric gestures as both constituting a broader category of *representational* gestures (where the hands depict aspects of meaning in speech literally or metaphorically - Alibali & Nathan, 2011), Edwards’ work provides great insight into the types of concrete experiences that may have influenced participants’ thinking. From her data, Edwards categorized two types of iconic gestures: *iconic-concrete*, that refer to concrete objects or processes “often related to tangible materials utilized in early instruction about fractions” (Figure 1), and *iconic-symbolic*, that refer to gestures that “re-enacted the physical process of writing out a mathematical procedure, or referred to visual locations and elements of mathematical symbols”. These categories highlight the potential to examine gestures in order to understand how prior perceptual and action experiences may have shaped thinking. What is not clear is whether gestures simulate *particular* experiences with materials, or whether there exists any underlying embodied metaphors for the range of gestures generated.



Figure 1. Iconic-concrete gestures (Edwards, 2009).

Embodied metaphors of number

In their seminal book, “*Where mathematics comes from*”, Núñez and Lakoff (2000) propose that there are two fundamental body-based metaphors underpinning concepts of number. The first is *Arithmetic is object collection (OC)* which is a mapping from experiences with physical objects to the domain of numbers. The second is *Arithmetic is motion along a path (MP)*, which is a mapping from experience moving along point locations in a linear direction to the domain of numbers (e.g. steps when walking). In the OC metaphor, numbers are “collections of objects of the same size” and in the MP metaphor numbers are “point locations on a path”. According to the authors, these metaphors draw upon early body experiences such as manipulating groups of objects or walking in discrete steps.

It is possible to re-examine the gesture categories identified in Edwards' research in light of the conceptual metaphors proposed by Núñez and Lakoff. As suggested in Edwards' definition of iconic-concrete gestures, many gestures seemed to portray simulated action with physical objects such as fraction pieces. Such gestures therefore appear relevant to the Object Collection metaphor – individuals conceptualizing numbers as if they were physical collections of objects. It is possible, although more difficult, to relate other iconic-concrete gestures described by Edwards to Lakoff and Núñez' MP metaphor. For example, describing fractions as a line being split at different points. Perhaps more difficult is to relate either conceptual metaphor to those categorized by Edwards as iconic-symbolic, where participants' gestures simulated interaction with mathematical procedures or symbols. Whilst it remains possible that individuals' concepts of these procedures or symbols pertain to particular metaphors (e.g. "+" symbol being conceptualized as the bringing together of object collections, or movement along a pathway), this cannot be as readily deduced from these types of gestures.

Other authors have proposed different categorizations for gestures explaining different number concepts (e.g. Arzarello, Robutti, & Bazzini, 2005). In this work, it is often possible to relate different gestures to the OC and MP metaphors, although perhaps more difficult for complex concepts which involve increasing levels of symbolization. This would suggest therefore that it might be easier to examine the role of the OC and MP conceptual metaphors in the development of more basic number concepts.

Examining metaphorical gestures in early number concepts

Whilst the notion that all number concepts can be reduced to underlying conceptual metaphors has been called to question (see Kövecses, 2008), the proposal can be scrutinized empirically. More recently, Marghetis (in Núñez, Marghetis, Cohen-Kadosh, & Dowker, 2014) examined the language and gestures of adults asked to explain the numerical concept of 'odd and even' – a relatively early number concept. In this study, adults were first primed according to the two metaphors. In the 'path' condition, participants were asked to imagine a bead moving along a string; in the 'object' condition, participants were asked to imagine combining different collections. In subsequent explanations of odd and even, participants deployed two main types of gestures: those simulating manipulation of objects (hand moving inward shaped as if grasping, pinching or holding), and those simulating movement along a line (pointing handshape, tracing motion along a horizontal axis). Moreover, as predicted, the priming mental imagery was reported to have had a significant effect on conceptualization.

Marghetis' metaphorical priming study design demonstrates the potential for investigating the role of the OC and MP metaphors in early numerical concepts, which therefore provides a window into the types of interaction experiences that might influence the development of these concepts. However, the reported study focused on a single concept and with adults, where the influence of priming may be quite different to children who are still developing their numerical thinking. More importantly, detail is missing about the range and prevalence of gestures generated that is required to evaluate the extent to which individuals' numerical thinking can be captured by the two embodied metaphors proposed by Núñez and Lakoff. The study reported in this paper was designed to address this important gap in the literature.

Study aims

The overarching aim of the authors' work is to understand the role of physical interaction in early numerical development; this paper addresses this aim by examining the role of the OC and MP metaphors in young children's numerical concepts. The paper reports a study that examined the effect of metaphorical priming on children's explanations of a specific numerical relationship: additive composition. Analysis initially focused on the effect on children's gestures during their explanations; however, this was then extended to examine the effect on metaphorical language. Although Núñez et al. (2014) refer to words that pertain to the OC and MP metaphors (see section 2.5.3), the summary of Marghetis' study does not indicate whether changes in the use of particular words were analyzed. We are not aware of existing work that has quantitatively analyzed changes in metaphorical language in this domain (although notable work has detailed more microgenetic changes, e.g. Roth, 2002).

The study design therefore echoes that reported in Núñez et al. (2014), with several key differences. Firstly, the study was carried out with children (aged 6-9 years old) not adults. Secondly a larger sample was involved (n=114). Thirdly, the study captured participants' pre-priming explanations, therefore providing more robust analysis of the effect of priming. Fourthly, this paper reports greater detail on how particular gestures and language were categorized according to the two metaphors. Fifthly, this study reports the effect on metaphorical language. Finally, the materials used for metaphorical priming were familiar classroom materials, thereby providing greater ecological validity in terms of educational implications.

The main research question for the study reported was: *What is the effect of metaphorical priming on young children's numerical explanations?*

Method

The study adopted an embodied cognitive theoretical paradigm, whereby cognitive processes are considered as grounded in the body's interaction with the world. The study methodology employed metaphorical analysis of concept explanations, interpreting gestures (and accompanying speech) in terms of metaphors within the concept domain.

Numerical concept explanation task: Additive composition of number

Additive composition is the concept that a number is composed of other smaller numbers, for example, that 8 can be decomposed into 2 and 6, or 1 and 7. Additive composition is core to numerical development (Martins-Mourao & Cowan, 1998), an overarching concept that connects numerous topics and applications, and is thought to form a conceptual base for the development of children's elementary arithmetic as well as their understanding of the decade numeration system (i.e. understanding that 12 is comprised of 10 and 2) (Nunes et al., 2007). This concept is also evident in the use of flexible addition strategies, for example, decomposing $6+7$ into $6+6+1$ in order to exploit knowledge of double facts (i.e. $6+6+12$).

As young children cannot be asked to simply explain their understanding of the concept term, a question was designed to elicit children's understanding of additive composition. The question draws upon the authors' prior work (Manches & O'Malley, 2016) that required children to explain the logical relationship between two addition parts. The problems were presented verbally (to minimize the effect of external stimuli) but could be represented symbolically as: $a + b = (a+1) + (b-1)$, e.g. $6 + 1 = 2 + 5$.

Design

Each child was asked to explain three additive composition questions consecutively, hence referred to as the 1st, 2nd, 3rd explanation, in a single session. For the 1st and 3rd explanation, children were given no relevant stimuli (e.g. materials or displays). For the 2nd explanation, children were randomly assigned (alphabetically) to one of three independent conditions: *No Materials* (control), *Object Collection* (OC), and *Motion along Pathway* (MP). The No Materials condition acted as a control, asking children to explain without materials (similarly to 1st and 3rd explanation for all children). In the OC condition children were given a collection of small yellow plastic Unifix® blocks. In the MP condition, children were given an 'empty number line'. The empty number line is a common numerical representation (Murphy, 2011) consisting of a horizontal line on paper marked with small vertical lines along its length. Children in this study were familiar with materials in both the OC and MP conditions, reflecting the attempts to ensure ecological validity in the study.

Participants

The study included 114 children aged from 6-9 years old (77-112 months), with parent and child consent. Three children were not included in the data as they were unable to answer the pre-explanation number questions, leaving 111 participants ($M=94.75$ mths; $SD: 10.05$). Children were from 9 classes across three year-groups within a non-denominational primary school in Edinburgh serving children from age 3-12 years.

Procedure

Children were interviewed in 1:1 sessions with the first author who is also a qualified Primary teacher in a quiet but visible location in the school. Each child was given three explanation questions, unless they were unable to answer the initial addition questions. Three children were not able (they were given other questions they could succeed at), leaving a total sample of 111 children.

The format for the 1st, 2nd and 3rd explanation question was the same; they only differed by the sum total: 7, 8 and 9 always in that order. Whilst learning effects were expected, the study focused on between condition differences. An interaction effect between materials and addition total amount was not expected. The format for each explanation was as follows. First, children were asked to count out to the total (e.g. "Can you count to n ?"), where n was 7, 8 then 9 for consecutive explanations. They were then asked the total of the addition sum of $1 + (n-1)$ e.g. "What is 1 add 6?". Children were subsequently asked the total of the addition sum $2 + (n-2)$ (e.g. "What is 2 add 5?"). Finally, for the explanation question, the interviewer reminded children of these two addition problems (e.g. "So, 1 add 6 makes 7, and 2 add 5 makes 7, they both make 7. Can you explain to me *why* 1 add 6 makes the same as 2 add 5?"). The presentation of this explanation task was developed from pilot work and was designed explicitly to draw out children's thinking about the additive relationship (rather than re-state that the sums shared the same result). The choice of language was developed from advice from the class teacher; however, the methodological issue of possible conceptual priming in the question presentation is discussed later.

The 2nd explanation total was always 8. For children in the No Materials condition, the question presentation was the same as the 1st and 3rd. For children in the materials (OC and MP) conditions, the interviewer

echoed the No Materials format, but adapted for the materials. In the OC condition, children were presented with 8 blocks and told to count how many there were. All children were able. The interviewer would then partition the materials into two spatial groups (1 and 7, then 2 and 6) whilst asking the corresponding addition problems. Blocks were recollected after each question to avoid modelling the actions for the explanation task. Finally, the interviewer would ask children to use the blocks to explain the numerical relationship. In the MP condition, the interviewer provided an A4 piece of paper (portrait) with 4 empty number lines in rows. Each would have 9 point-locations (i.e. number line to 8 as first point represents zero). Children were first given a pencil and asked to mark 8 on the number line (the last point on the right of line). If children made the error of counting the first point location (a common and interesting error in light of metaphors), they were corrected (several children did). The interviewer would then use the materials to represent amounts (drawing two arcs on the number line, e.g. from 0 to 1st point then 1st to 8th point) whilst asking the addition problems. Finally, for the explanation question, children were asked to use the number line and pencil to explain the numerical relationship.

There are clear challenges of balancing the presentation and usability of two different representations; however, the following points should be highlighted. First, all children were familiar with materials. Secondly, all children had been shown how to use the materials by the time of the explanation question. Thirdly, the study was not looking at performance with materials, but rather how materials primed particular forms of visual and motor imagery that would be evident in their gestures and language in a subsequent explanation without materials.

Analysis

Out of 111 children, 97 children provided a verbal explanation of some form. Of these children, 59 (61.5%) used at least one *representational* gesture (iconic/metaphoric). The language and gestures employed by the young participants varied substantially, where this paper focuses on three forms of coded data that is proposed to capture children's metaphorical thinking. These are illustrated in Figure 2 and discussed subsequently.

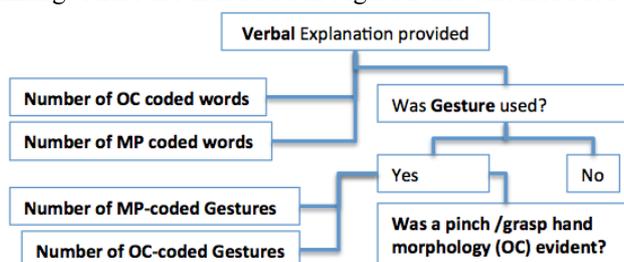


Figure 2: Metaphoric coding.

Metaphorical gesture: OC and MP

The 59 children who gestured generated a wide range of iconic gestures. Many gestures echoed Edwards' iconic-concrete categorization. Within these, it was possible to distinguish between those that seemed to simulate manipulation of objects. These were coded as OC gestures and included simulated splitting a collection of objects with two hands (Figure 3a) or collecting objects (Figure 3b) or others such as sweeping with a curved arm toward the body. The most prominent gesture, simulating collecting objects with two hands (Figure 3b), reflects that described by Núñez et al. (2014) for the OC metaphor. Our previous work (Manches & Dragomir, 2015) also reports a number of gestures appearing to simulate interaction with an imaginary linear pathway. Gestures including pointing up or down (Figure 3c), tracing multiple arcs (Figure 3d), or gesturing movement toward a space to the far right of the child. We coded these as MP gestures. The pointing up or down gesture seems to relate most closely to Núñez et al.'s (2014) description for the MP metaphor. Each explanation received a score for the number of OC and MP gestures used.



Figure 3a-d: OC and MP coded gestures: Splitting, Collecting, Pointing upwards, Arc Tracing.

Metaphorical hand morphology: OC

As illustrated by the work of both Edwards (2005) and Núñez et al. (2014), many children pinched or grasped when enacting an OC gesture; although some did not (e.g. flat hands). Moreover, many children created this hand morphology when not generating an identifiable OC gesture, for example, as part of a simpler deictic (pointing) gesture. It is possible that many children used a pinch or grasp to simulate manipulation of symbols, or materials such as a pencil. However, closer inspection of their language, coupled with the younger age group, suggested that this form of hand morphology simulated grasping physical objects. Therefore, children's explanations were coded according to the presence of at least one pinch or grasp during explanation.

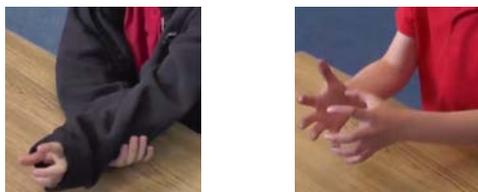


Figure 4a-b: Pinch and Grab hand morphology.

Metaphorical words: OC and MP

Whilst metaphorical language typically refers to collections of words, Núñez et al. (2014) refer to how particular words tend to suggest the OC metaphor (e.g. “big”, “take”) or MP metaphor (e.g. “up”, “far”). In this paper, we report an initial examination of the frequency of words used in explanations that suggested the OC or MP metaphor.

The following procedure was used to code words used in explanations. Firstly, all videoed explanations were transcribed. The authors then coded a list of unique words independently on a five-point scale: *Confident OC*, *Possible OC*, *Neither/Either*, *Possible MP*, and *Confident MP*. A word was then coded as OC or MP if *both* authors independently rated the word as ‘Confident’ or ‘Possible’ OC or MP. In summary, from an initial list of 247 words (having removed non-words, e.g. fillers such as *uhm*), a list of 37 OC metaphor and 46 MP metaphor words were identified. This list corroborated and significantly extended words suggested in Núñez et al. (2014). Whilst there are clear limitations to the de-contextualized process of coding individual words, the focus of analysis for this paper was on the *change* in coded words between explanations according to condition.

Findings

The distribution of group data for three dependent measures (metaphorical gestures, hand morphology, language) was tested (Kolmogorov-Smirnov) and revealed significant departures from normality, therefore non-parametric analyses are reported. Analyses were also carried out on differences in measures between conditions for the 1st explanation. As expected from the random allocation of children to conditions, there were no significant differences found between groups for the 1st explanation.

Metaphorical gestures

In total, there were 13 OC gestures used by 13 children in the 1st explanation, and 49 used by 18 children in the 3rd explanation. A non-parametric between subjects analysis (Kruskal-Wallis) was carried and revealed a significant difference in the change of OC gestures between conditions ($\chi^2(2)=12.263$, $p=.002$). Mann-Whitney tests showed that a positive increase in OC gestures was significantly greater in the OC condition and the No Materials condition than the MP condition ($U=455$, $p=.002$, $r=.36$; $U=544$, $p=.011$, $r=.30$). There were no differences found between the OC and Control condition ($U=605$, $p=.159$, $r=.16$).

For MP gestures, 21 children used 24 gestures in the 1st explanation, and 17 children used 35 in the 3rd explanation. A non-parametric between subjects analysis (Kruskal-Wallis) was carried and revealed no significant differences in the change of MP gestures between conditions ($\chi^2(2)=3.166$, $p=.205$).

Pinch grasp morphology

Of the 59 children who gestured, just under half (25) used a pinch or grasp hand morphology in the 1st explanation, while 24 children did in the 3rd explanation. Each child was given a score of -1 if they used at least one pinch/grab in the 1st explanation but not in the 3rd, a score of 1 if they used at least one pinch/grab in the 3rd but not in the 1st explanation and a score of 0 if there was no change. A non-parametric between subjects analysis (Kruskal-Wallis) was carried and revealed a significant difference in the change of children using pinch/grab between conditions ($\chi^2(2)=7.17$, $p=.028$). Mann-Whitney showed that a positive increase in pinch/grab was significantly greater in the OC condition than the MP condition ($U=507$, $p=.012$, $r=.29$). Although the change in in pinch/grab was greater

in the OC condition than the No Materials condition, this difference was not significant ($U=569$, $p=0.057$, $r=.22$). There was no difference between No Materials and MP conditions ($U=650$, $p=0.446$, $r=.09$). Therefore, there was a tentative priming effect: children who used physical blocks in the 2nd explanation increased the number of pinch/hand morphology in their explanations afterwards significantly more than children who used the number line, and more than the control although not significantly.

Metaphorical words

Each child was given a score for OC and MP word change by subtracting the number of metaphorical words used in Explanation 1 from number of words used in Explanation 3. A non-parametric between subjects analysis (Kruskal-Wallis) was carried and revealed a significant difference in the change of OC words between conditions ($\chi^2(2)=7.43$, $p=.024$). Mann-Whitney showed that positive increase in OC words was significantly greater in the OC condition than both the No Materials ($U=482$, $p=.015$, $r=.28$) and MP ($U=485$, $p=.024$, $r=.26$) conditions. There was no difference between No Materials and MP conditions ($U=679$, $p=.79$, $r=.03$). In other words, there was a priming effect: children who used physical blocks in the 2nd explanation then increased the number of OC words they used without materials in the 3rd explanation.

A non-parametric between subjects analysis (Kruskal-Wallis) was carried and revealed a significant difference in the change of MP words between conditions ($\chi^2(2)=6.73$, $p=.035$). Mann-Whitney showed that a positive increase in MP words was significantly greater in the MP condition than the OC condition ($U=473$, $p=.012$, $r=.29$). Although the change in MP words was greater in the MP condition than the No Materials condition, this difference was not significant ($U=559$, $p=.081$, $r=.20$). There was no difference between No Materials and OC conditions ($U=630$, $p=.391$, $r=.10$). Therefore, there was a tentative priming effect: children who used the number line in the 2nd explanation increased the number of MP words they used relative to children who had used blocks, but not significantly more than children who used no materials.

Discussion

Much research has attempted to understand and evaluate the role of concrete materials in children's thinking and learning. In this regard, Embodied Cognition has offered a valuable framework to consider how children's sensory and motor experiences may shape cognition, and gesture research has provided a methodological tool with which to examine how prior concrete experiences may be simulated when articulating thinking. Although previous work has demonstrated the transition from action experiences to gesture (Roth, 2002), there is limited empirical work illustrating how particular action experiences significantly affect the types of words and gestures children produce when subsequently explaining their numerical thinking. The study reported in this paper addressed this gap by building upon the priming study design used by Marghetis (as cited in Núñez et al., 2014).

The current study examined the effect of action experiences on the language and gestures of young children explaining the numerical concept of additive composition. By drawing upon Núñez and Lakoff's (2000) conceptual metaphor theory, it was possible to codify and quantify three measures of metaphorical thinking: children's gestures, hand morphology, and spoken words. By randomly allocating children to one of three conditions, the study examined how particular experiences shaped children's subsequent thinking. As predicted, the use of physical objects in the second explanation significantly influenced (although effect sizes were generally small) the words and gestures children subsequently used in comparison to using a number line, and often more than the control condition. The finding that the effect was stronger in comparison to the number line condition than control is in itself interesting as it suggests that the actions in the number line group may have interfered with particular metaphorical thinking.

Limitations

There are a number of limitations in the study reported, many of which are discussed. These include the specificity of the task, context, materials, as well as the potential for interpretation in the methods used to quantify children's gestures and words. It is also important to emphasize the immediacy of the post-materials (3rd) explanation, after which the priming effect may soon have been lost. One interesting methodological issue was the format of the problem question. Although the study had been refined in various ways from pilot work (e.g. decision not to provide a the numerical problem visually, ensuring the interviewer did not gesture), the words used in the problem question ("add", "make") are arguably metaphoric (OC). This might explain the salience of the OC priming effect. However, this seems unlikely to fully explain the findings, particularly the effects found for the number line condition on MP metaphorical thinking. It is interesting to reflect on how these methodological issues have been addressed in other work, in particular the possibility of using metaphorical language or stimuli (e.g. number symbols are presented using particular spatial conventions).

Conclusion

The study reported in this paper indicates the potential to analyze the language and gestures children use to investigate how concrete experiences have influenced their thinking. The study contributes to previous work in providing support for the embodied nature of cognition, and the value of gesture research in examining the internalization of action experiences. The study is limited in evaluating whether such experiences have supported learning, yet provides a window into how this approach may address such a goal, thereby informing classroom practice or the design of new materials.

References

- Abrahamson, D. (2009). Embodied design: constructing means for constructing meaning. *Educational studies in mathematics*, 70(1), 27-47.
- Alibali, M. W., & Nathan, M. J. (2011). Embodiment in mathematics teaching and learning: evidence from learners' and teachers' gestures. *Journal of the learning sciences*, 21(2), 247-286.
- Arzarello, F., Robutti, O., & Bazzini, L. (2005). Acting Is Learning: Focus on the Construction of Mathematical Concepts. *Cambridge Journal of Education*, 35(1), 55-67.
- Bakker, S., Antle, A. N., & Van Den Hoven, E. (2012). Embodied metaphors in tangible interaction design. *Personal and Ubiquitous Computing*, 16(4), 433-449.
- Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2012). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, 105(2), 380-400.
- Edwards, L. D. (2009). Gestures and conceptual integration in mathematical talk. *Educational studies in mathematics*.
- Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology*, 91(4), 720.
- Kövecses, Z. (2008). Conceptual metaphor theory Some criticisms and alternative proposals. *Annual review of cognitive linguistics*, 6(1), 168-184.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by* (Vol. 111): Chicago London.
- Manches, A., & Dragomir, M. (2015). *Gesture as a means to examine the role of physical interaction in early numerical development*. Paper presented at the 2015 annual meeting of the AERA, Chicago, US.
- Manches, A., & O'Malley, C. (2016). The Effects of Physical Manipulatives on Children's Numerical Strategies. *Cognition and Instruction*, 34(1), 27-50.
- Martins-Mourao, A., & Cowan, R. (1998). The Emergence of Additive Composition of Number. *Educational Psychology*, 18(4), 377-389.
- McNeil, N. M., & Jarvin, L. (2007). When theories don't add up: disentangling the manipulatives debate. *Theory into Practice*, 46(4), 309-316.
- McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.
- Murphy, C. (2011). Comparing the use of the empty number line in England and the Netherlands. *British Educational Research Journal*, 37(1), 147-161.
- Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology*, 25, 147-166.
- Núñez, R., & Lakoff, G. (2000). *Where mathematics comes from*. New York: Basic books.
- Núñez, R., Marghetis, T., Cohen-Kadosh, R., & Dowker, K. (2014). Cognitive Linguistics and the Concept (s) of Number: Oxford Handbook of Numerical Cognition. Oxford University Press.
- Pouw, W. T., van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review*, 26(1), 51-72.
- Roth, W.-M. (2002). From action to discourse: The bridging function of gestures. *Cognitive Systems Research*, 3(3), 535-554.
- Segal, A. (2011). *Do gestural interfaces promote thinking? Embodied interaction: Congruent gestures and direct touch promote performance in math*. (Doctoral Thesis), Columbia University.
- Tellier, M. (2009). The development of gesture. In K. de Bot & R. Schrauf (Eds.), *Language development over the lifespan* (pp. 191-216): Routledge.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625-636.

Acknowledgments

We would like to thank the Economic Social Research Council for funding this work (ES/K001531/1), the reviewers, and all the children who took part and made this research so rewarding.

Expanding Coordination Class Theory to Capture Conceptual Learning in a Classroom Environment

Lauren A. Barth-Cohen, University of Miami, l.barthcohen@miami.edu
Michael C. Wittmann, University of Maine, mwittmann@maine.edu

Abstract: This article presents an extension to coordination class theory—a theory of conceptual change that was built to capture an individual’s learning in an interview setting. Here we extend that theory to capture group and individual learning in classrooms. The proposed extension focuses on different contexts in the sense of groups’ and individuals’ different interpretations of the same student-generated artifact. A classroom of 9th grade earth science students created embodied models for a specific scientific concept, the steady state energy of the earth. We describe that the students encountered difficulties aligning their embodied models with their conceptual understandings, and yet, they were able to make progress by changing their models to better align their understanding of the scientific concept with their newly modified model—instances of individual and group learning. We conclude with discussing implications for designing classroom learning environments.

Keywords: conceptual change, group learning, embodied modeling

Introduction

This article extends a theory of learning that was built to capture individual’s learning in an interview setting to capture both groups and individuals learning in a classroom. The theory of learning known as coordination class theory (diSessa & Sherin, 1998), describes the organization of knowledge in a well-developed understanding of a concept, the learning difficulties that may arise, and the process of acquiring a well-developed understanding. Our proposed extension describes how a group of students who encountered difficulties were able to make progress at both the individual and the group level. The students generated and modified a series of embodied models for the steady state energy of the earth, they largely modeled energy leaving the sun, entering the earth, and dissipating to space. First we describe the theory and our extensions and then we present several illustrative examples. We conclude with discussing implications for designing classroom-learning environments.

Coordination class theory

Coordination class is a theory of conceptual change that describes the learning of often difficult concepts in science. A coordination class is a specific type of concept that is meant to provide a means for determining information from the world. Knowing that the world is not transparent, learners need to learn how to access relevant information (diSessa, 2002). Different contexts require different means for determining information, and this theory describes how individuals determine information across contexts.

Canonical coordination class theory studies focus on the process of acquiring a well-developed understanding of a concept and the learning difficulties that arise when working to determine the same information in different contexts. For example, diSessa and Sherin (1998) describe how a university student, J, had difficulties determining what is and is not a force in the context of a book sliding across a table and a coin toss. In another empirical study, Parnafes (2007) examined the development of students’ understanding of the motion of simple harmonic oscillators through interactions with physical and simulated oscillators. With the physical oscillator, the students’ intuitive coordination of “fastness” was different from a scientific coordination of velocity and frequency. Then, with a simulation the students were able to distinguish between velocity and period as they determined the same information and refined their knowledge. Thereby the students’ coordination classes were extended to apply to additional relevant contexts of oscillatory motion.

Theoretical machinery

The central function of a coordination class is to determine information from the world, specific classes of information out a variety of contexts in which a specific concept is applicable, for instance determining forces in homework problems, laboratory experiments, and everyday situations. The challenge of “having” the concept is being able to “see” the pertinent information in an appropriate range of contexts, such as “seeing” information about acceleration in physical objects, equations, and with motion detectors in laboratory settings. Learning difficulties may involve having insufficient conceptual knowledge to “see” force or acceleration within a context or having trouble applying the concept across contexts. The organization and structure of a coordination class,

learning processes, and difficulties encountered can be broken down according to the following machinery. (See diSessa & Wagner (2005) and Levrini & diSessa (2008) for additional information.)

Architecture: Rather than directly exacting the pertinent information, people *extract* related information (e.g. shape, color, size, or speed), and then generate an *inference* to turn that information into the pertinent information, such as about acceleration or forces. The latter is known as the causal or inferential net, and it is all of the relevant knowledge used to determine how the information extracted relates to the desired information.

Learning: There are two processes for how prior knowledge contributes to learning: 1) Elements from prior conceptualizations can be *incorporated* into new conceptualization. 2) Elements from a prior conceptualization can be dismissed or *displaced* from a new conceptualization. These processes transpire as one applies or “works” a concept across a range of contexts while incorporating and displacing a variety of knowledge elements.

Difficulties: Given that a certain amount of context-specific knowledge is needed to properly work a concept across contexts, there are two difficulties: 1) The *difficulty of span* is not having adequate conceptual knowledge to apply the concept across contexts, for instance, not being able to access sufficient conceptual elements about forces on static objects. 2) The *difficulty of alignment* is having trouble applying the same concept reliability across contexts, for instance, seeing forces on only moving objects and not static objects.

Contextuality is important; different contexts are where work is done to a concept to increase span and improve alignment while also incorporating new knowledge elements and dismissing old ones. When a student applies a concepts to a variety of contexts, that students may encounter difficulties because they do not have sufficient knowledge to apply the concept (lack of span) or they may encounter difficulties because in some contexts they cannot reliability apply the concept (alignment problem). Either way, incorporating new knowledge elements or dismissing old ones can increase span or improve alignment. Through this machinery we can track learning at a moment-by-moment level.

Existing modifications and extensions to coordination class theory

Several empirical studies of learning have used this machinery while also extending it. Initially the machinery was constructed to describe an individual’s learning of Newtonian Mechanics concepts in a clinical interview setting (diSessa & Sherin, 1998). More recently, Wagner and diSessa (diSessa & Wagner, 2005; Wagner, 2006) presented an explanation of knowledge transfer based on an incremental refinement of knowledge elements across contexts. Over eight weekly interviews one student came to see various probability problems that she had initially perceived as different as being alike. For her, operating the same concept across different contexts was particularly challenging while determining the same information from different contexts was not challenging. Additional empirical studies have applied this machinery to learning in interview settings that involve other math and science concepts (e.g. Thaden-Kock, et al., 2005; Wittmann, 2002).

Using the same theoretical machinery, Levrini and diSessa (2008) conducted an empirical analysis of a classroom of 18-19 year old students who were learning about the concept of proper time within special relativity. Initially, the students had trouble with proper time, but they were able to make progress and the analysis highlights how students “coordinate” proper time. The students incorporated a more careful definition of proper time in a new context, but they encountered difficulties, prematurely concluding that a new context was the same as a prior context, when it was in fact different, which led to a misalignment. Eventually they were able to make progress when the teacher introduced a new definition of proper time that increased span (through increased conceptual knowledge) with an associated cost of misalignment (reduced ability to apply the concept across contexts). Finally the students achieve alignment through an additional context that harnessed the new knowledge. In the analysis, the entire classroom of students was treated *as if* they all were one individual and the analysis documents classroom-learning difficulties instead of individual learning difficulties. Yet, that work did not distinguish between individual and group learning, which is important because individuals may have different interpretations of the same context, and an analysis of classroom learning would want to take that into account as some individuals might experiencing a lack of span or alignment problem while others do not.

Proposed extension to capture alignment difficulties in a classroom environment

Contextuality is important. In prior coordination class theory analyses, often a student has been presented with a variety of contexts where the mathematical or scientific concept applies, for instance forces on moving and stationary objects or in homework exercises and laboratory activities. Yet, multiple individuals observing the same concept, in a context that is supposedly the same, may not, in fact, view the concept in the same way. Their interpretations of a particular concept in a particular context could be different, and articulating those differences of interpretation can be moments of learning. Building from this rough notion, we outline several kinds of contexts

in a classroom-learning environment and then describe the individual and group learning difficulties that may occur.

The data to be presented comes from several team-taught 9th grade earth science classrooms. In the first step of the activity, each classroom of students worked together to collectively build an embodied model for the steady state energy of the earth using a pedagogical activity known as Energy Theater (ET) (Scherr et al., 2012). Next, each classroom presented their model to their peers in the adjacent classroom and there was an opportunity to ask questions. Finally, the two classrooms worked together to build a joint model. Given this classroom-learning environment, we distinguish the following three notions of context. First, there is the *scientific context*, similar to the notion of context used in the aforementioned literature (e.g. Levrini & diSessa, 2008), this notion of content refers to the different scientific phenomena in the external world where the concept is applicable. In our data, the scientific context is the steady state energy of the earth. Second, there is the *classroom context*, which is an entire classes interpretation of an ET model. In our data there were three classroom contexts. The two classrooms each separately built their own models (two classroom contexts), and then, after presenting their model to peers, they build a joint model (the third classroom context). Finally, there is the *student context*. In this activity everyone participates by acting out the motion with their bodies. Each student has his or her own interpretation of what the model is showing. Each person's interpretation acts as the context for that person's thinking and acting, and that becomes pertinent to the group of students and accessible to us as researchers when an individual describes their thinking and it is explicitly considered by others.

Previously the difficulty of alignment has been defined as the difficulty of applying the same concept across contexts, where "context" has previously meant what we call the *science context*, different situations in the external world where the concept is applicable. However, since we have delineated the student context and the classroom context we needed to carefully re-define and explain alignment in regards to each of those *contexts*. Thus, we denote three types of alignment difficulties: (1) the difficulty of applying the same concept across sciences contexts (not relevant to the current analysis, but seen elsewhere, for instance diSessa and Sherin, (1998)), (2) the difficulty of applying the same concept across classroom contexts (a classroom of students common interpretation of an ET model), (3) the difficulty of applying the same concept across student contexts (individual student's interpretations of their ET models as seen from a peers perspective).

An example of the second type of alignment can play out in the following manner: there could be a classroom-context/classroom-context disagreement. Imagine one classroom of students that generated and presented an ET model, while the other classroom of students was observing and interpreting their peer's model. For the observing students it may become apparent that their peer's model did not align with their conceptual elements. Perhaps the observing students extracted information that was unexpected, and the resulting inference did not match with their conceptual elements—resulting in a lack of alignment. For instance, perhaps they extracted more energy (people) entering the earth than leaving, leading to an inference that the earth would increase in temperature. Yet, this would be unexpected as the assigned task was to model the steady state energy of the earth and their conceptual elements would have likely supported them to expect equal amounts of energy entering and leaving. The lack of alignment could be resolved by either a modification to the model, which in turn would result in a change to the information extract by cuing different conceptual elements, such as modifying the number of people who enter and leave the earth. Or by directly changing the information extracted, while not changing the ET model, in other words, having the students extract different information from the original model, such as recognizing that individuals are not equal in their amount of energy, leaving open the possibility of despite more people entering than leaving, the earth being in steady state. Either of these modifications could result in learning given the possible incorporation or displacement of knowledge elements.

The third type of alignment can be an instance of a classroom-context/student-context disagreement and play out as follows. Imagine a classroom context in which students are (momentarily) in agreement as they collectively create an ET model. Then, when discussing the model, one student "sees" something new; they might recognize something in the model that is unexpected or perplexing. When that happens, a student might be extracting specific information from that model (the classroom context) that does not align with his or her own conceptual elements. This creates a new interpretation for that student, different from that of the rest of the class. This lack of alignment would likely be apparent in what the student says aloud and heard by peer(s). A result of this lack of alignment might be an attempt to bring that student's interpretation and their concept back into alignment. One option would be to modify the classroom context, change the ET model in such a way as that going forward the future information extracted would align with that students conceptual elements. Thus, after modification the student who previously recognized something being "off" and their peers would all be in agreement. Another potential solutions could involve changing the information extracted from the model thereby having that student cue different conceptual elements.

The problems of alignment and solutions to those problems become complicated because there are three kinds of context (science content, classroom context, and student context) and several ways to fix this kind of problem: modifying the model or modifying the information extracted. Changes to the information extracted can result in changes to the conceptual elements cued through either incorporation or displacement, which are potential instances of learning.

Data collection and analysis

The previously described theoretical extensions are illustrated with empirical examples from classroom data concerning 9th grade earth science students who were using an embodied modeling activity known as Energy Theater (ET) to generate and revise models for the steady state energy of the earth. In ET students create models with their bodies' positions and movements. This is a research-based and validated learning activity based on a substance metaphor for energy and designed to promote conceptual understanding of energy (Scherr, et al., 2012). In this article ET forms the backdrop for the analysis, for readers interested in this pedagogical activity we direct them to Scherr et al., (2013; 2012).

Classroom setting

Data were gathered from in-class observations from the classrooms of two 9th grade earth science teachers, Ms. Girard and Mr. London (aliases). Neither Ms. Girard nor Mr. London had done ET in their classroom before, Ms. Girard had attended an evening professional development activity in which Energy Theater was used and she contacted one author signifying that she and Mr. London wanted to repeat the activity with their students. She invited the authors to observe and study her class doing this activity. The teachers assigned their students the task of "Model the energy transfer/flow to show how Earth remains at a fairly constant temperature." Students were expected to use the tools of ET, which were written on the board. As is typical of this activity, each student was to portray a unit of energy. Ropes on the floor were used to delineate objects in the system. Movement from one area on the floor to another indicated movement of energy between objects in the system (e.g. earth and sun). A total of six classrooms (three periods taught by Ms. Girard and Mr. London, each, in parallel) were observed. One period was not studied further because class sizes were so small that nearly no discussion happened. In the remaining two periods, we observed a total of six enactments of ET—for each period, one individual class enactment for each class was followed by a joint enactment at the end of each period which included all the students from both classes.

There are some additional points about important aspects of the classroom discussion that will not be discussed further in this paper. ET is purposefully designed to be underspecified as to what one should do. Students had to figure out the model, and this led to a great deal of discussion, often in parallel, with many students talking at once. At times, the discussion was utterly chaotic, and we were unable to adequately discern the conversations. Our data analysis focuses on those moments when only one or a few speakers were talking, which was often facilitated by the teachers asking whole class to listen to one student's ideas.

From our observations, it became clear that the topic was not new to them. Students had studied the different wavelengths of light (ultraviolet, visible, and infrared) and what percentage of each wavelength transfers energy from the sun to the earth. They had also talked about the flow of energy away from the earth, but only in one wavelength. Overall, students entered into this activity with experience in the topic. Suggesting that the activity was not, for them, an opportunity to discover new ideas, but instead to build on prior ideas.

Knowledge analysis

Data consists of video and audio recordings of the classes participating in ET. Given that the discussions were utterly chaotic at times, we strategically decided to focus on the moments in which only one, or a few, speakers were talking. Those moments were transcribed in detail and figures were created in order to analytically track the ET models that were being created and modified. We noticed that many of these moments corresponded with instances in which there were disagreements that often resulted in modifications to the models. We began by examining the conceptual substance of these disagreements and searched for patterns. Eventually, there was a grounded process of coding for conceptual elements, relevant contexts, and moments of misalignment.

The empirical analysis was conducted using Knowledge Analysis. This is a methodological approach to studying knowledge with the purpose of examining learning and it has commonly been used in research that utilizes coordination class theory. We engaged in an iterative cycle of observing, schematizing, and systematizing the theoretical constructs across the data set (Parnafes & diSessa, 2013) while building the proposed extensions. Knowledge Analysis is a developing methodology and details about its history, theoretical foundations, and practical implementation are available in diSessa, Sherin, & Levin (2015).

Analysis

We illustrate our theoretical extensions with three cases of learning in a classroom environment. In the first case the alignment problem was fixed by a modification to the classroom context. In the second case the alignment problem was fixed by a modification to the information extracted. Finally, the third case is more complicated as fixing the alignment problem involved multiple solutions.

Case 1—Modifying the classroom context: Adding a tagging mechanism

The first case presents a lack of alignment between a student context (one student's interpretation of a classroom context) and their relevant conceptual elements, which was determined by the information that the student extracted from the model. This alignment problem was subsequently fixed by way of a modification to the classroom context. Mr. London's classroom of students generated a model of the energy entering and leaving the earth with no sun. Students walked in small circles around the North Pole, South Pole, and Equator, representing energy entering and leaving the earth at those three locations (see Figure 1). During the process of acting out the model, one student stopped the group to point out that the earth was not in a steady state ("It's not in a steady state though.") There was a lack of regulation for when energy entered and left the earth ("That is not steady state because it's not really regulated.")

For this student, at least, a conceptual understanding of steady state included knowledge about a need for a constant number of people inside the circle at a time (to represent a constant amount of energy in the earth) and while acting out this model that student extracted information that violated what the student expected to see as based on available conceptual elements. The number of people inside the circle varied because it was not "well regulated." This was a lack of alignment between the student's interpretation of the classroom context and that student's conceptual elements. Subsequently a solution was proposed, a tagging mechanism to control when individuals enter and leave: "We should have a few people walking around in there, and then at some point switch a person out and, there has to be, an input and an output. So say, two people on the pole and three on the outside. So one person walks out and one person walks in... Give them a tag out and you'll switch... just a tap out." This modification to the ET model, having individuals tag each other when entering and leaving the earth so as to insure an equal input and output (Figure 1) had the effect of returning to alignment the student's interpretation of the classroom context and their conceptual elements. It also had the effect of creating agreement between the student context (that single student's interpretation of the classroom context) and the entire classroom context.

A lack of alignment, making a change to the classroom context, and a subsequent return to alignment is a type of learning. Individual learning occurred for the student who pointed out that the earth was not in steady state because of a lack of energy regulation; that individual worked the concept of steady state across contexts—the initial problematic student context and the subsequent classroom context—while incorporating the regulation of energy flow into a new conceptualization. Additionally, group learning occurred through the incorporation of a new element, regulation of energy flow, into the group's conceptualization of steady state; this is evident in the classroom agreement about the lack of regulation as suggested by the proposal and implementation of the tagging mechanism.

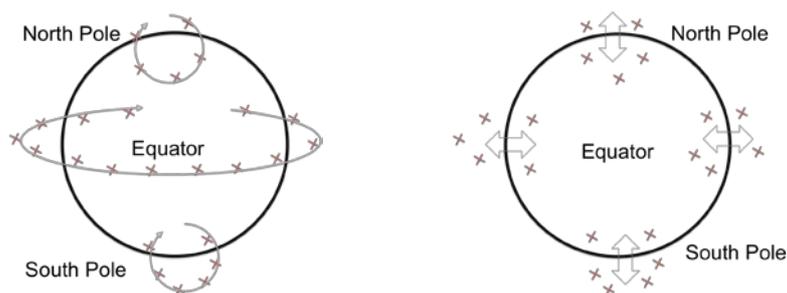


Figure 1. Mr. London's first period class' ET model before (left) and after (right) modification.

Case 2—Modifying the information extracted: The earth would explode

The second case involved one classroom presenting their ET model to the other classroom. There was a lack of alignment between a classroom context (the ET model one class generated) and information that the other class extracted from the former's classroom context. The lack of alignment was eventually fixed through a modification to the information extracted. Mr. London's second period class generated an ET model with three circular ropes on the floor, representing the sun and earth, and a third larger rope surrounding both to represent the boundary of the model. Three people at a time left the sun representing three types of light (ultraviolet, infrared, and visible)

and entered the earth (see Figure 2). Afterwards, one person representing infrared light left the earth and went into space. Mr. London's students had intended that the three people entering the earth were worth the same amount of light as the one person leaving, but they represented different types of light.

While observing the model Girard's students questioned why there were three people entering the earth and one person leaving ("How come there is only one thing leaving but three coming in? I think the earth would explode eventually.") Likely the Girard students were accessing conceptual elements that led them to expect equal amounts of energy entering and leaving the earth and they were expecting to extract information aligned with those elements, but that did not occur. In response to the question London students explained that three people entering were meant to be "worth" the same amount of energy as one leaving ("He was worth all three of us. Cause we were just representing all the different kinds of energy that were going through the earth. And he was the one kind of energy coming out of the earth....Oh, so if we were all equal in parts, then the earth would explode. But we're not. There is more of Ian than all of us.") In this episode there was a lack of alignment between the information Girard students extracted, as related to their conceptual elements, and the London student's classroom context. This lack of alignment was fixed by a suggestion that the Girard students extract something different: equal amount of energy in and out of the earth being represented by the one student leaving the earth ("Ian") being worth the same amount of energy as three people entering. This change in the information extracted is an example of working the concept across contexts resulting in a return to alignment.

Modifying the information extracted has the potential to be a good opportunity for learning. It could result in displacement or incorporation of knowledge elements into a new conceptualization. In this case Girard's class' conceptualization of steady state likely contained an element of equal amount of energy (light) entering and leaving the earth. Likely those students' conceptualization contained this element before and after this moment, thus this may not have been an instance of learning for them. But, differently, London's class' conceptualization might have been modified during the discussion. Specifically, this student might have been experiencing a displacement of knowledge about unequal numbers of people entering and leaving with knowledge about equal people entering and leaving. It's not clear from the discussion whether or not the students' knowledge was displaced, but there is additional evidence for the existence of this displacement to be found in the final second period joint ET model in which there were equal numbers of people entering and leaving the earth, suggesting that for London's class there might have been some group learning.

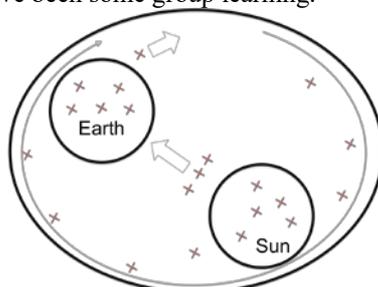


Figure 2. Mr. London's second period classes ET model with three people entering the earth and one leaving.

Case 3—Modifying both the classroom context and the information extracted

The final case presents an example of a lack of alignment in which there was multiple solutions proposed. The first solution involved modification to the ET model; the second solution involved modification to the information extracted. The first solution was eventually taken up in the final joint model. Mr. London's class from Case 1 above presented their model (with the tagging mechanism) to Girard's class. This model did not contain a sun, and during the presentation there arose an alignment problem about the lack of a sun. The problem was a lack of alignment between the London class' classroom context and the Girard class' conceptual understanding. Girard's class' conceptual elements likely involved a sun as a source of energy; their model, not discussed, included a sun as a source. The Girard students observed London's class' model and failed to extract a sun. ("Where is your, if you have energy on the outside, where is it coming from since you don't have a sun?") The London students understood the predicament and they entertained the possibility of changing their classroom context (their ET model) by adding a sun ("You didn't give us anything to be the sun! // Yeah, we didn't have anything for the..! // Mr. London // I am the sun! // Mr. London is the sun.") Modifying their model by adding a sun would have had the effect of solving the alignment problem.

Yet, a second solution was also proposed to the problem of alignment. As the London students recognized that their classroom context did not contain a sun, they recognized the lack of alignment predicament this created for the Girard students. However, rather than accept it as a problem to be fixed by modifying their ET model, they instead asserted that the problem does not exist ("It's pretty much, not, where it's coming from but how it would

affect the earth itself. // This is just a model, this is the system boundary.”) The London’s students’ classroom context involved only the inflow and outflow from the earth, not the source. If the Girard students were to change their conceptualization of the phenomena, by changing which conceptual elements were applicable to also only focus on inflow and outflow, that modification would fix the alignment problem. Of the two solutions proposed to address the lack of alignment problem— modifying the model by adding a sun, or modifying the Girard’s students’ conceptualization of the phenomena—the former was taken up in the final joint model which included a sun (see Figure 3).

In this case, the lack of alignment afforded multiple opportunities for group learning. The London students considered adding a sun to their classroom context. Adding a sun (a source) to their conceptualization of steady state could have been an example of incorporation of a new knowledge element into a prior conceptualization. Another learning opportunity occurred when the London students suggested that the Girard students change their conception to focus only on inflow and outflow. If the Girard students had done this, it could have been an instance of group learning by displacing a knowledge element about needing a source of energy with a new element about focusing on only inflow and outflow. In the end, the final joint model included a sun and earth and the movement of people was from the sun, to the earth, and then outside the bounds of the model where the students prepared to re-enter the model at the sun. The suggestion to focus only on inflow and outflow was not followed up on. As evidenced by the final joint model, the London students incorporated knowledge about a source into their conceptualization, thereby learning, while in this instance, the Girard students did not observably change their conception.

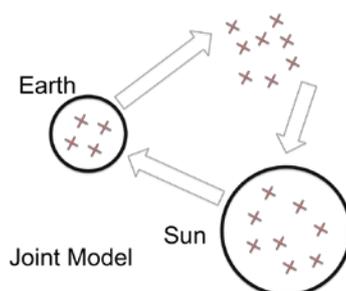


Figure 3. Illustration of the joint model from the first period classes.

Discussion and conclusion

These cases illustrate group and individual conceptual learning in a classroom environment. In the first case, one student’s understanding of steady state was not apparent in that student’s interpretation of the model that had been generated by their class. This problem was fixed by a modification to the classroom context that resulted in aligning that students’ interpretation of the classroom context with the student’s conceptual understanding— illustrating individual learning. Furthermore, the modification may have resulted in group learning through incorporating a new element. The second case involved one class presenting its model to the other classroom and a lack of alignment between the presenting class’ classroom context and the information that the observing class extracted. The fix involved a modification to the information that the observing class extracted; quite literally, they changed their interpretation of what they saw. Finally, the third case, similar to the second case, involved one classroom presenting its model to the other classroom and a lack of alignment between the presenter’s classroom context and the observer’s conceptual understanding. Two potential solutions were proposed: modifying the classroom context or modifying the information extracted. Both solutions were learning opportunities and the first solution was taken up in the final joint model. All three cases involved a difficulty, lack of alignment, and a solution; all three were case of learning. The first case incorporated individual and group learning, while in the second and third cases only group learning was observed.

To capture and delineate individual and group learning we needed to distinguish the individual and classroom contexts which necessitated understanding that different individuals can have differing interpretations of the same model that then affect their subsequent reasoning and conceptual learning during key moments of classroom discussion. Practically, we were able to capture these moments through verbal disagreement and negotiation in the classroom, this is important as previously coordination class analyses tended to focus on learning during cognitive-clinical interviews. For this analysis, we expanded the notion of “context” within coordination class theory, thereby expanding our understanding of learning difficulties and successes in classroom environments. A limitation of the current work was only one scientific context, energy of the earth, while elsewhere the focus has been on individual’s understanding across multiple scientific contexts (e.g. Parnafes,

2007). Focusing on multiple scientific contexts in a classroom where individual and group learning is present would be a natural direction for future work.

Capturing both individual and group learning in a classroom setting is unusual. Both the individual and collective group levels are critical for analyzing classroom learning (Cobb et al., 2001). One should not focus on collaborative learning to the detriment of individual learning (Conlin & Hammer, 2015). But, all too often, theories of learning and conceptual change have emphasized either group learning or individual learning and not integrated the two. Success in building on a theory of individual learning to expand it into a group-learning environment in a manner that allows integration of the two suggests new utility and areas of relevance for coordination class theory. More generally, this machinery helps us understand how a single student, or a single utterance can rather dramatically change a model, and when that happens, the kind of learning that can ensue, for that individual, and the entire class.

Finally, there are implications for the design of learning environments. Designing for this kind of learning requires facilitating an environment in which students, both collectively and individually, have opportunities to generate and modify the artifact or means by which they present their conceptual ideas. An important feature of this data was that the modifications to the ET models were practically instantaneous, relatively easy to execute, and had the potential to be observed by all individuals. When things did not go as intended or expected, there was a low bar to recognize it and that facilitated quick modifications. The classroom lessons were only 45-minutes, and the students had no previous experience with Energy Theater. Possibly these features had a strong influence on facilitating learning; recognizing these key features and building them into other learning environments may be crucial for promoting this type of group learning in the future.

References

- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences, 10*(1-2), 113-163.
- Conlin, L. D. & Hammer, D. (2015) Commentary: From the Individual to the Ensemble and Back Again. In A. diSessa, M. Levin, & N. J. S. Brown (Eds.), *Knowledge and interaction: A synthetic agenda for the learning sciences*. New York, NY: Routledge.
- diSessa, A. A. (2002). Why “conceptual ecology” is a good idea. In *Reconsidering conceptual change: Issues in theory and practice* (pp. 28-60). Springer Netherlands.
- diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences, 13*(1), 77-103.
- diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change?. *International Journal of Science Education, 20*(10), 1155-1191
- diSessa, A. A., Sherin, B., & Levin, M. (2015). Knowledge Analysis: An introduction. In A. diSessa, M. Levin, & N. J. S. Brown (Eds.), *Knowledge and interaction: A synthetic agenda for the learning sciences*. New York, NY: Routledge.
- diSessa, A. A., & Wagner, J. F. (2005). What coordination has to say about transfer. *Transfer of learning from a modern multi-disciplinary perspective*, 121-154.
- Livrini, O. & diSessa A. A. (2008), How Students Learn from Multiple Contexts and Definitions: Proper Time as a Coordination Class. *Physical Review Special Topics-Physics Education Research, 4*, 1-18.
- Parnafes, O. (2007). What does “fast” mean? Understanding the physical world through computational representations. *The Journal of the Learning Sciences, 16*(3), 415-450.
- Parnafes, O., & diSessa, A. A. (2013). Microgenetic learning analysis: A methodology for studying knowledge in transition. *Human Development, 56*(1), 5-37.
- Scherr, R. E., Close, H. G., Close, E. W., Flood, V. J., McKagan, S. B., Robertson ... and Vokos, S. (2013). Negotiating energy dynamics through embodied action in a materially structured environment. *Physical Review - Special Topics: Physics Education Research 9*(2), 020105 020101-020118.
- Scherr, R. E., Close, H. G., McKagan, S. B., and Vokos, S. (2012). Representing energy. I. Representing a substance ontology for energy. *Physical Review - Special Topics: Physics Education Research 8*(2), 020114 1-11.
- Thaden-Koch, T., Mestre, J., Dufresne, R., Gerace, W., & Leonard, W. (2005). When transfer fails: Effect of knowledge, expectations and observations on transfer in physics. In J. Mestre (Ed.), *Transfer of learning: Research and perspectives*. Greenwich, CT: Information Age Publishing.
- Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction, 24*(1), 1-71.
- Wittmann, M. C. (2002). The object coordination class applied to wave pulses: analysing student reasoning in wave physics. *International Journal of Science Education, 24*(1), 97-118.

Examining the Impacts of Annotation and Automated Guidance on Essay Revision and Science Learning

Libby Gerard, University of California, Berkeley, libby.gerard@gmail.com

Marcia C. Linn, University of California, Berkeley, mclinn@berkeley.edu

Jacque Madhok, University of California, Berkeley, jjmadhok@gmail.com

Abstract: Automated guidance can facilitate student revision of explanations and arguments in online inquiry science units. We explore ways to design guidance for short essays that promotes meaningful revision rather than superficial changes. Specifically, we compare the affordances of annotation of a fictional essay to knowledge integration guidance on revision of science writing. 293 middle-school students were randomly assigned to condition. Students who annotated an essay made significantly greater pre to post test gains and were also better able to use automated guidance on a posttest item than students who only received knowledge integration guidance. These findings suggest ways to support students to revise science writing and build integrated understanding of science.

Keywords: technology, revision, writing, student learning, science, knowledge integration

“I like when there is writing because with typing it feels like I can explain more instead of...just doing multiple choice, because then its kind of like I want to explain it really bad but you can only put an answer.” (7th grade study participant)

Introduction

Writing in science gives students the opportunity to defend claims and integrate multiple pieces of evidence gathered from a variety of contexts. A substantive literature demonstrates the value of writing and revising activities that prompt students to build connections between their initial ideas and new ideas (Fitzgerald, 1987). The centrality of writing and revision in scientific practice are reflected by the Next Generation Science Standards. Students are expected to construct explanations and arguments, and, to “identify flaws in their own arguments and modify and improve them in response to criticism” (NGSS, 2013). Writing and revision of arguments is integral to long-term science learning (Rivard, 1994). In spite of the importance of writing and revision, these activities are rare in the science classroom. This is due largely to the time required of teachers to provide individualized feedback for often at least 150 students. A NAEP report found that about half of Grade 11 students (52%) had never completed a written report of any kind in science (Mullis & Jenkins, 1988).

Advanced natural language processing tools can help science teachers integrate writing and revision activities by automatically scoring student written short essays about an inquiry topic. Individualized guidance can be assigned immediately based on the score to help students revise their essays (Liu et al, 2014; Roscoe & McNamara, 2012). Some essay writing tools (e.g. Criterion) that give diagnostic feedback on essay mechanics and organization have been widely deployed in classrooms to increase writing and revision opportunities. We explore ways to design guidance that strengthens the argument in the essay.

Students generally find revision in science unfamiliar and challenging. We characterize revision as involving revisiting or gathering new evidence, comparing initial ideas with new ideas, and reformulating connections. Class norms typically support completeness and correctness rather than refinement. In analyses of classroom argumentation Berland & Reiser (2011) found students rarely revised their ideas in light of challenges or questions posed. Studies consistently find that students most often make mechanical revisions to their essays when given feedback, rather than substantive changes to the content as suggested (Roscoe & McNamara, 2012). Less competent writers tend to make changes to spelling and grammar rather than to revise for meaning (Fitzgerald, 1987). This is particularly problematic for science because most pre-college students are novice science writers and have limited understanding of what makes a good science explanation (Sato & Linn, 2014).

Our prior work suggests that individualized, automated guidance for students’ short essays can strengthen students’ revisions and science learning when designed to promote knowledge integration (Linn & Eylon, 2011). Guidance designed according to the knowledge integration perspective prompts students to distinguish between their own scientific ideas and new ideas introduced in instruction. The knowledge integration guidance led to significantly more productive essay revisions, and subsequently more coherent and accurate science essays than did generic guidance (e.g. “add more evidence) or specific guidance (e.g. “Incorrect. Energy transforms from light energy into chemical energy) (Author et al., 2015).

Analyses also revealed difficulties students faced in integrating ideas during the essay revision process. Only a small percentage of the students integrated a new idea into their essay meaning that they edited their initial essay when revising to build a connection between their initial ideas and the new information they added. Rather the majority of students added a new idea without connecting it to their initial response; the new idea was “tacked-on”. On a post test taken a week after the writing and revision activities, the students who had made integrated revisions demonstrated a significantly deeper understanding of the science concepts than those students who had added new but disconnected ideas in their revision (Gerard & Linn, 2015). This suggests that the practice of revision, when done in a way that builds and refines connections among ideas, can set in motion a successful learning strategy for the rest of the unit.

In this study, we explore ways to help students learn to revise essays by integrating new ideas. We designed an essay annotation activity to help students learn how to use guidance to revise science writing. In the annotator activity, students identify gaps in a fictional student’s essay and place pre-determined labels on the essay as hints for revision (Figure 1). The hints call for the fictional student to edit her essay to connect new ideas. We compare students’ use of annotation and then automated guidance to revise short essays in an inquiry unit, to students’ use of multiple rounds of automated guidance to revise their short essays.

Methods

This instructional comparison study investigates two approaches to guiding students essay revision and examines their impacts on science learning. Specifically, we compare the affordances of annotation of a fictional essay versus additional practice with automated guidance on revision of science writing.

Participants

Two-hundred and ninety-three students in three teachers’ classrooms in one public middle school were randomly assigned within class periods to either the Annotation + revision condition, or, Two revisions condition. The school serves a moderately diverse student population (47% of students are an ethnicity other than Caucasian; 26% receive a free or reduced price lunch). Students worked in pairs in the unit and completed the pretest and posttest individually.

Curriculum and embedded assessments

We used the Web-based Inquiry Science Environment (WISE, <http://wise.berkeley.edu>) to randomly assign conditions to students within classes. Students studied the WISE Photosynthesis and Cell Respiration unit, which guides students to investigate energy flow in plants and animals. The unit incorporates dynamic visualizations and generative activities, including writing short essays, to help students gather evidence and integrate ideas. Students studied the WISE unit, led by their regular classroom teacher, for 6-8 class periods (50 minutes each) spread over 2 weeks.

We selected three short essays in the unit for annotation, guidance and revision. The selected essays call for students to integrate multiple pieces of evidence from the unit to explain energy transfer and transformation. Our analysis focuses on one of the three essays, called GreenRoof (Figure 1):

GreenRoof: Mary heard that growing plants on a roof could lower a house’s energy usage. Mary does not understand why and how plants could help. Write an energy story to explain to Mary what happens to energy from the sun in the picture? How could growing plants on the roof reduce the house’s energy usage?

Each of the short essays included automated scoring by c-raterML™, a natural language processing tool developed by the Educational Testing Service (see Liu et al, 2014 for model development info). WISE assigned personalized guidance for the student’s essay immediately based on the automated, craterML score (Table 1).

Essay revision

For each of the three short essay steps students received personalized, automated guidance (Table 1). To assess the value of the annotation activity the two conditions were:

1. Annotation+revision: students first annotate the response of a fictitious student, then they get personalized automated guidance on their own response.
2. Two revisions: students get two rounds of personalized, automated guidance on their own response. We constrained the algorithm to assign unique guidance to the second revision even if the score did not change.

In the Annotation+revision condition, students were prompted to write their short essay, and then to annotate Mary's essay using pre-defined labels. Mary's essays included vague ideas as well as a mix of normative and non-normative ideas (Figure 1). The pre-defined annotation labels included questions about energy flow that were general enough to be used across all three short essays: (a) Where does energy come from? (b) How does energy change – what is the process? (c) How does the energy move? (d) Why are plants important – what would happen with no plants? (Figure 1). We used the same labels on each essay to model for students the questions they could ask themselves when they wrote and revised future essays. After annotating Mary's essay, students were prompted to submit their own essay for one round of automated guidance and revision.

Annotation activity

GreenRoof essay and auto guidance

Figure 1. Revision activities.

In the Two revisions condition, students were prompted to write their short essay (Figure 1). They were then given two chances to submit their essay for automated guidance and to revise. All guidance was designed to help students move up one level in the knowledge integration rubric (Table 1).

Table 1: GreenRoof knowledge integration scoring rubric and personalized, automated guidance

Score, Criteria	Student Essay Examples	Automated Guidance Round 1
1 Off Task	IDK	{Student name}, think about the Sun's rays. Where do they go? How do they move? Check out Step 2.1 for a hint. Then, rewrite your story below.
2 Incomplete Scientifically non-normative OR irrelevant OR repeats the question	It helps reduce energy usage because you can get energy that you can use for your house through the plants.	{Student name}, expand your story. How does the energy change inside the chloroplast? Check out Step 3.2 for a hint. Then, revise your story below.
3 Partial Partial link of key ideas OR normative and non-normative ideas linked	The sun gives light energy to the plants so they can photosynthesize and the plant give off oxygen and energy to the world.	{Student name}, explain the connections between your ideas. How does the plant transform light energy to energy it can use? Check out Step 2.11 for a hint. Then, revise your story below.
4 Full 1 full link [ok if response also has non-normative ideas, as long as not connected to link]	One day Brent and Emilioobserved the green roof. They noticed the sun's light energy was being absorbed by the plants on the roof....They thought the plant turned the light energy to chemical energy.	{Student name}, good progress. Now expand your story. Why are the plants important - what would happen to sunlight that hits the side of the roof with NO plants? Check out Step 1.4 for a hint. Then, revise your story below.

		Step 1.4 for a hint. Then, revise your story below.
5 Complex 2 or more full links	Plants can lower energy usage when there on a roof because they absorb the sun and turn that energy into chemical energy. Without the plants on the roof it get hot since sunlight energy will turn into heat	{Student name}, great reasoning. Now, check over your story to make sure it describes the energy flow from the sun to both sides of the roof. Revise as needed.

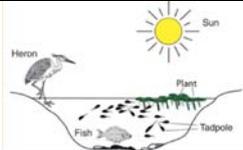
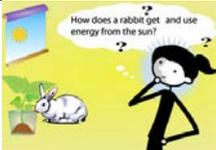
Data sources

We used a pretest, embedded assessments, and a posttest to capture student science learning and student ability to use guidance to revise. The pretest was taken one day prior to starting the unit. The posttest was taken approximately three days after completing the unit and one week after the last essay revision activity. The pre/post test included two short essays, Tadpole and Rabbit [Table 2].

To assess student ability to use automated guidance, for the Tadpole essay, students received one round of personalized, automated guidance and revised their response. The two essay activities were the same on the pretest and posttest.

We conducted student interviews and classroom observations to gather further information on how students used the writing and revision tools, and student views on writing and revision in science in general.

Table 2: Pre/Post Test Short Essays

Tadpole*		Explain how the tadpole in the picture gets and uses energy to grow. After you are done writing, press “ Check Answer ”. You will have 1 chance to get feedback and revise your story.
Rabbit		There is a rabbit in the classroom. Mary wonders how the rabbit gets and uses energy from the sun. Write a story using scientific evidence to explain to Mary how the rabbit gets and uses energy from the sun.

* Students received one round of personalized, automated guidance and were prompted to revise.

Data analysis

We scored students’ embedded and pre/post test responses for scientific accuracy and coherence using knowledge integration rubrics (see Table 1). The rubric rewarded students for making links among ideas (Liu, Lee & Linn, 2011). We scored students’ initial and revised essays on the embedded items, and students initial and revised essays on the Tadpole pretest and posttest item. Two researchers blind coded all of the student essays. Researchers reached 97% agreement and worked out the disagreements until coming to a consensus.

In addition to scoring the data with knowledge integration rubrics, we also coded the students’ embedded essays for the kind of revisions they made based on the automated guidance. We compared their initial essay to their revised essay and categorized the type of writing change (see Table 3). In this analysis, we coded the changes in the students’ science writing not the scientific accuracy of the change.

Table 3: Revision Rubric

Score	Category/Description	Example	
		Initial Essay	Revised Essay
3	Integrated: Incorporated an idea into the middle of the essay. Edited the initial essay to formulate the connection to the new ideas.	The plants on the roof will take the energy to create photosynthesis and grow. Since the plants are on top of the roof it is closer to the sun taking energy from the house.	The plants on the roof will take the energy to create photosynthesis and grow. The plant uses carbon dioxide fro which we breathe out, water, and light energy to create glucose. Since the plants are on top of the roof it is closer to the sun taking energy from the house.

2	Added Different: Added a new idea to end of essay that is different than ideas in initial essay. Does not make any changes to initial essay. Or, deleted initial essay and wrote new essay with different ideas.	The plant uses the left over sunlight from what the house leaves behind	The suns energy goes to the plants then, plants absorb the light energy. The energy goes to the house. The plants change the energy into the chemical energy when it goes through the plants
1	Added Similar: Added a new idea to the end of the essay that expanded with more details, or repeated in different words, the ideas in the initial essay. Does not change initial essay.	The plants soak up the energy and provide light to the house. This will make it so that you will not have to use lights during the day.	The plants soak up the energy and provide light to the house. This will make it so that you will not use lights during the day. Also it will help the energy bill go down because it is using solar energy.
0	No Change		

Findings

Embedded assessment

In both conditions, students used the automated guidance tools to significantly improve their GreenRoof essay. Students assigned to annotation+revision outperformed students assigned to two revisions [N=152 pairs Annotate M=.46 SD=.72; Auto Guidance Two Rounds M=.28 SD=.56, $t(149)=1.72$, $p=.09$]. The difference between conditions was significant for students who began the writing activity with mid to high prior knowledge [Gain, High Prior, n=92 pairs, Annotate+revision M=.46 SD=.55; Two revisions M=.18 SD=.39, $t(90)=2.9$, $p<.01$]. We defined mid/high prior knowledge as students who scored a 3 or above on the knowledge integration rubric on their initial GreenRoof essay.

Students' revision approach differed significantly across the annotation+revision and two revisions conditions [Pearson $\chi^2(3)=14.11$ Pr=.003]; Table 5]. The annotation+revision condition led significantly more students to add new and different ideas to their essay when they revised. Alternatively, students who received two rounds of automated guidance were more likely to add ideas that were similar to those they had already written. Further, the annotation condition led substantially more mid/high prior knowledge students to make a revision; 88% of high prior knowledge students in the annotation+revision condition compared to 78% of high prior knowledge students in the two revisions condition. Approximately 19% of students in both conditions integrated a new idea into their initial essay. There was no difference between conditions in the percentage of students who made integrated revisions. Mid/high prior knowledge students were more likely to integrate than low prior knowledge, consistent with the literature on revision (Fitzgerald, 1987).

Table 5: Percent of students making each revision type on embedded essay by condition and prior knowledge

	All Students		High Prior		Low Prior	
	Two revisions	Annotation + revision	Two revisions	Annotation + revision	Two revisions	Annotation + revision
No revision	24	16	22	12	27	20
Added Similar	39	30	40	37	36	23
Added Different	18	34	14	24	27	46
Integrated	19	20	24	27	9	11

Student interviews and classroom observations suggest that the annotation+revision condition helped students use guidance to critically examine their own essay. As one student reported, “*I think that the Mary’s story [annotate activity] ... helps you change yours [essay] because you take the feedback you are giving Mary and then you kind of apply it to your own.*” Students used the prompts that were repeated in annotation labels (e.g. how does energy change?) as they reflected on how to revise their essay. Another student reported: “*Since you have to move [the labels] around and see where it goes, it makes you think that if that was you doing the question and the teacher gave you that feedback...it kind of makes you think ‘Oh ya, that should go there’ and then it can help you when you have a question similar you can be like oh I remember that and I should add that in mine now.*”

Pre/post test

We report the results of each pre/post test essay separately since students received automated guidance on the Tadpole short essay, and did not receive automated guidance on the Rabbit essay. Students in both conditions significantly improved from pretest to posttest [Rabbit $t(282)=8.72$, $p<.0001$; Tadpole $t(276)=7.78$, $p<.0001$] [Figure 2]. There was no difference in gains between the two conditions on the Rabbit essay.

On the Tadpole short essay where all students received automated guidance, students in the annotation+revision condition made significantly greater pre to posttest gains than the students in the two revisions condition, Figure 2 [N=261 students, Annotate+revision $M=.58$ $SD=1.02$; Two revisions $M=.29$ $SD=1.03$, $t(259)=2.29$, $p<.05$].

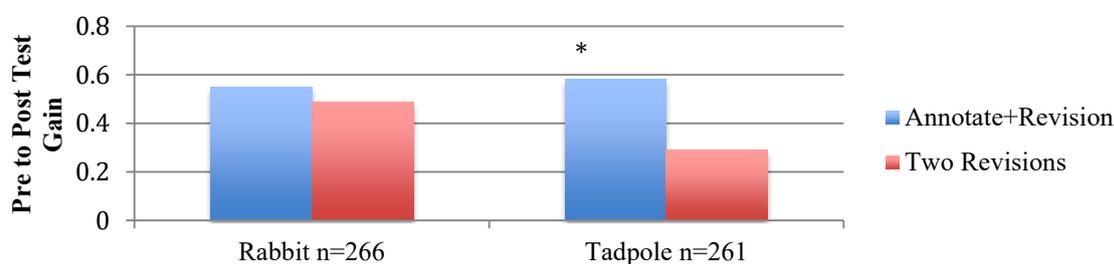


Figure 2. Student pre to post test gains on two short essays. The tadpole essay included automated, guidance.

Further, by posttest the students in the annotate+revision condition took better advantage of the automated guidance on the Tadpole essay than students in the two revisions condition. The analysis suggests that the annotation activity helped students learn how to use guidance to revise their writing and improve the coherence of their understanding.

Students who made an integrated change when revising their GreenRoof essay during the unit, made the greatest pre to post test gains on both short essays [Figure 3]. This strengthens our earlier findings demonstrating the benefits of integrating as a revision approach (Author, 2015). Further design work is needed to improve the percentage of students who integrate when revising.

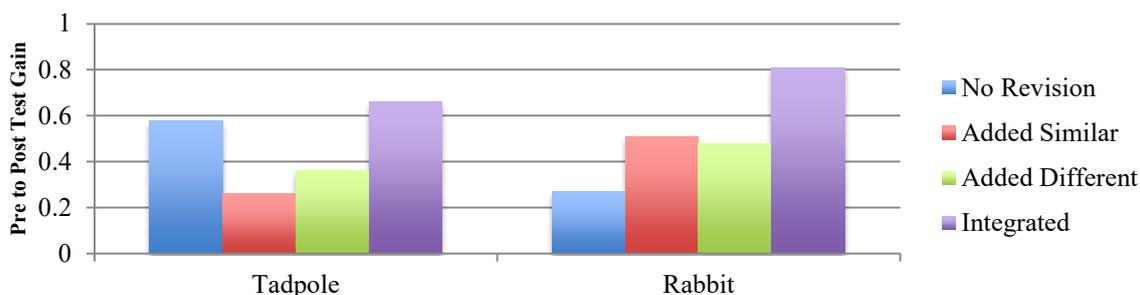


Figure 3. Pre to post test gain by type of revision made on an essay embedded in the unit.

Example of student revision using annotation and auto guidance

We selected a case of a student, Erica, whose work was representative of those in the annotate+revision condition. The case documents how Erica used annotation and automated guidance to strengthen her essays.

Erica began the Photosynthesis/Cell Respiration project with a mix of ideas about energy transfer and energy forms. On the pretest Tadpole essay, she explained a tadpole “gets energy straight from the sun. The tadpoles use the energy to store and create kinetic energy.” This response seems to draw on vocabulary that is not fully understood.

Erica gathered new ideas from visualizations of energy flow in photosynthesis. She connected these ideas with her intuitive views about energy in her initial GreenRoof essay [Figure 4]. She explained that plants get energy from the sun and give energy to the house. Erica then annotated Mary’s GreenRoof essay placing the labels “where she wanted [Mary’s essay] to explain more”. Next, Erica received automated guidance on her essay. The guidance helped Erica see that her essay generally needed improvement, but she did not know what to change. Erica drew on her annotation experience to consider what to revise. “It [the annotator] helped me use [automated] feedback because I could go back and look at those questions [the annotator labels] I had put down and use it for myself.” Erica then revised her GreenRoof essay to add a new idea about photosynthesis, while still holding onto her non-normative ideas about the transfer of energy from plants to the house.

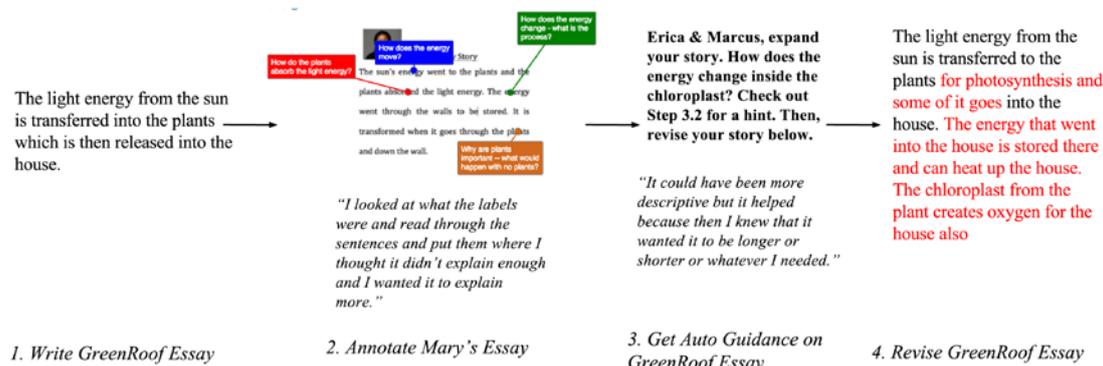


Figure 4. Erica's writing, annotation, and revision on the GreenRoof essay.

After the GreenRoof essay activity, Erica explored visualizations of energy flow in cellular respiration. She then wrote the EnergySun essay, the last essay writing activity in the unit [Figure 5]. Erica wrote a vague essay indicating she was still confused about energy transfer. Erica next annotated Mary's EnergySun essay "placing the labels where it was hard to understand [what Mary wrote] and [she] wanted to know more." These comments seem to reflect Erica's growing realization that she needs to find out more about energy transfer and transformation. Next, Erica received personalized, automated guidance on her essay. The guidance helped Erica understand what she needed to find out whereas at first she said, she "didn't know what to write." Erica revised her essay to integrate detailed links about energy transformation and storage in glucose.

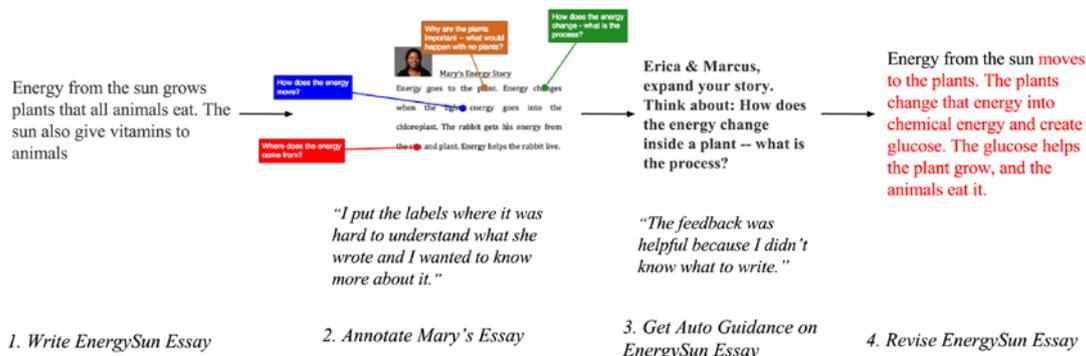


Figure 5. Erica's writing, annotation, and revision on the EnergySun essay

By the time of the posttest, Erica had formulated a coherent account of energy flow. She explained on the posttest how a tadpole gets and uses energy. "...The sun gives off radiation to the plant which absorbs it into its chloroplast. The chloroplast transforms it into chemical energy to make glucose which helps it grow. The tadpoles then eat the plants..." The annotate+revision condition helped Erica learn how to revise a science essay. She drew on the big questions about energy that were in the annotator labels she used to critique Mary's essay to inform her own writing and revision.

Conclusions and implications

This study investigated two different approaches to supporting students' use of automated guidance to revise and strengthen their science writing and understanding. These results demonstrate that automated knowledge integration guidance can help students refine their ideas about energy transfer and transformation in photosynthesis and cellular respiration. It is noteworthy that both conditions guided students to figure out how to revise their own essays rather than providing them with the answer.

In providing students a fictional peer's essay to annotate, we found that students transferred the hints they used to critique a peer's essay to their reflection on how to apply guidance to improve their own essay.

Students' reflective process may explain why students in the annotation+revision condition later showed greater gains on the posttest when revising an essay with automated guidance than those students in the two revisions condition. The annotation+revision condition helped students learn how to revise using guidance.

Students have difficulty revising scientific arguments and often tack ideas on to their essays rather than integrating them to increase the coherence of their responses (Gerard & Linn, 2015). We designed the annotation activities to model the process of editing within an essay, as opposed to only adding ideas to the end. A minority of students however in both conditions (19%) integrated new ideas when revising essays during the unit. The annotate condition increased the likelihood that students would add new and different ideas rather than elaborate on their existing ideas. Research shows that when students are asked to add new ideas they integrate their ideas more effectively than when they are asked to add similar ideas (Matuk & Linn, 2015). The case study reinforces these interpretations by showing that the annotate condition helped Erica add key ideas in energy transfer and transformation.

Students' propensity to add-on ideas, similar or different, rather than integrate new ideas may indicate the difficulty and the unfamiliarity that integrating ideas presents to students as they learn a new topic. Traditional school culture prioritizes adding a "right" idea over distinguishing and building connections among one's initial views and the new information presented. This finding also indicates areas for improvement in the annotation task design. Although the annotation hints (e.g. "how does energy change?") were intended to prompt the fictional student to both fill gaps in her essay *and* modify non-normative ideas, our student interviews suggest students perceived the hints as guidance suggesting the fictional student add more information about this idea. Future design iterations will aim to focus students' attention on both adding and modifying during the annotation activity.

In all, this study reveals how tools like annotation and automated guidance can help students to engage in meaningful writing and revision practices in the science classroom. Knowledge integration guidance as well as using annotations to model the process of using guidance can significantly strengthen essay revisions.

References

- Fitzgerald, J., (1987). Research on revision in writing. *Review of Educational Research*, 57(4), 481-506.
- Gerard, L., Ryoo, K., McElhaney, K., Liu, L., Rafferty, A., & Linn, M.C. (2015, online first). Automated guidance for student inquiry. *Journal of Educational Psychology*
- Gerard, L., & Linn, M.C. (2015). *Writing and revising in science*. Paper accepted for the annual meeting of the American Education Research Association, Washington D.C
- Linn, M. C., & Eylon, B.-S. (2011). *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration*. New York: Routledge
- Liu, O. L., Brew, C., Blackmore, J., Gerard, L. F., Madhok, J., & Linn, M. C. (2014). Automated scoring of constructed response science items: Prospects and obstacles. *Educational Measurement: Issues and Practice*, 33(2), 19-28
- Liu, O.L., Lee, H.S., & Linn, M.C. (2011). Measuring knowledge integration: Validation of four-year assessments. *Journal of Research in Science Teaching*, 48(9), 1079-1107
- Matuk, C. & Linn, M. C. (2015). Examining the real and perceived impacts of a public idea repository on literacy and science inquiry. In *Proceedings of the 11th Intl Conference for Computer Supported Collaborative Learning*, Gothenburg, Sweden: International Society of the Learning Sciences.
- Mullis, I.V. S. & Jenkins, L.B. (1988). *The science report card: Elements of risk and recovery. Trends and achievement based on the 1986 national assessment (National Assessment of Educational Progress)*. Princeton, NJ: Educational Testing Service.
- NGSS Lead States (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press
- Rivard, L. O. P. (1994), A review of writing to learn in science: Implications for practice and research. *Journal of Research in Science Teaching* (31), 969-983
- Roscoe, R. D., & McNamara, D. S. (2013). Writing Pal: Feasibility of an intelligent writing strategy tutor in the high school classroom. *Journal of Educational Psychology*, 105(4), 1010-1025
- Berland, L. & Reiser, B. (2011) Classroom communities' adaptations of the practice of scientific argumentation. *Science Education*, 95(2), 191-216.
- Sato, E. & Linn, M.C. (2014). Designing critique to improve conceptual understanding. In Polman, J. L., et al. (Eds.). *Learning and becoming in practice: The International Conference of the Learning Sciences (ICLS) 2014*, Vol 1. Boulder, CO: International Society of the Learning Sciences.

Understanding Middle School Teachers' Processing of Student-Generated Resources in Science Classrooms

Xinyang Liu, Shandong Normal University, liuxinyang_sdn@163.com

Shu Chen, Yinglian Jin, Xinning Pei, Tainian Zheng, and Sihan Xiao
cherryshucs@gmail.com, jinyinglian@126.com, xnpei@kcx.ecnu.edu.cn, tnzheng@iice.ecnu.edu.cn,
shxiao@kcx.ecnu.edu.cn
East China Normal University

Abstract: Teachers constantly encounter various responses generated by students in the classroom. These responses, which can be regarded as instructional “resources” that have potential value for facilitating students’ deep learning, are not fully used in teachers’ everyday practice. By proposing the notion of “student-generated resources” (SGRs) and adopting a “resources use” rather than sociolinguistic perspective, this study investigates what SGRs are there in the classroom and how teachers process them. Based on classroom observations, we identify five types of SGRs. Following the “perceive–interpret–mobilize” processing flow, teachers demonstrate three processing results (utilize, feedback, and abandon). Further, we identify three approaches of teachers’ utilization of SGRs and five approaches of their feedback to SGRs. Our findings suggest that teachers tend to give feedback to SGRs rather than utilize them as resources. This study broadens theoretical landscape about classroom resources and helps teachers take full advantage of SGRs in empowering students’ learning.

Introduction

The classroom is a place full of dynamic and unpredictable events, such as students’ answers to teachers’ questions and questions raised by students. Teachers’ processing of students’ responses shapes and features teacher-student interaction as well as instruction and learning in the classroom. The structure and content of teacher-student interaction is the focus of studies following the discourse analysis perspective (Frances, 2002). The famous IRE/F (Initiate-Response-Evaluate/Feedback) framework, first proposed by Mehan (1979) and Sinclair and Coulthard (1975) and revised by Lemke (1990), has dominated this field for a long time (Louca, Zacharia, & Tzialli, 2012). However, IRE/F-based analyses fail to capture the complexity of the discourse with regards to what the students are contributing, most of these studies view teacher’s responses to students as relatively passive reactions. Little research moves forward to trace effects of teacher’s responses and their potential influence on consequential instructional activities.

Our study adopts an alternative perspective that views student-generated events as potential instructional resources, termed here as student-generated resources (SGRs). Specifically, in this study, SGRs include students’ responses (especially non-normative answers) to teachers’ questions as well as students’ ideas, questions or experiment findings expressed in class. We not only focus on how teachers deal with SGRs, but also explore their perceiving and interpreting of SGRs, mobilizing and integrating other related resources with SGRs, and the final utilizing of them. In line with Chin and Osborne (2008), we viewed students’ questions as “a potential resource.” Such resource not only helps students in the learning process, but also serves useful functions as a pedagogical tool for the teacher. The resource perspective expands our vision from teachers’ in-the-moment decision making procedure (Erickson, 2001) to the entire processing flow, providing us with a more systematic view of teachers’ classroom practice.

Our perspective draws on the emerging research field of curriculum (resource) use (Cohen, et al., 2003; Remillard, 2005). The underlying assumption of this field is that teachers are central players in the process of transforming curriculum ideals, captured in the form of mathematical tasks, lesson plans and pedagogical recommendations, into real classroom events (Lloyd, Remillard, & Herbel-Eisenmann, 2009). Over the last decade, studies in this field have grown tremendously, generating several enlightening frameworks (e.g. Brown, 2002; Guedet & Trouche, 2012), enhancing our understandings of how teachers use curriculum resources. Research in this field broadens researchers’ understandings of curriculum materials or resources, not only textbooks, instructional guides and digital learning environments as well as interactions among teacher, student and curriculum are also regarded as an important origin of curriculum resources.

By adopting the resource perspective, we try to bridge the fields of teacher-student discourse interaction and of curriculum (resource) use. Building on the intersection point of the two fields, our study tries to answer the following questions:

1. What SGRs are there in middle school science classrooms?
2. How do teachers process these SGRs?

Theoretical framework

Teacher-student discourse interaction

Following a sociolinguistic perspective (Carlsen, 1991; Frances, 2002), Teacher-student discourse interaction studies has stressed the importance of classroom discourse for teaching and learning. The main approaches for studying teacher discourse followed primarily the Initiation–Response–Evaluation (IRE) or Initiation–Response–Followup or Feedback (IRF) structure (Lemke, 1990). These approaches have received criticism, especially regarding the focus of research approaches in using IRE exclusively on a teacher’s role in initiating and maintaining a conversation. Research needs to provide a more coherent framework for studying teacher discourse in relation to student discourse (Aguilar, Mortimer, & Scott, 2010; Cazden, 2001; van Zee, Iwasyk, Kurose, Simpson, & Wild, 2001).

Studies guided by other perspectives have broadened the vision of classroom discourse. Scott (1998) defined the notion of “teacher responsiveness” including three elements: 1) monitoring (monitor the present performance of students), 2) analyzing (analyze the nature of any differences between present performance and the target performance), and 3) assisting (respond with an appropriate intervention to support students).

Louca, Zacharia, and Tzialli (2012) argued that it might be more useful for the research community to have a framework that focuses on student contributions during the conversation and the teacher decisions and responses based on those contributions. Therefore, they developed a framework for investigating classroom discourse by focusing more broadly on what the teacher responds to, how she responds, as well as the process of deciding how to respond. This framework includes *Identification*, *Interpretation—Evaluation*, and *Response*. The Identification part concerns what the teacher responds to (student discourse contributions), the *Interpretation—Evaluation* part concerns how the teacher *interprets* and *evaluates* students’ discourse contributions, while the Response part concerns how the teacher responds to students’ discourse contributions.

Teachers’ use of curriculum (resource)

For researchers in this field, “use” means a variety of interrelated pedagogical activities, including how teachers engage or interact with these resources as well as how and the extent to which they rely on them in planning and enacting instruction, and the role resources play in teachers’ practice. (Lloyd, Remillard, & Herbel-Eisenmann, 2009)

Research on curriculum (resource) use (e.g. Cohen, et al., 2003; Remillard, 2005; Gueudet & Trouche, 2012; Jin, et al., 2015) proposes that teachers *perceive* and *interpret* what happened in the classroom before they draw upon resources in the subsequent instruction. Hammer (1997) argued that teachers should coordinate relevant resources to achieve instructional goals before they have the opportunity to implement instruction. Brown (2002) proposed the notion of “pedagogical design capacity” to characterize teachers’ capacity to perceive, interpret and mobilize curriculum resources in the pursuit of desired outcomes. Adler (2000, 2012) proposed to turn the noun word resources into a verb “re-sources”, emphasizing the nature of teacher’s use of curriculum materials is a reconstruction procedure. Similarly, Gududet and Trouche (2012) also pointed out that through a procedure of what they called instrumentalization, teachers transform external resources (e.g. textbook, software, unpredictable student responses) into “lived” resources applicable in their own classroom contexts.

A framework for analyzing teachers’ processing of SGRs

By reviewing literature, we found an intersection point of the above two fields: they both have interests in teachers’ processing flow of SGRs. Viewing students’ responses as instructional resources, we constructed our framework for analysis of teachers’ processing of SGRs by integrating the key procedures derived from widely used curriculum (resource) use research (e.g. Brown, 2002) and teacher-student discourse interaction research (e.g. Louca et al., 2012), as shown in Figure 1. The verbs “perceive-interpret-mobilize” in rectangles defined three key steps of teachers processing of SGRs. In curriculum (resource) use research, “perceive” means teachers’ awareness of the resource that has potential value for instruction (Brown, 2002). Louca et al named the similar step that teachers noticed students’ verbal responses as identification (Louca et al., 2012). “Interpret” refers to teachers’ evaluation of SGRs and current contexts so as to decide whether or not to use SGRs in the next step. If teachers decide to use a certain piece of resource, they need to choose and integrate other related resources. This step is named “mobilize”. For instance, the teacher refers to his/her PCK to give a comment to one student’s response to an open-ended question or gives a clear explanation to one student’s question about the experiment

design. As a result, at the end of the processing procedure flow, there are three different types of processing results, labeled as *utilize*, *feedback* and *abandon*. Specifically, our framework illustrated branches (not perceived, not decided to use) in the processing flow and three types of processing results (utilize, feedback and abandon), these details were mentioned or implied in the two fields but not thoroughly discussed. We hope to promote understanding of these details with our framework.

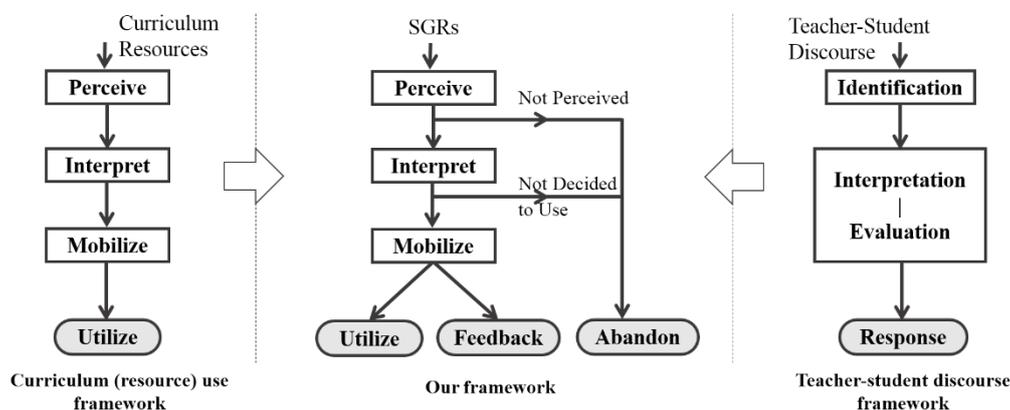


Figure 1. A Framework for analyzing Teachers' Processing of SGRs.

Methods

Participants and context

This study took place in a middle school located in downtown Shanghai. Two experienced teachers were purposefully selected from 4 science teachers in that school according to criteria proposed by Berliner (1994, 2004) they had taught science in school for more than 7 years, and 2) their achievement in science teaching had been recognized by both principals and colleges. We chose experienced teachers in our study because they have well-developed professional competences and more stable teaching behaviors, which would enable us to capture the patterns and features of teachers' processing of SGRs. Teacher A had 19 years teaching experience while Teacher B had 8. They taught science for 7th graders and 6th graders respectively.

According to *Shanghai Middle School Science Curriculum Standards*, Science, as a subject, is designed for students in grade 6 and 7 and aims at developing students' comprehensive science understanding and knowledge before taking physics, chemistry, life science and geography in grade 8. After discussing with both teachers, we finally chose "photosynthesis and respiration" for Teacher A and "features of living things" for Teacher B. Both teachers said they had interest and confidence for teaching their own topic. This context was helpful for us to capture abundant SGRs and explore teachers' sophisticated processing of them. Each topic included 45 minutes of classroom instruction and 45 minutes of laboratory experiment, and was taught by one teacher in two classes respectively. Students in all the classes had similar performances and achievements in science learning.

Data sources

Eight lessons (each topic included two lessons, one teacher taught for two classes repeatedly) were recorded using video camera. Immediately following each lesson, short interview (3-5 minutes) were conducted and recorded to help us verify some ideas about teachers' teaching behaviors we observed. All of the video and audio records were transcribed. We used Nvivo 10 to manage data and conduct analysis.

Data analysis

Categories of SGRs

We first viewed all the lesson videos and identified every piece of SGR according to the definition mentioned above. Then by using open coding method (Glaser & Strauss, 1967), we labeled and categorized each piece of student-generated resource. We then looked through all labels and verified their fitness. Finally, we adjusted and refined the following list (see Table 1) to develop codes for categorizing SGRs.

Table 1: Categories of SGRs

Category	Description	Typical Example
NR-Close	Non-normative Response to teacher raised Close-ended question	T: Where does respiration take place? S: In our lungs.
NR-Open	Non-normative Response to teacher raised Open-ended question	T: Do you know other common features of animals? S: Animals live together.
RR-Open	Reasonable Response to teacher raised Open-ended question	T: Give me an example of animal's adaption to environment. S: Bears hibernate in winter
SQ	Student raised Question	A student asked Teacher A what is the black thing in the bottom of the test tube.
SE	Student Experiment phenomenon or problem	A student reported his snail showed no reaction to light.

*As correct answers to close-ended questions are predictable, this kind of student responses was not considered as *generated* resource in our study.

Procedure of teacher's processing SGRs

The theoretical framework mentioned above served as a reference for analyzing teacher's processing procedure of student-generated resources. Meanwhile, we kept open-minded and sensitive to any evidence emerged from the data that might bring revision to the theoretical framework. After a similar process mentioned above, we developed codes to identify different processing stages. (Table 2)

Table 2: Stages of teachers' processing SGRs

Stage	Description	Typical Example
Perceived	The teacher noticed a piece of student-generated resource.	When discussing "whether coral is animal or plant", some students said "animal" others said "plant", We observed that both were noticed by Teacher B.
Not Perceived	We noticed a piece of resource but no evidence showed that the teacher noticed it.	(not occurred in the lessons we observed)
Decided to use	The teacher showed response to a piece of resource she noticed.	Teacher B asked students who said "coral is animal" to explain their reasons.
Not decided to use	The teacher noticed a piece of resource but showed no response to it.	Teacher B noticed some students said "coral is plant", but never asked them to explain.
Utilized	The teacher made use of the resource in the following instruction.	A student asked "why should we first put the plant in a dark room for 24 hours", then Teacher A used it as a question to ask the class.
Feedback	The teacher only gave the student(s) some kind of feedback	When a student replied "chlorophyll" for the question on the products of photosynthesis, Teacher A said "No, chlorophyll is a kind of matter of the plant itself."

Results

Categories and frequencies of SGRs

According to the codes of SGR categories, we finally identified 36 pieces of SGRs in Teacher A's classes and 26 pieces of SGRs in Teacher B's classes. (Considering the same lesson repeated in different classes, SGRs observed in different classes with both same content and teacher processing mode are counted as one single resource).

Figure 2 shows the frequencies of different types of resources generated by students in science class. Overall, the majority of SGRs (Teacher A: 89%, Teacher B: 88%) are students' responses to teachers' questions (including NR-Close, NR-Open, RR-Open). In Teacher A's class, we identified all five types of SGRs (36 pieces), of which up to two-thirds were NR-Close type (22 pieces), while in Teacher B's class, there's no SQs, RR-Open type occupies 50% (13 pieces) of all SGRs in her class.

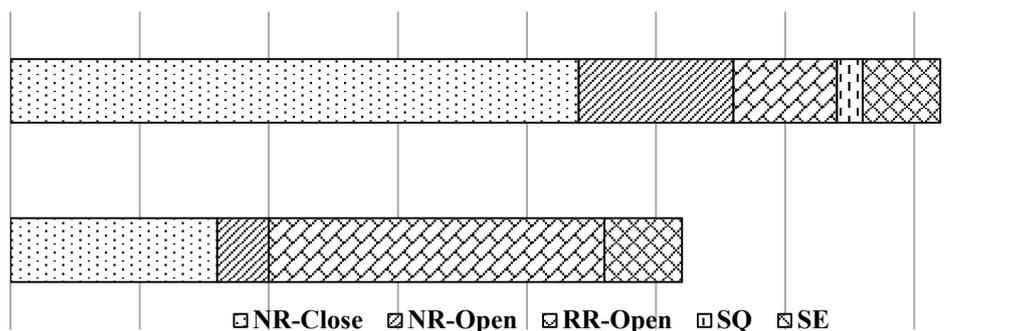


Figure 2. Frequencies of different categories of SGRs in two teachers' classes.

Procedure and results of teachers' processing SGRs

By coding teachers processing stages of SGRs, we found teachers basically followed the "perceive-interpret-mobilize" procedure and both teachers could perceive and interpret all the SGRs. In terms of the processing results, teachers have similar preferences. They abandoned only two pieces of SGRs each, all the four pieces of SGRs belong to NR-Close category. As shown in Figure 3, when science teachers successfully incorporated other resources to mobilize the SGRs, they might not *utilize* them. Rather, they were more likely to give *feedback* for the SGR's own sake. For instance, a teacher might directly correct students' non-normative answers or only tell students whether their answers are right. Both Teacher A and Teacher B gave lots of feedback to SGRs in their science class, accounting for 61.1% and 69.2% of all the processing results respectively. However, Teacher A utilized one third of SGRs in the class and Teacher B utilized less than one quarter of all the SGRs. (see Figure 3, width of the arrow lines indicates quantity of SGRs in the branches)

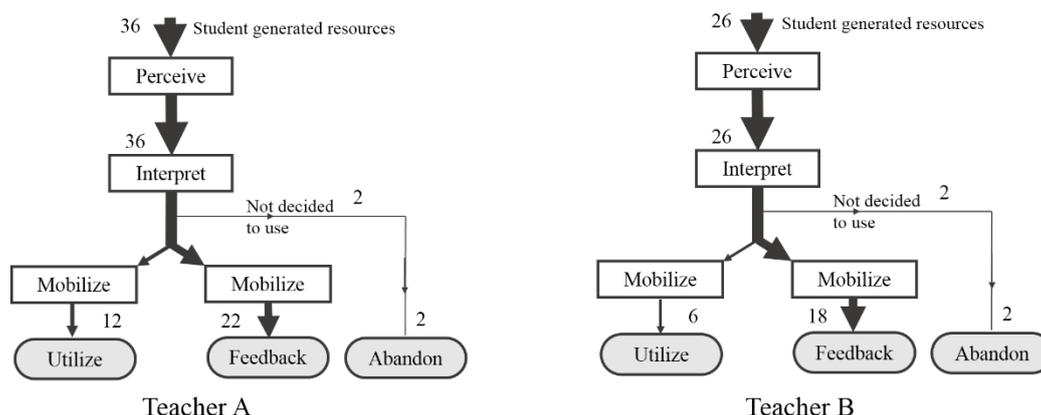


Figure 3. Teachers' processing flow of SGRs.

Distribution of utilization and feedback among five categories of SGRs

In general, teachers *decided to use* the majority of the SGRs regardless of their categories. Figure 4 illustrates the distribution of utilization and feedback among five different types of SGRs. We found that the two teachers *utilized* three of the five categories of SGRs: NR-Close, RR-Open and SQ. Both teachers could *utilize* NR-Close type of SGRs, of all the SGRs utilized by each teacher, NR-Close type of SGRs took 92% and 67% respectively. In terms of each type of SGRs, both teachers utilized more than one half of NR-Close type of SGRs in their class. In the contrast, neither NR-Open nor SE type of SGRs were utilized by two teachers.

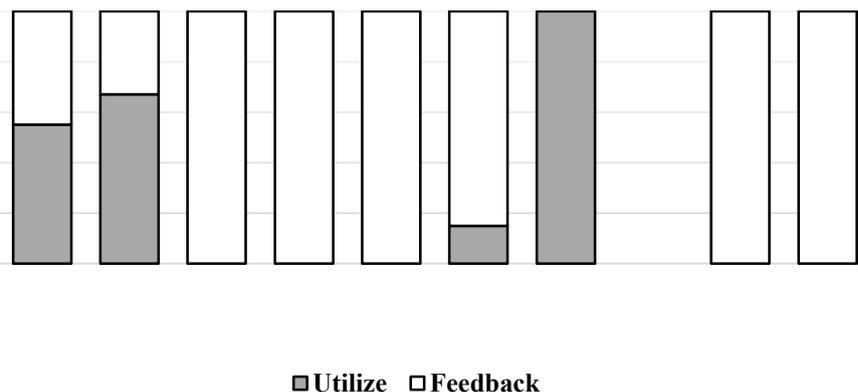


Figure 4. Distributions of utilization and feedback among five categories of SGRs.

Teachers' instructional approaches of utilization and feedback

When teachers *utilized* SGRs, they demonstrated 3 types of approaches: 1) *Testing* students' knowledge. For example, a teacher might repeat or revoice a student's answer and then ask the class "Do you agree on his/her idea?" or "Who can tell us what's wrong with his/her answer?". Teacher A used about 50% of the utilized SGRs to test students' mastery of knowledge, while for Teacher B, the proportion was 66.7%. 2) *Creating problem contexts*. For instance, a student proposed a question to Teacher A, "Why we should use alcohol to dissolve and remove chlorophyll in leaves?" Then Teacher A used this question as a trigger to engage all the students in thinking about the purpose of this key step in the experiment. Teacher A adopted this approach to process 42% of the utilized SGRs and Teacher B processed 33% of the utilized SGRs in his/her class. 3) *Extending content* in the textbooks such as introducing the concept of carbohydrate one the basis of students' responses. Two pieces of SGRs were utilized by Teacher A for content extension, while none utilized by Teacher B.

In terms of *giving feedback* to SGRs, both teachers had five specific ways to process them: 1) *Judge*, which means telling the truth of false was the most common way adopted by both teachers. In Teacher B's feedback to SGRs, 56% of them were processed in this way. For Teacher A, the proportion was 35%. 2) *Query*, referring to pointing out problems in students' responses. For example, a teacher asked the students who reported an experiment design that "could you guarantee that your design would prevent air from entering into the test tube?" Teacher A gave query to 1/3 SGRs received feedback, and Teacher B only processed one piece of SGRs in this way. 3) *Redirect* students' responses to the topic they were talking about. For example, one student argued that running away in the face of danger was a common feature of animals, and then teacher told her that "this is animals' adaptability to surroundings". Only one piece of SGR was redirected in Teacher A's class, while in Teacher B's class, 28% feedback were given by this approach. 4) *Correct*, that is, teachers correct mistakes or inaccuracy in students' responses. 5) *Explain*, namely give explanations to problems students encountered in doing their experiment. These two approaches were used seldom by both teachers.

Discussion and conclusion

This study shows that science classes were short of student initiated SGRs that might be of highly value for teaching utilization. The majority of SGRs were students' responses to teachers' questions (including NR-Close, NR-Open, RR-Open), whereas students initiated questions (including SQ and SE) were much less. This illustrates that students did not have sufficient opportunities or get necessary scaffolding to generate their own questions. This finding resonates with the results of many studies reviewed by Aguiar et al. (2010), which suggest that students' questioning is both infrequent and unsophisticated in the science classroom. Among students' responses, NR-Close type of SGRs took 56%. Though RR-Open SGRs occupy the second largest portion, most of those "open questions" can be characterized as directives, such as "Please give us an example of ..." or "Tell me a phenomena in your everyday life that can illustrates ..." which could be attribute to what Scott (1998) defined as "authoritative discourse", both of the teachers admitted that they made efforts to control students' utterance "on the expected way" to textbook topics. In that case, science teachers should be prompted to be less dominated in the class and give students more chances to propose their own questions that could be utilized as precious resources for enhancing students' science learning.

The analysis of teachers' processing procedure and results of SGRs reveals that the framework for analyzing teacher-student interaction in the curriculum (use) perspective could be available and practical for researchers. In our study, both teachers were able to notice and perceive all the SGRs, and they would

automatically mobilize other resources they owned to utilize or give feedback to these SGRs. Besides, as we paid more attention to how teachers' responses influenced students' learning rather than the formal pattern itself, we identified specific instructional approaches teachers adopted in utilizing or giving feedback to SGRs. When teachers decided to utilize SGRs, they would flexibly use them to test, create problems situations or extend contents in the textbooks, which showed their general awareness of valuable SGRs and competences in processing SGRs. As classes moved, we found that students' understanding of science concepts or phenomena was enhanced. As a result, teachers' practical wisdom could contribute to further academic researches on the usage of SGRs, and then maximize the value of SGRs for students' science learning and understanding.

Our study also illustrates teachers' disabilities and weakness in dealing with these in-the-moment events. These would help us give more effective and targeted suggestions for teachers to optimize their processing of SGRs, thus empower students' science learning. The distribution of teachers' processing results among five categories of SGRs shows that most of the SGRs utilized by teachers were NR-Close type, while NR-Open, RR-Open and SE types of SGRs were not fully employed. It could be inferred that NR-Close type of resources were much easier for teachers' recognition and utilization with the least variability and uncertainty among all five types of SGRs. Therefore, we should develop teachers' abilities to address SGRs with high uncertainty and complexity.

In our view, giving feedback which was the most common processing result of SGRs, could also benefit students' learning. The key point lies in specific ways teachers adopt to approach the resources. In our analysis, the five types of feedback we identified and named in this study are similar with four different types of teacher feedback (Chin, 2006). The analysis suggests that when teachers gave feedback to SGRs, they mostly judged the answers and pointed out directly whether the answers were right or wrong. While query regarded as higher-order cognitive feedback (Chin, 2006) took less than one quarter of all the pieces of SGRs. The effectiveness of feedback to SGRs was reduced because of the low-order cognitive ways teachers adopted. In this sense, we should facilitate teachers to give more high-order cognitive feedback in order to facilitate students' higher-order thinking.

In this study, we proposed the notion of "student-generated-resources" (SGRs). We argued that SGRs have much potential value for enhancing students' science learning if it can be made full use of by science teachers. We identified five types of SGRs and constructed a theoretical framework which links researches on curriculum (resource) use and teacher-student discourse interaction for analyzing teachers' processing of SGRs. In particular, we focused on teachers' processing results of different types of SGRs and their specific ways to utilize and give feedback to SGRs in the class. These findings helped us to figure out characteristics of teachers' processing of SGRs and thus give practical suggestions on teachers' instruction which may empower students' learning in science class.

References

- Adler, J. (2000). Conceptualising resources as a theme for teacher education. *Journal of Mathematics Teacher Education*, 3(3), 205-224.
- Adler, J. (2012). Knowledge resources in and for school mathematics teaching. In Ghislaine, G., Pepin, B. & Trouche, L. (Eds.), *From Text to 'Lived' Resources: Mathematics Curriculum and Teacher Development* (pp. 3-22). Netherlands: Springer.
- Aguiar, O. G., Mortimer, E. F., Scott, P. (2010). Learning From and Responding to Students' Questions: The Authoritative and Dialogic Tension. *Journal of Research in Science Teaching*, 47(2), 174-193.
- Berliner, D.C. (1994). Expertise: The wonder of exemplary performances. In Mangieri, J.N. & Block, C.C. (Eds.), *Creating powerful thinking in teachers and students* (pp. 161-186). Fort Worth, TX: Holt, Rinehart & Winston.
- Berliner, D.C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24(3), 200-212.
- Brown, W. M. (2002). *Teaching by Design: Understanding the Interaction between Teacher Practice and the Design of Curricular Innovations* (Unpublished doctoral dissertation). Northwestern University, Evanston, IL.
- Carlsen, W.S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research*, 61(2), 157-178.
- Cazden, C.B. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.
- Chin, C. (2006). Classroom Interaction in Science: Teacher questioning and feedback to students' responses. *International Journal of Science Education*, 28(11), 1315-1346.
- Chin, C., Osborne, J. (2008). Students' questions: a potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1-39.

- Cohen, D.K., Raudenbush, S.W., & Ball, D.L. (2003). Resources, Instruction and Research. *Educational Evaluation and Policy Analysis*, 25(2), 119-142.
- Erickson, F. (2011). Accessing mathematics teachers-in-the-moment noticing. In Sherin, M. G., Jacobs, V.R., & Philipp R. A. (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 79-94). New York and London: Routledge.
- Frances, C. (2002). *Classroom discourse analysis*. London: BookEns Ltd.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. New York: Aldine.
- Gueudet, G & Trouche, L. (2012). Teacher's Work with Resources: Documentational Geneses and Professional Geneses. In Ghislaine, G., Pepin, B. & Trouche, L. (Eds.), *From Text to 'Lived' Resources: Mathematics Curriculum and Teacher Development* (pp. 23-41). Netherlands: Springer.
- Hammer, D. (1997). Discovery learning and discovery teaching. *Cognition and Instruction*, 15, 485-529.
- Jin, Y., Liu, X., & Pei, X. (2015, April). Science Teachers' Utilization of Teaching Resources in Junior High School in Shanghai. *Paper presented at the annual meetings of the American Educational Research Association*. Chicago, USA.
- Lemke, J.L. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex.
- Lloyd, G. M., Remillard, J. T., Herbel-Eisenmann, B. A. (2009). Teachers' Use of Curriculum Materials: An Emerging Field. In J. T. Remillard, B. A. Herbel-Eisenmann & G. M. Lloyd (Eds.), *Mathematics Teachers at Work: Connecting Curriculum Materials and Classroom Instruction* (pp. 3-14). New York & London: Routledge.
- Louca, L. T., Zacharia, Z. C., Tzialli, D. (2012). Identification, Interpretation—Evaluation, Response: An alternative framework for analyzing teacher discourse in science. *International Journal of Science Education*, 34(12), 1823-1856.
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.
- Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75 (2), 211-246.
- Scott, P. (1998). Teacher Talk and Meaning Making in Science Classrooms: a Vygotskian Analysis and Review. *Studies in Science Education*, 32(1), 45-80.
- Sinclair, J.McH., & Coulthard, R.M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press.
- van Zee, E.H., Iwasyk, M., Kurose, A., Simpson, D., & Wild, J. (2001). Student and teacher questioning during conversations about science. *Journal of Research in Science Teaching*, 38(2), 159-190.

Analyzing Patterns of Emerging Understanding and Misunderstanding in Collaborative Science Learning: A Method for Unpacking Critical Turning Points

Alisha Portolese, Lina Markauskaite, Polly K. Lai, and Michael J. Jacobson
Alisha.Portolese@sydney.edu.au, Lina.Markauskaite@sydney.edu.au, Kuan.Lai@sydney.edu.au,
Michael.Jacobson@sydney.edu.au
The University of Sydney

Abstract: When students learn science in a computer-supported, collaborative, delayed-instruction environment, how does understanding (and misunderstanding) emerge? Are there patterns in the pivotal moments when emerging understanding turns for the better or worse? While components such as modeling software, delayed instruction methods such as productive failure, and analogical-encoding methods such as contrasting cases have all been shown effective at supporting deep learning in science, little is known about the micro-level mechanisms explaining how and why students might be more or less successful when working in an environment combining all three. This paper details our refinements of an innovative method for unpacking the micro-level mechanisms contributing to turning points in the successes and failures in collaborative understanding when learning science with computer modeling. In unpacking our methodology, we discuss work including Sanderson and Fisher's (1994) exploratory sequential data analysis (ESDA) guidelines and the productive multivocality project (Suthers, 2013) to frame our approach.

Keywords: conceptual change, collaborative learning, science education, turning points, pivotal moments, delayed instruction

Introduction

When learning science, *experiencing* science and engaging in collaborative, authentic, and active learning can be highly beneficial for deep learning (Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001). Providing a space for domain-grounded discussions is particularly beneficial for students to activate prior knowledge, become aware of existing beliefs, realize the extent of their current understanding, and discuss emerging understanding (Greeno, Collins, & Resnick, 1992; Roschelle, 1992; Vosniadou et al., 2001). Additionally, learning environments are strong when they include problems that resemble the complexity inherent in authentic, real-world situations (Jacobson, 2000; Kapur & Kinzer, 2007). Environments with this type of authentic depth foster the rich discussions that support the development of deep learning and transfer.

One method of sculpting a learning environment with these features involves incorporating computer modeling. Models provide the opportunity to manipulate and interpret complex relationships as they happen, and can serve as a visual tool to ground abstract concepts (Nersessian, 2008). Students can manipulate a shared computer model in small groups (collaboration), which can be situated in a delayed instruction learning sequence. For example, in productive failure, students begin by collaboratively exploring possible solutions to a complex, authentic problem they have not yet been taught how to solve. This is followed by a teacher-led consolidation highlighting the critical solution features and comparing and contrasting student and expert solutions (Kapur & Bielaczyc, 2012). Collaborative problem solving with computer modeling software and contrasting cases can be effectively included as part of productive failure's initial idea generation and exploration phase (Jacobson & Markauskaite, 2015, April; Portolese, Markauskaite, Lai, & Jacobson, 2015, April). The effectiveness of contrasting cases is grounded in analogical encoding theory (Gentner, Loewenstein, & Thompson, 2003), which proposes that explicit comparison of multiple cases with different surface features but similar underlying principles can enhance learning the critical features of the core concept.

Despite strong evidence for promoting deep learning and transfer with modeling, delayed instruction, and contrasting cases, the mechanisms underlying how and why these processes work, particularly with all combined, remains relatively unknown. One thread of our research has examined some important questions – when learning in this way, how does collaborative understanding and misunderstanding emerge? And how specifically does understanding take a turn for the better or worse? From a sociocultural perspective, knowledge is co-constructed through interactions (Säljö, 1991). Knowledge building interactions are productive when they allow students to build partial meanings that are gradually refined towards increasingly expert understanding

(Damşa, 2014). Therefore, the micro-level interactional mechanisms fueling emerging knowledge co-construction are a critical yet little understood aspect of deep, collaborative learning. Few studies have explicitly addressed how emerging knowledge relates to interaction over time (Kapur & Kinzer, 2007; Damşa, 2014).

This paper aims to respond by contributing to theory and methodology by providing a detailed, multi-layered, temporally-sensitive approach for analyses designed to unveil the micro-level mechanisms of conceptual change and emerging understanding in complex, rich, computer-supported collaborative learning environments. In preliminary analyses, we explored students' collaborative learning at three parallel grain sizes using an impact coding approach (Portolese et al., 2015, April). With this, we found that superior performance on a far transfer item may have been associated with an idea generation process characterized by producing substantially more ideas, particularly more correct suggestions, predictions, experimental questions, experimental designs, and explanations. We also found that students' collaborative processes seemed to be characterized by small segments of misunderstanding propelling extended correct understanding. However, it remained to be explored what exactly occurred at these critical turning points in understanding; a deeper investigation was required. As we will present in this paper, we expanded our analysis technique to identify and explore turning points in understanding in a rich and meaningful way. We briefly presented our early ideas for how these turning points could be unpacked in a poster at The Computer Supported Collaborative Learning Conference (Portolese, Markauskaite, Lai, & Jacobson, 2015, June). With the support of peer feedback, we present here a refined expansion of this methodology for unpacking turning points in substantially more detail. This detailed explanation of our exploration and conceptualization of turning points is an important contribution to understanding how the mechanisms of emerging understanding and misunderstanding in model-based collaborative learning might be revealed.

In line with the conference's thematic strands, our aim is to unpack the micro-level mechanisms underlying conceptual change and knowledge construction in a science learning, computer-mediated collaborative environment. We believe our main contribution is the detail of an innovative method for unpacking the micro-level mechanisms of turning points in the successes and failures in collaborative understanding when learning science with modeling software. In addition, we present the patterns of emerging understanding and misunderstanding from our data, including insights regarding which aspects might be more and less "productive" to include or withhold scaffolding. While there has been much recent activity regarding such issues (see Kirschner, Sweller, & Clark, 2006; Kapur & Bielaczyc, 2012; Jacobson, Kim, Pathak, & Zhang, 2013) at the larger learning design level, much work is still needed regarding unpacking the mechanisms within powerful learning designs (see Loibl & Rummel, 2014, for a productive move in this direction).

Methods

Context and participants

The two dyads chosen for detailed analyses were selected from a larger study conducted across four Year Nine science classes at a selective girls high school in Australia (see Jacobson & Markauskaite, 2015, April; Portolese et al., 2015, April). Dyads worked collaboratively on inquiry activities that required experimentation with NetLogo (Wilensky, 1999) models. The students from the two selected dyads improved substantially on the target complexity concepts from pretest to posttest within their own groups, in line with the overall results for each group (see for details Jacobson & Markauskaite, 2015, April; Portolese et al., 2015, April).

Details of our coding approach, Phase 1: Impact coding at three-parallel grain sizes

The first phase of our analysis involved exploratory coding at three parallel grain sizes. The approach was developed from our understanding of the data (not a pre-defined model). The incorporation of impact coding to allow for context and time-sensitivity was based on Kapur, Voiklis, and Kinzer's (2008) method; see Portolese et al. (2015, April). As Kapur et al. (2008) argue, this method preserves the temporal sensitivity that is critical in understanding emerging understanding (or misunderstanding). With this, data segments were coded as +1 (moving towards the solution), -1 (moving away from the solution), or 0 (not changing progress). In addition to numeric codes, we also used descriptive labels to add richness to our interpretation.

The micro grain size was the idea level. We defined ideas as a single train of thought (one or multiple speakers) and associated computer actions (e.g. model manipulations). In addition to impact coding, ten categorical labels emerged: the first three about process and the remaining seven about content (see Table 1, and see Portolese et al., 2015, April, for additional details and examples). The content categories that emerged could be understood as cycles of the scientific method (see Figure 1). In summary, each micro segment was assigned a category label and numeric code.

Table 1: Summary of Micro Level idea category descriptions

Idea category	Description
Task	Orienting to the task, understanding instructions
Technical	Technical issues often related to the NetLogo modeling software
Representation	What an element of the model represents, and/or how various parts of the model are related to each other (e.g. simulation interface vs. graphical representations)
Suggestion	Novel contributions or solution directions
Experimental Question	Inquiries related to or leading to student experiments
Prediction	Guesses for the outcome of a current or upcoming experiment
Experimental Design	Planning and execution of a modeling experiment
Observation & Data Collection	What the students see in the model (as made overt)
Explanation	How students interpret their observations and experiments
Understanding	Indication of broader comprehension (or miscomprehension) of the target concept

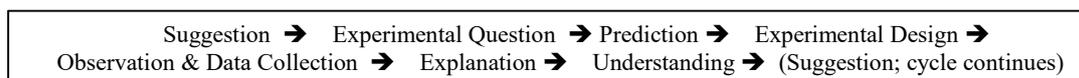


Figure 1. The seven content based idea-level categories can be understood as the cycle of the scientific method.

The meso grain size was about change in understanding the epistemic task. Generally, one or multiple idea-level moves indicated this progression of a new phase of understanding. For example, if a student made a prediction and ran an experiment with particular parameters, this could represent their preconception of the concept at hand. Then, during the experiment, students might make observations and explanations, and at this point students might *change* their understanding due to the events they are observing. As students discuss and elaborate ideas, understanding might *change again*. This example broadly outlines a progression of three meso-level segments. In addition to the impact coding, a non-categorical description was provided (see Table 2).

The macro grain size was the experimental level. Students were given guiding questions to softly scaffold their interaction with the models. Working through the questions, students manipulated the model, running simulation experiments. We segmented each macro grain size between when students planned an experiment to when students completed related conclusions. A prototype would include idea segments representing an entire cycle of the process in Figure 1. However, since this was a minimally guided activity with novices, students often did not follow this prototypical method – sometimes they made ad hoc choices with the modeling parameters, or did not make overt observations or discuss their inferences. In these instances, we created segment boundaries based on indicator activities such as manipulating model parameters and refreshing the model, as these actions demonstrate the intention of creating a new experiment. As with the meso level, both a numeric impact code and a non-categorical description was associated with each segment. See Portolese et al. (2015, April) for examples of coding at this level. Importantly, the impact code was assigned based on the position students demonstrated they were in at the end of the segment – as macro segments could contain diverse meso segments within it, understanding could be turbulent and changing throughout the segment.

The three grain sizes were coded somewhat in parallel – the broader context of students’ activity regarding their experiments helped make sense of their actions and words at the micro and meso levels subsumed within in (see Figure 2). With the broader context of a macro segment generally understood, the grain sizes were then coded from the micro level segments until reaching the end of a meso segment, and then moving onto the following meso chunk and starting again with the micro segments within it. Labels and descriptions were usually coded before the numeric impact code. Multiple time-aligned factors in our rich data were taken into account, including computer actions, words, and written workbook answers (integrated as they occurred in time). Students’ words were considered regarding tone of voice (e.g. sarcasm), implications (utilizing context to infer likely meaning when possible), and focus of attention (e.g. cues from eyes and classroom events). Other relevant events such as interactions with nearby groups or with researchers and teachers were also considered and coded. Overall, a simple yet effective way to help determine how to code a segment at any level was asking ourselves, “At the end of this segment, has the students’ understanding changed? If so, in what way?”

Table 2: Meso Level Coding

Impact Level	Definition	Example	Example Description
-1	New or reinforced misunderstanding	“Okay, maybe make it equal” • Computer: wolf-reproduce #4 • “Make this four” • Computer: wolf-gain from food #4 • “But we’re gonna keep that number the same, ok?” • “Mhmm” • Computer: Go (start) (no response) • “Oh wait, no. Set up” • Computer: Set up, Go (start)	The students incorrectly thought that manipulating the parameters “wolf gain from food” and “wolf reproduce” so that the values were equal would make the model self sustaining.
0	Do not understand what is happening or no overt change in understanding	Computer: Go (start) • “Wow okay” • “Gosh” • “Where’s the discussion?” • “Wait – where are the wolves?” • “Oh wow, oh gosh” • “Woah” • “Oh wow. That, that’s horrible. Stop. How do I stop it?” • “Woah” • “What just happened?” • “I don’t know, but the population of sheep is just going up.”	At the beginning of their work, the students do not yet understand what is happening with the model.
+1	New or reinforced understanding	Computer: Go (stop) • “How do you stop it?” • Computer: Go (start) • “Oh there we go... okay yeah it’s still going” • Computer Go (stop) • “ok, um”	The students learn how to stop the model.

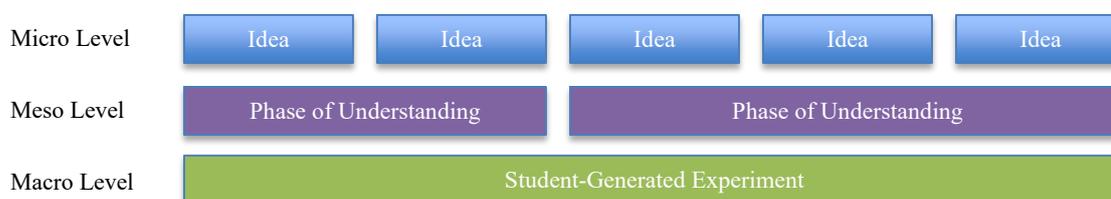


Figure 2. How the three parallel grain sizes fit together.

Details of our coding approach, Phase 2: Turning points analysis

As discussed above, we found that the initial phase of analysis provided useful insights, however, we felt the need to go deeper to truly understand why understanding changed, and what (if any) patterns existed in what we were observing. The impact coding afforded us the possibility to graph a group’s cumulative, unfolding understanding and misunderstanding. There were many possible ways this could be done for each group – at each of the grain sizes, and within the micro category using a filter for one or some categories. In our meso level, which we determined focused at the core of what was happening and changing with understanding, we were particularly interested to delve deeper to understand points where understanding seemed to turn. We defined turning points as critical moments in the development of collaborative understanding when understanding changed in an incremental way. The numeric impact coding allowed for an opportunity to provide a clear boundary for identifying turning points; we operationalized turning points as when the impact direction changed and continued for at least two segments. Positive turning points were changes from misunderstanding towards understanding (e.g. -1, +1, +1) and negative turning points were changes from understanding towards misunderstanding (e.g. +1, -1, -1; Portolese et al., 2015, June; see Figure 3).

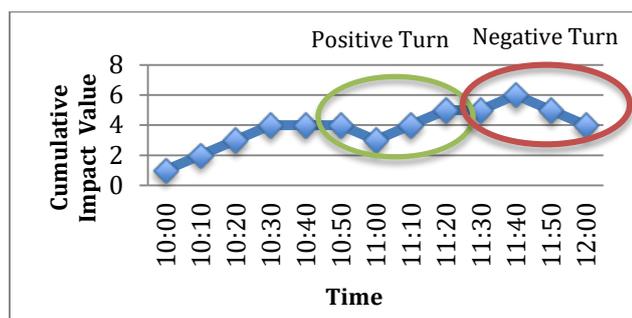


Figure 3. A prototype of how the turning points could be visualized graphically.

In order to understand these turning points, we zoomed into the micro segments in two locations: a) the meso level segment before understanding turned and b) the subsequent meso level segments during the change until two segments in the new direction occurred. These were not always the immediately subsequent segments as sometimes neutral segments spaced in between, such as -1, +1, 0, +1. For example, using the positive turn in Figure 3, the analysis for part (a) would unpack the ideas in the meso segment at time 10:50-11:00, and the analysis for part (b) would be between 11:00-11:20. Table 3 is an example of this breakdown from a turning point in our data. Following this, turning points were interpreted and grouped based on patterns of understanding and misunderstanding which emerged based on a) what happened before and b) what happened during the change in understanding. The analysis aimed to identify the nature of the events that caused the change, to provide us with tangible information on the development of the group processes and understanding.

Table 3. Example from our data of unpacking and analyzing a turning point

Turn Direction	(a) Ideas in Meso Segment Before Turn	(b) Ideas in Meso Segment (i) During Turn		(b) Ideas in Meso Segment (ii) During Turn
Positive	Experimental Design (-1)	Observation (+1)	Technical (+1)	Experimental Design (+1)

Findings and discussion

Patterns of turning points in our data

In total, we identified and analyzed 26 turning points across the two dyads (17 positive, 9 negative). See Tables 4 and 5 for a summary of our turning points analyses. We grouped the turning points based on thematic patterns that emerged. Missing the bigger picture was a common theme, both as a precursor to positive turning points and as the substance of negative turning points. This type of problem can be elusive, because due to the nature of the problem, students likely do not realize there is an issue. Students' understanding turned for the better through additional observations, experimentation and elaborated discussions. Similarly, making incorrect observations was common, both as a precursor to positive turning points and as the substance of negative turning points. Additional correct observations helped students re-ground their developing understanding in correct ideas. We found one instance of students' confusion on a conceptual level as a turning point – engaging in the experimental cycle of ideas rebuilt understanding from the ground up. Similarly, misunderstanding caused by poor experimental designs turned via focused predictions that improved subsequent experimental designs. A less productive pattern was students experiencing technical or representational confusion or errors, which unfortunately in some cases led to deeper misunderstanding – this type of floundering did not appear productive for deep understanding. We would lastly like to highlight a group of negative turning points (last row in Table 5) as examples of instances where basic declarative understanding, even when correct, does not necessarily indicate or lead to deeper conceptual understanding. As seen in our cases, this was even found when the students were largely doing all “right” things regarding following instructions and using a good experimental procedure.

Table 4: Positive turning points turning towards understanding (moving towards solution)

What went wrong: Pre turning point	Turning point frequencies	Key micro level characteristics during turn	How understanding turned & emerged
Conceptual confusion	1	Correct suggestions, experimental design, and observation	Cycles of suggestion, experiments, and observation required to rebuild understanding from the ground up
Technical or representational errors	1	Task ideas (reorientation)	Moved along despite unresolved technical challenges or representational misunderstandings
	2	Correct observation, explanation, and representation ideas	Extended observations and discussions about technical and representational aspects
Incorrect observations	4	Correct observations; Correct suggestions (for some)	New observations corrected misconceived observations
Poor experimental design	4	Correct predictions and experimental design	Predictions typically fueled improved experimental designs

Missing bigger picture	5	Correct observations, explanations, understanding, and experimental design	Additional experimentation and elaborated discussion
------------------------	---	--	--

Table 5: Negative turning points turning away from understanding (moving away from solution)

What went well: Pre turning point	Turning point frequencies	Key micro level characteristics during turn	How misunderstanding turned & emerged
Understanding based on extended experimentation and discussion	2	Incorrect technical and representation ideas	Misunderstanding based on technical errors
Correct explanation; Solved technical problem; Challenged incorrect idea	3	Incorrect observations	Incorrect observations
Correct task orientation; Focused experiment; Partially correct understanding	4	Incorrect or correct observations, incorrect explanations and misunderstanding	Incorrect elaboration and explanation of understanding; Focusing on the wrong details (missing bigger picture)

Discussing, comparing and evaluating our coding approach

Our analyses of turning points can be related to the pivotal moments in the multivocality project (Suthers, 2013), in particular with Chiu's (2013) statistical discourse analysis of a fractions lesson. Chiu used statistical modeling to map the characteristics of conversation turns (their micro level) within a broader context of the classroom (their meso level). Similar to our method, Chiu evaluated conversation turns with consideration of the context of the previous action, and utilizing a -/+/0 scale as we did. Different to our approach, Chiu considered various dimensions of the micro level, separately considering if each turn: was correct/valid, invited further participation, contained novel content, and was an agreement with the previous turn. Our grain sizes increased in smaller increments, and we kept the students' understanding at the center throughout – our largest macro grain size was still about the students' work (their experiments) and we integrated the classroom context as relevant throughout the grain sizes. We agree with Chiu that student disagreements, even when incorrect, could be productive at stimulating thought, action, and other perspectives.

Also in the multivocality project, Sawyer, Frey, and Brown (2013) found that two main collaborative moves that enable knowledge-building discourse were: (1) collaborative elaboration of ideas and (2) self-monitoring content understanding. On the flip side, they found that groups experienced problems when: (1) the critical features of the problem were not explicitly focused on, (2) students asked closed questions, and (3) a lack of elaboration. Our findings are very much in line with this; we also found that elaboration of ideas was a critical component for the progression of student understanding. The problem of missing critical features we believe is related to the problem of missing the bigger picture, which can be a particularly elusive problem as students might have a false sense of confidence and not realize what they do not know. Similarly, Mameli and Molinari (2011) analyzed the interactive micro-processes and turning points in classroom discourse, and similar to us, found that students had challenges with focusing on the wrong details or making incorrect observations which can lead down a garden path of misunderstanding. They highlight the importance of describing the "order and disorder" in a classroom, which we believe is similar to our consideration of when things turn for the better or for the worse – we believe there is great value in examining when both understanding and misunderstanding emerge. However, their work was limited in that they did not reach a precise definition of turning points – one of the utilities in our methodology is that a turning point could be defined quite precisely.

Sanderson and Fisher (1994) outline "8Cs" as 8 different general transformations that can be done as "primitive smoothing operations" on rich, sequential, human computer interaction data; they include chunking, comments, codes, connections, comparison, constraints, conversion, and computation. As these can be conceptualized as the components to be considered in an analysis that involves analyzing video or observation data in an exploratory way (Dyke, Lund, Suthers, & Teplovs, 2013) such as ours, we have evaluated our coding approach against these 8Cs. Regarding chunking, which refers to how the data is grouped into phases – with great consideration for how our choice of grain size might influence the understanding gleaned, our approach incorporates three parallel and hierarchical grain sizes. Regarding comments, which refers to the informal ideas and notes about the data an analyst might have – in our analysis software (Elan), we preserved multiple tiers for these kinds of notes. We had tiers marked for notes to self, ideas for new coding directions, and notes to discuss at group meetings. It was useful that we built an organized way to record the unorganized ideas that emerge during analysis. Regarding codes, which refer to labels assigned to chunks to reduce variability while preserving meaning

– we found that such codes were useful at the micro idea level, but we feared that with the complexity of the larger meso and macro grain sizes that meaning would be lost if codes were used. However, at the end of our turning point analysis, we were able to meaningfully group together larger-scale patterns. Regarding connections, we have a method that is very strong at providing connections at multiple grain sizes in temporality, and also very strong at connecting related actions (e.g. written responses, computer actions, speakers, classroom activities) in temporality, but perhaps more could be done to connect events in a context that is not organized by temporality or idea category type. Regarding comparison, we had an initial inter-rater reliability of 87% (with 100% agreement following discussion) with a second coder (third author) analyzing approximately 10% of each video. Inter-rater reliability for segment boundaries was not measured, but could be worthwhile in future applications. Within our data in related work (Portolese et al., 2015, April; 2015, June) we compare dyads working in slightly different conditions, and we broadly compare our work to other related collaborative learning and productive failure work. Regarding constraints, which refers to filtering and selecting a part of the data for further analyses – we have done this in the greatest sense with our phase two turning points analyses of what we consider pivotal moments in changes in understanding. We also found it useful when understanding our data to play around with including and excluding various idea categories at the micro level. When doing the initial analyses, we often filtered codes of the same kind to ensure consistency. Regarding conversion, which refers to transforming the data, one of our aims for our next application of this scheme is to experiment with new ways to change and improve the way we represent the data, including the multiple layers and emphasis on patterns in turning points. Finally, regarding computation, our use of numeric impact coding and code categories at the micro level allowed for meaningful numerical summaries. Overall, considering our approach within this framework, our approach has much strength and some specific, tangible areas for us to continue to improve the design and representation of our approach.

Conclusions and implications

Our method of analysis is a useful strategy for unpacking how groups' understanding and/or misunderstanding emerges over time in a deep and rich way. Multiple, integrated grain sizes allow for a deep understanding of critical moments such as turning points by being able to “look up” and “look down” (Russ, Scherr, Hammer, & Mikeska, 2008) a level at the explanatory content and mechanisms. Our results suggest that it is critical that students attempt to elaborate for themselves regarding what they are observing, and perhaps it could be wise that teachers/facilitators check in on their elaborated understanding. There may be utility in explicitly encouraging frequent observation and potentially strengthening students' observation skills. In line with Kapur and Bielaczyc (2012), we saw a benefit when students persevered and generated as many diverse ideas as possible. Students may need to engage in written responses to demonstrate understanding or lack of understanding, and formative feedback on this developing content understanding at a deep level would be productive. When students have misunderstanding, rebuilding understanding from the ground up utilizing the classic experimental procedure (Figure 1) can be useful. It appears that floundering in relation to representative and technical elements is less productive, and more extensive support could be useful in this area. We look forward to continuing our research program by continuing to refine our methodology and strengthen our conclusions as we apply and expand it with larger data sets. Overall, the development of group understanding is an incremental process (Jeong, 2013); understanding the mechanisms of students' developing success and failure in these increments is an important key to understanding how collaborative scientific understanding emerges.

References

- Chiu, M. (2013). Social metacognition, micro-creativity, and justifications: Statistical discourse analysis of a mathematics classroom conversation. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 141-159): Springer US.
- Damşa, C. (2014). The multi-layered nature of small-group learning: Productive interactions in object-oriented collaboration. *International Journal of Computer-Supported Collaborative Learning*, 9, 247-281.
- Dyke, G., Lund, K., Suthers, D. D., & Teplovs, C. (2013). Analytic representations and affordances for productive multivocality. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (639-658): Springer US.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95, 393-408.
- Greeno, J., Collins, A., & Resnick, R. (1992). Cognition and learning. In B. Berliner & R. Calfee (Eds.), *Handbook of Educational Psychology* (pp.15-46). New York: Simon & Shuster MacMillan.
- Jacobson, M. J. (2000). Problem solving about complex systems: Differences between experts and novices. In B. Fishman & S. O'Connor-Divelbiss (Eds.), *Fourth International Conference of the Learning Sciences* (pp. 14-21). Mahwah, NJ: Erlbaum.

- Jacobson, M. J., Kim, B., Pathak, S., & Zhang, B. (2013). To guide or not to guide: Issues in the sequencing of pedagogical structure in computational model-based learning. *Interactive Learning Environments*. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/10494820.2013.792845>
- Jacobson, M. J., & Markauskaite, L. (2015, April). *Understanding complex systems and climate change: Learning designs with agent-based models, productive failure, and analogical encoding*. Paper presented at the 2015 Annual Meeting of the American Educational Research Association, Chicago, IL.
- Jeong, H. (2013). Development of group understanding via the construction of physical and technological artifacts. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 331-351): Springer US.
- Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences*, 21, 45-83.
- Kapur, M., & Kinzer, C. K. (2007). Examining the effect of problem type in a synchronous computer-supported collaborative learning (CSCL) environment. *Educational Technology Research and Development*, 55, 439-459.
- Kapur, M., Voiklis, J., & Kinzer, C. K. (2008). Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups. *Computers and Education*, 51, 54-66.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75-86.
- Loibl, K., & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction*, 34, 74-85.
- Mameli, C., & Molinari, L. (2013). Interactive micro-processes in classroom discourse: Turning points and emergent meanings. *Research papers in education*, 28(2), 196-211.
- Nersessian, N. J. (2008). Mental modeling in conceptual change. In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 391-416). New York, NY: Routledge.
- Portolese, A., Markauskaite, L., Lai, P. K., & Jacobson, M. J. (2015, April). *Model-based learning with productive failure and analogical encoding: Unpacking learning dynamics with contrasting designs*. Paper presented at the 2015 Annual Meeting of the American Educational Research Association, Chicago, IL.
- Portolese, A., Markauskaite, L., Lai, P. K., & Jacobson, M. J. (2015, June). *How collaborative successes and failures become productive: An exploration of emerging understanding and misunderstanding turning points in model-based learning with productive failure*. In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.). Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 2 (pp. 655-656). Gothenburg, Sweden: The International Society of the Learning Sciences.
- Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, 2, 235-276.
- Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science education*, 92, 499-525.
- Säljö, R. (1991). Learning and mediation: Fitting reality into a table. *Learning and Instruction*, 1, 261-272.
- Sanderson, P. M., & Fisher, C. (1994). Exploratory sequential data analysis: Foundations. *Human-Computer Interaction*, 9, 251-317.
- Sawyer, K., Frey, R., & Brown, P. (2013). Knowledge building discourse in peer-led team learning (PLTL) groups in first-year general chemistry. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 191-204): Springer US.
- Suthers, D. (2013). The productive multivocality project: Origins and objectives. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 3-19): Springer US.
- Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction*, 11, 381-419.
- Wilensky, U. (1999). *NetLogo*. Evanston, IL: Center for Connected Learning and Computer-Based Modeling. Northwestern University (<http://ccl.northwestern.edu/netlogo>).

Acknowledgments

We thank participating teachers and students, and the anonymous reviewers for their productive suggestions on earlier phases of this work. This research was supported by the Australian Research Council grant #LP100100594.

Hacking a Path In and Through STEM: Unpacking the STEM Identity Work of Historically Underrepresented Youth in STEM

Edna Tan, University of North Carolina at Greensboro, e_tan@uncg.edu
Angela Calabrese Barton, Michigan State University, acb@msu.edu

Abstract: In response to the vast inequities minoritized youth experience in STEM, we investigate how minoritized youth imagine and author pathways towards becoming in STEM. Drawing from social practice theory, identity and mobilities of learning, we investigate, through critical longitudinal ethnography, how youth hack a pathway in and through STEM as they remix and repurpose tools, practices, relationships and artifact within and across the STEM-related worlds in their lives. We present one case study to articulate how youth engaged in pathhacking identity work to both gain access and transform both the process and outcomes of STEM participation in both informal programs and school science. The process, salient factors and implications of pathhacking identity work are discussed.

Keywords: identity work, minoritized youth, STEM, equity

“When I think of STEM pathways I think of science fairs and stuff. I don’t think I was ever that kind of a person... It makes me feel like an imposter to call my pathway a STEM pathway.” - Cathy, pre-med university freshman

Introduction

Large gaps in achievement and interest in engineering and the physical sciences persist for youth growing up in poverty, and in particular for African American and Latino youth from low-income communities. These gaps persist across all levels of educational attainment. In the United States the percent of bachelor degrees awarded to African Americans in engineering in the U.S. has hovered around 4% (Yoder, 2014). The research literature documents many reasons for these persistent gaps, including inequitable access to resources, quality instruction or role models, along with cultural barriers and stereotypes (Margolis, et al., 2008; Oakes, 2005). Yet, success in the sciences and engineering is one viable route towards personal and/or community economic advancement for youth growing up in poverty. It also factors into opportunities for informed, meaningful, and empowered democratic participation. That lower-income communities of color experience the greatest levels of environmental injustice in the US, and often have the least voice in STEM-related decisions affecting their communities are further evidence of the impact of persistent inequities in access to the sciences and engineering (National Academy of Engineering, 2010).

Our work is in response to these vast inequities that youth from lower-income communities of color face. We know that institutional and classroom level practices in schooling and society have unfairly positioned lower-income youth and youth of color as non-experts and outsiders to STEM. Even when such youth succeed in science academically, they disproportionately do not view themselves as having a future in science. Take, for example, Cathy, whose quote opens this paper, is an African American female growing up in a lower-income community and attending a majority minority school, Cathy performed well in school and earned a full scholarship to college. Later, Cathy points to the people and out of school experiences that kept her in STEM despite not having access to the more traditional STEM enrichment experiences of the prototypical STEM person. She sees herself as an imposter in STEM, one who does not belong, despite her academic success.

The overarching question that guides this manuscript is: *How do youth imagine and author pathways towards becoming in STEM?* We use the term “imagine” to capture, as Boal (1974/1979) might say, the places where expectations can become desettled, generating new practices grounded in the world as it is, and the world as it could be. In particular, we are interested in the ways in which youth navigate their worlds, and author spaces of learning, doing and becoming in STEM. Tensions arise, however, when the different institutional, social, cultural and political forces push back on youth as they work to author these pathways, as Cathy noted in her comments, causing her to feel like an imposter.

Conceptual framework

We draw from both social practice theory and mobilities of learning studies to frame identity as constructed in social interaction, over time, and across many different spatial (physical and virtual) locations. We take as a

starting point the idea that identity is constructed in practice – practice that requires knowledge, skills, and ways of thinking that characterize the community in which one is engaging (Holland, Lachicotte Jr., Skinner, & Cain, 2001). Identity is also a highly social and dynamic construct and is related to how individuals are recognized by others at any given moment, in ways that support or work against who they are and want to be. However, neither one’s agency to act in particular ways or how one is recognized by others for what one does, are constant. Such things can quickly change as individuals move across space and time where their individual and collective actions are differentially enabled and constrained by social structures-in-motion.

We therefore find it more productive to focus on *identity work* rather than identity itself. By identity work we refer to the actions that individuals take and the relationships they form towards becoming particular kinds of people. The reception, or recognition, of these positionings by the community highlights the dialectical nature of identity work. Because identity work happens within and against local norms and expectations and as a part of longer standing sociocultural and historical narratives, its outcomes are always uncertain and gain new meaning as they get traced in time.

Leveraging on a mobilities of learning framework (Engeström & Sannino, 2010; Gutiérrez, 2012) helps us situate identity work across multiple time scales, different spatial (physical and virtual) locations, and both vertical and horizontal dimensions. The real and imagined geographies of learning experienced by youth as they work on, and within, the powered boundaries of science and community, all play into how young people position themselves with and against the normative expectations of becoming in science (Bright, Manchester & Allyndale, 2013). Studies have documented how individuals navigate and bridge the worlds of home, school, and community, including how they move people, practices, tools, and ideas across these settings (e.g. Ehret & Hollet, 2014). These studies remind us that there is continuity between youths’ worlds and that of science and that we best understand these worlds as “generative resources in learning new ideas and traditions of inquiry” (Warren, Ogonowski, & Pothier, 2005, p. 121). They have also led to the recognition that youths’ mobilities among a vast range of learning arrangements makes learning and identity work always “tangled up” among practices in complex ways (Rahm, 2012).

Identity work across time and space is one way to understand how youth author paths to possible futures in STEM. Possible paths are created and facilitated by power-mediated opportunity structures that some youth can traverse, but others struggle to understand, let alone be welcomed by. Often times, youth-centered actions towards path authoring are not recognized by traditional structures and gatekeeping authorities (Authors, 2013). A focus on how youth author paths through their identity work highlights the forward trajectory of young learners’ decisions and opportunities toward promising possible futures in/with science. Seeing such movement allows for nuanced examinations of the twists and turns a possible path takes throughout individual learning landscapes. How youth choose to engage in science, for what purposes, where and when all shape, and are shaped by, the people, places, events, and power structures that constrain or expand activity. This approach also foregrounds the multiple directions one may take with science, the various on/off ramps into/through science, and the agency youth have to author within/across the multiple layers and contexts of learning experiences in science.

Methods

We draw from data spanning 12 years of research (2013–2015), across four studies, including two that are ongoing. These studies have all utilized critical, longitudinal ethnography, where we have followed youth across sites (four states), spaces within sites (community-based, home, informal science, school science), and time (focused, youth case studies that are between one year long to 5 years and ongoing). For the two cases presented, we observed Quentin for 6 years in both formal school science and informal science settings, in a Midwestern state. For Melanie, we observed her for 1.5 years in her formal school science setting in a Northeastern state. We purposefully select these two cases to highlight the importance of two different grain-sizes of longitudinal ethnography and the focusing on different spaces. These two cases also allow us to unpack the kinds of pathhacking identity work that can take place across these varying spatial and temporal domains. Data sources (Table 1) include extensive fieldnotes across sites for each case study youth, narrative and artifact interviews with youth, focus group interviews with youth and their peers, interviews with teachers and adult mentors, and STEM artifacts youth produced.

Table 1: Date Types Generation Strategies

Data Form	Specific Data Generation Strategy	Quentin	Melanie
Participant Observation	<ul style="list-style-type: none"> Informal science setting (Green Club): Video recordings of twice weekly sessions and field notes 	60hrs/yr 20hrs/yr	n/a 130hrs

	<ul style="list-style-type: none"> Formal science classroom: Observations and video recordings 		
Interviews	<ul style="list-style-type: none"> Focus groups & solo interviews 	10hrs/yr	5 hrs
Artifact Think Aloud	<ul style="list-style-type: none"> Allowing youth opportunities to talk about their engineering design work in detail (mid and end of year) 	2hrs/yr	2hrs
Artifact Collection	Youth's science related work collected: examples include school science projects and classwork; for Quentin, also artifacts from Green club, including sketch up notebook, 3D Google SketchUp model of design, worksheets, prototype, movie, etc.		

Analysis

Data was analyzed by our research team in the grounded theory tradition, using a constant comparative approach (Straus & Corbin, 1998). We first engaged in open coding by thoroughly perusing all generated data (e.g. transcribed interviews, observation fieldnotes) to surface a) critical episodes of STEM engagement that youth subsequently leveraged in some way for future STEM engagement; b) how youth positioned themselves during these critical episodes; c) how youth responded to how others positioned them within these episodes; and d) the artifacts youth chose as most representative of them in STEM, and why. We took these codes to be signifiers of the identity work youth were engaging in. Weekly conversations were held between the authors on these insights as a way to work towards a more “expansive consensus”; that is to say that any differences in view were debated until new meaning was generated as a result of our differences. A detailed list of emergent open codes were kept with analytic memos attached to them which we brought to bear on the second pass at axial coding.

With our theoretical framework as a guide, we then looked for evidence of how youth repeated performances of salient identities-in-practice, across space and time. In particular, we looked for what youth did (the actions they took), and how they went about it (the resources they recruited), to create new opportunities and spaces for becoming in STEM that were previously denied. The relationships and connections identified in this second stage of coding, in turn, guided our selective coding, and became categories and themes, from which our example cases were selected for a final round of analysis and presentation

Findings

We purposefully selected two case study youth, Quentin and Melanie, as focal cases. We highlight these two cases because they reflect two highly dynamic cases of identity-in-the-making, where youth respond to and push back against the people, contexts, sociocultural histories and normative views and expectations, all of which shape their opportunities to become in STEM. In both cases, the paths authored by youth are non-traditional and non-linear, involving identity work that required creative cross-leveraging and re-mixing of available resources.

Their stories help us to see how identity work, in the moment and over time, make authoring paths into STEM more or less possible. At the same time, we selected these two cases because they reflect a profound difference in how identity work in the moment and over time juxtapose in ways that productively yield possibilities for the future, both real and imagined. Whereas Quentin eventually both narrates and embodies a possible future in STEM, Melanie struggles to connect who she is and what she brings to the STEM table as something that holds potential, despite her moments of success. These differences in their stories help us to describe the complicated process of path authoring through the on-going and high contentious interactions among self, science and social contexts. Finally, these two cases were selected because they help us to interrogate how the matrix of race, class and gender intersect and play an active role in the on-going opportunities youth imagine as possible in path authoring. The differences in timescales between the two cases are further leveraged to shed light on how the simultaneous use of different timescales helps to surface the insight on the role of in-the-moment identity work on imagining and authoring pathways in STEM.

For the purposes of this proposal, we present only Quentin's case.

Quentin

Quentin is a gregarious African American young man growing up as an eldest son in a single parent family. In the six years we have known Quentin, his family has moved three times in search of more affordable living conditions. We first met Quentin when he joined the Green Energy Club [GEC] afterschool program in the fifth grade, at the insistence of the Club's main director. He told us that he did not mind being in GEC because he

could “use the computers to make movies” and “do things” for his community. However, he also indicated that he was not interested in science.

Vignette 1: Quentin’s letter to his 5th grade teacher

In January of his 5th grade year (6 months into the school year), Quentin wrote a letter to his science teacher. He entitled his letter, “BRINGING FUN SCIENCE TO YOU! (A.K.A) MR.B!” The text of the letter is as follows:

Hi Mr. B, this is your student Quentin the first one in the 2nd row. I’m going to tell you about things that we should do in Science. I’m in [GEC] and [GEC] helps me get my grade for science to like a B or a A. I really don’t like getting lectures and getting assignments out of those 20 year-old science books, it doesn’t have the latest science news and new experiments.

I do things out of school and out of GEC that involve science. I went door to door and ask adults if they use CFL lights. The majority of the adults did NOT use CFL lights, I will try to decrease the amount of people who use incandescent lights. I did it on Wain Wright Ave. and I did it because people’s bills are up because they use just Incandescent lights. In GEC we made a word that is probably not going to be in the dictionary, but anyways the word is called Fcience, it is Fun and Science put together and I want you to make our science class into a Fcience class. Here’s the definition of Fcience: Doing things that are interesting that include scientific things.

Its not so much for energy that I get attention in school, but for being a smart alek. I think that should be good. GO FCIENCE!”

Vignette 2: Power-Sucking Pigs, 6th grade & 6th grade science

In the summer leading into 6th grade, Quentin participated in a two week GEC summer program focused on solar energy, which took place at the local university. Quentin investigated solar energy as a viable renewable energy source for the energy challenges his community faces, both economic and environmental. As part of this work, he along with another young man, Cam, produced a 60 second video-based public service announcement – “Power Sucking pigs” – aimed at educating his local community on the importance of these issues (see <http://getcity.org/blog/2012/10/06/power-sucking-pigs/>).

In the video, the two boys start out acting out the role of business associates in a meeting using their laptop computers. Suddenly, the overhead lights in the room flash and the electrical power goes out. The two business associates look at each other and ask if the laptop lost power. Trying to figure out what caused the power to go out, they scan the room and discover there are “power sucking pigs” draining all the energy from the wall outlets. The pigs are also played by the two boys who have donned pig noses. The boys infuse their sense of humor as we see the business associates chase the pigs set against the sounds of rhythmic drums. The video transitions to a serious tone as the boys narrate and show graphics of energy use in the US compared to the rest of the world. They also add information and images for ideas about how to save energy. The boys included a picture of a home with solar shingles from a field trip taken in the summer program. The video closes with the boys standing in front of a mirror. The boys in unison remove the pig snouts and point to the reflection in the mirror and declare, “It’s you!” The video finishes with a link to the group website where viewers can find additional information.

In this PSA, Quentin highlights the need to unplug electrical cords from outlets when they are not in use in order to reduce carbon dioxide emissions and save energy. He provides information about energy consumption in the US and how using solar panels can reduce electric bills and emissions. All of the information provided in the PSA came from his scientific investigation (e.g., investigating how much CO₂ is emitted from burning fossil fuels, testing how solar panels work, and measuring standby power) and field trip experiences (e.g., visiting a solar powered house and a solar research lab). He also provides information about energy consumption in the US and how using solar panels can reduce electric bills and emissions. He wanted their movie to speak to people like his mom – making ideas “real” not distant, and making change something that was possible for lower-income families. Quentin decided to bring his movie later that fall into school to show his science teacher. Of the power sucking pigs movie, he said, “It’s the movie that changed how people thought about me... We showed people how they can save electricity, which will help with CO₂. Mainly, I was excited to show my teacher because he saw that I could do it. That I got it done. And that I know a lot. I had to get it done. I’m not really that C and D person.”

At the time Quentin shared this video, he was struggle in school, in general. We often found him sitting in the hallway, having been sent out of class for “clowning around.” However, Quentin felt that this punishment was unfair, as he felt he was not clowning around. He said he was just a funny person. Because he spent so much time outside of class, sitting in the hallway for his behavior, he missed many assignments, leading to low grades.

He did not seek to make up the work because, as he says, “what is the point?” Quentin says his teachers did not know who he really is or what he is capable of.

Vignette 3: Grand Climate Change (7th grade)

In the 7th grade, Quentin and his peers were studying how the GEC might “get off the power grid”. Over the course of 24 weeks, Quentin engaged in activities investigating the electrical production system and its alternatives, with a particular focus on energy transformations in these various pathways. About mid-way through this unit, Quentin learned about a contest for “youth innovators” hosted by a local non-profit organization focused on promoting entrepreneurship. Any youth under the age of 18, could submit to the contest a prototype of an idea that could be brought to market. On his own time Quentin worked on prototyping a game, which he named Grand Climate Change.

Quentin’s reason for focusing on a video game was that he wanted to teach young people about the impact of climate change, which he felt was already “in action.” As he states, “My idea was a Video Game. As kids, we have already seen climate change in action. Hurricane Sandy. Hurricane Katrina. We have seen lots of tornados and big storms all across the country. We have had a warm winter with hardly any snow. This year was Dustbowl 2012.” He was concerned, however, that his peers would not be interested in playing science game, even if was a fun computer-based game. His idea for getting kids interested was to model the game after what he knew to be a very popular game among his peers, Grand Auto Theft. As he states:

I wanted to create a video game that teaches other kids about climate change. One game that I like and that lots of other kids like is Grand Theft Auto. This game is about taking missions from the masters and completing them. It’s kinda violent. But, it’s fun. It’s popular. I like taking missions. So, my climate change game would be like this, but it would not be violent. The missions would be about doing thing to help CO₂ from not building up. Each mission, you have to know more, or learn more about the causes and effects of climate change.

Like Grand Auto Theft, the game would involve main character, Tony, whose job it was to solve the missions and rise through the ranks and take on more difficult and complicated gameplay. Tony would encounter antagonists in his journey, who would attempt to impede his progress. Grand Climate Change not only draws upon the gameplay of Grand Auto Theft, but also draws heavily upon what he had been learning overtime in his after school program, in both in the 6th and 7th grade. As Quentin states,

My game . . . would have a main character Tony, who goes around the city and does missions. He helps people do stuff to help prevent climate change. You have to have ideas. You have to understand how the things you do contribute to climate change.

I started with three missions. Mission #1 is to go to someone’s home and remove all incandescent lightbulbs and replace them with CFLs. So Tony might come to a home of a person who sells incandescent lightbulbs and they refuse to let him in. So, he has to convince him why it is important. How it works. You have to think of all the reasons why someone might care. You need to think about the strategy because the more angles you hit, the better you do.

Quentin was motivated to make his game because of a contest that he learned about from a peer in the GEC program. He was excited by the prospect of winning both a camera and a monetary prize (\$250) if he placed in the top three. Winning the contest was also another way that Quentin felt like he would be able to prove to people that he was smart and capable in science. In the end, Quentin won second place in the state-wide contest.

Vignette 4: Summer Engineering Program (7th and 8th grades)

Quentin applied for and received a full scholarship to a summer engineering program at the state’s technical university, about 7 hours bus ride from home, and in a part of the state he had never been. He felt that having a chance to take engineering classes on a university campus would help him to be better prepared for a future in STEM. As he stated, “When I finished 7th grade, I decided to apply to Tech University for a summer camp program to learn about engineering. We do not learn about engineering in school, only in GEC. I thought I needed more of chance to see what it is about. I was accepted to the program with a full scholarship. When I was there took civil engineering classes, and we built bridges and things there and learned about it. It was interesting to me.” Quentin enjoyed the program so much that the following summer, after completing the 8th grade, he applied again. The eligibility requirements explicitly state that students could only receive the scholarship one time. However, he requested to be able to apply again because engineering is not taught at his school, and that science was also

barely taught at his school. Quentin felt that if he could attend the summer residential program again he would be able to catch up on what other students in other schools probably had the chance to learn. He referenced his accomplishments, such as Grand Climate Change and Unacceptable Heat, as indications of his ability to work hard over time, and as indications of his desire to become an engineer. As he stated, “I knew I had to apply again after 8th grade. No one has ever been able to go twice, but I knew I had no choice but to try. I felt lucky that I got accepted.”

Discussion

Quentin’s story reminds us how youths’ identities are always in-the-making, responding to and pushing back against the people, contexts, sociocultural histories and normative views and expectations which shape their lives. Over time, and across the different spaces of their lives, we have seen these young people grow into particular kinds of STEM people who challenge normative expectations for who can be a STEM person and what it means to be a STEM person. We have seen them author paths into STEM that are nontraditional, and far from linear. They have strategies in re-mixing and re-purposing both traditional and nontraditional resources – material and symbolic – alike, in order to engage in the kinds of STEM work they care about, and to be recognized for who they are and what they can do. Institutions, people, tools, and practices, all imbued with and embedded in histories, have structured the youths’ opportunities to become, and have also provided points of resistance as youth refuse marginalization.

Rather than a STEM identity “pathway”, which suggests a pre-laid out route with helpful and visible signposts or at least an obvious track for one to walk on, we have come to see minoritized youths’ identity work, over time, as a form of *pathhacking*. We use the term ‘hacking’ because it refers to the need to wield creative force (or agency) to imagine a way forward, most of the time through unclear territory, with unknown outcomes or stopping points along the way. We also use ‘hacking’ to convey the characteristics inherent in authentic, hacker subculture; that of playfulness, excellence, and boundary pushing, all undergirded in egalitarian principles (Himanen, Linus & Castells, 2001). We see minoritized youths’ identity work in STEM reflect these similar characteristics, as they seek for more elbow room at the STEM table, and opportunities to transform that table. There is force in our view of hacking because there is resistance. The traditional pathway laid out for minority youth is AWAY from STEM (Berry, 2005; Gándara, 2006, Triana & Rodriguez, 1993)

For example, if we return to Quentin’s letter that he wrote in the 5th grade, we can see some of the struggles that he faces as he considers his own pathway into STEM in the 5th grade. He points out which student he is – “This is your student Quentin, the first one in the second row” – suggesting that he worries he does not know whether his teacher really knows who he is, despite having been in his class for 6 months. He also contrasts his feelings about being a passive recipient of science (“I don’t like getting lectures”) with his active engagement in community (“I do things outside of school... I went door to door.”). He points to the creative ways in which he leverage his community funds of knowledge in order to help make science more accessible and salient to his community. However, he also indicates, in the last paragraph, how important it is for others to recognize and value his efforts for what they contribute to community, not for how they reflect negative stereotypical views.

As we consider the question of how youth imagine and *hack* paths into STEM, it appears that a salient feature of this identity work relates to how these young people try on new ways of being through their varied forms of engagement and actions in response to particular norms or sanctions of the worlds they inhabit. As such, the youth are engaged in an on-going process of negotiating between their individual agency and community response to their efforts, in order to have their identity work legitimized as an important part of becoming in STEM. However, their hacking is often seen in the contemporary sense of hacking, that of *trespassing*. Such judgment lead youth to see themselves, as Cathy does, as an imposter. We present three ideas derived from our case studies that describe the salience of the identity work that youth do across towards their authoring a way into STEM: 1) the nature of pathhacking; 2) the tools of pathhacking, and 3) the pitstops, dangerous intersections and cul-de-sacs of pathhacking.

Trying on new ways of being with/against norms and expectations: The nature of pathhacking

The youth in our study are engaged in the on-going practice of trying out new ways of being in STEM, both with and against the norms of expectations of the worlds they inhabit. These practices are not always fully intentional towards identity building, but they reflect moves to preserve, defend, validate, or call attention to the lived experiences they bring to learning and becoming in STEM. While such practices, over time, can position youth against oppressive norms in a bid to seek new avenues of agency, they also bear potential academic risks.

The youth appear to try out these new ways of being not for the sake of resisting sanctioned norms themselves (e.g. in opposition to the authority figure, often teachers, and sometimes peers, who positioned them on the margins of STEM). Rather, they appear to do so in order to gain recognition for the non-traditional forms

of “STEM capital” they possess, in order to transform the discourse, and what counts as legitimate ways of being and doing in STEM (Archer et al, 2015). In Quentin’s case, he authors the letter to his teacher not in antagonism, but to remind his teacher that he is a hard worker and one who cares about science and his community. By introducing the term “Fcience,” he suggests, however, that the world of *school* science has not yet offered the same affordances for being in STEM that his community has offered, and suggests changes that could be made to his classroom. He is aware of the constraints his teacher and school face (e.g., having to use 20 year old textbooks), but suggests that it is still possible to do “fcience.”

Another path hacking maneuver youth perform towards productive STEM identity work is to re-organize and/or create new figured worlds more suited to who youth are in-the-moment. Quentin drew on compelling reasons (he has no access to engineering in his public school), his deep personal interest in engineering, and his future career aspirations in order to secure for himself another scholarship to the engineering summer camp, even when he was, being a previous awardee, technically ineligible for another application,. Such hacking acts open up new possibilities for thinking about oneself and one’s futures differently (and what one needs to learn and do to get there). Quentin further laminated his STEM identity in the second camp as a result of this hacking move.

Critical STEM identity artifacts and allies: The tools of pathhacking

STEM is not an easy world to navigate for the young people with whom we have worked. A fun activity, a personal connection, or a scientific understanding is often not enough for youth to see themselves a part of that world, although these are the solutions often offered in classrooms or reform documents. The youth have had to author new routes that demand re-organizing worlds and/or creating new worlds for their identity work to be recognized and valued. Pathway hacking requires both tools (what we call “critical STEM identity artifacts”) and allies (social relationships) for figuring out these unknown worlds one is working within (and its affordances and constraints) and to work towards reconfiguring these worlds. In particular, we noticed three important roles that hacking tools and allies jointly play.

First, both tools and allies help to break down marginalizing binaries, such as that of novice/expert, insider/outsider, successful/unsuccessful. Second, tools and allies legitimized ways of being that are more inherently germane to youths’ sense of selves instead of the privileged “other”. Third, they broker connections across potentially deep chasms. Quentin’s Grand Climate Change game (tool) reflects his desire to engage others in broad environmental issues related to everyday practices such as lighting and driving. He is deeply aware of the precarious nature of these practices for the people in his community. In authoring the game, he had to draw from, and synthesize information from various domains: peer culture, video gaming culture amongst teens, video gaming infrastructure, nuanced content understanding about the different issues salient to Climate Change and translating content into a gaming format. Indeed, this suite of expertise that Quentin demonstrated was recognized when he won for the \$250 prize. His GC adult mentors (allies) supported his game development. When Quentin referenced these accomplishments in his application to Tech U’s summer program, he further leveraged these tools towards new inroads into STEM (brokering connections).

The pitstops, dangerous intersections, and cul-de-sacs of pathhacking

We have noted how identity work accrues over time, as small events disrupt oppressive forces in youths’ lives. Each disruption becomes a moment where ideas, tools, and bodies can refigure learning and becoming, giving rise to new relationships and opportunities for becoming in STEM. We view these moments of re-grouping and re-building as pitstops, or rest areas where youth and their allies can begin to imagine for themselves what the possibilities of becoming in STEM might be. We also view pit stops as safe place to rest or get some help, but also as a place to regroup and to figure out how to tackle the terrain ahead. For Quentin, pit stops included various Green Club experiences and consecutive summer programs at the local technical university.

At the same time, such pathhacking is not always forward moving, and young people can find themselves facing dangerous intersections, crossings or even deadends. As Quentin entered his teenage years, the structures of his afterschool club required him to spend his spare time in the teen room, where the rules for engagement were distinctly different from the younger spaces and the GEC. He eagerly joined the teen room, and relished in his status there. He spent more time hanging out, and less time finishing homework. More energy was put into social relationships, time on his phone, and appearance. It was not that any of these practices were in conflict to his vision of becoming in STEM, but they did not mesh with the club director’s view of what it meant to be a teen leader. His participation in GEC and attendance at Tech U was threatened, and he felt that he was forced between worlds he had found ways to marry in the past. That he spent his time in GEC designing a cell phone charger (while being a teen) speaks to one way he sought to push back against dichotomous choices he felt were presented to him, and at least implicitly, sought to re-organize his worlds so that his new teen status mattered in STEM. But Quentin was lucky in many ways. He carried with him a history of relationships through GEC (allies) that made

this identity work as a teen possible in his work on the cell phone charger, and he had a wide repertoire of tools for bridging the chasm between these worlds.

Conclusions and implications

In many ways it is impossible to see how Quentin navigated the pitstops and dangerous intersections without calling into question timescales. We cannot look at any individual act as taking place apart from a history of engagement across the spaces that make up their lives, what they do in those spaces, and how they sought to move ideas, practices, and tools from one space to another. While achievement scores offer a form of evidence for student learning, they reveal only a narrow slice of it --they tell us little about what students understand science to be or the mechanisms by which students come to engage in meaningful science learning, and to see themselves as a part of science. Indeed, *who individuals are and want to be* in science has serious implications for how or why one might engage in science class, enroll in science courses, or use scientific ideas and practices outside of the classroom.

References

- Boal, A. (1974/1979). *Teatro del Oprimido*. Urizen Books (translation to English 1979 by C. A. McBride & M. O. Leal McBride).
- Berry, R. Q. (2005). Introduction: Building an infrastructure for equity in mathematics education. *The High School Journal*, April/May, 1-5
- Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*, 5(1), 1-24Bright, Manchester & Allyndale, 2013
- Ehret, C., & Hollett, T. (2013). (Re)placing school: middle school students' counter-mobilities while composing with iPods. *Journal of adolescent & adult literacy*, 57(2).
- Gutiérrez, K. (2012). Re-Mediating current activity for the future. *Mind, Culture, and Activity*, 19, 17–21.
- Gándara, P. (2006). Strengthening the academic pipeline leading to careers in mathematics, science, and technology for Latino students. *Journal of Hispanic Higher Education*, 5(3), 222-237.
- Holland, D., Lachiotte, D., Skinner, D., & Cain, C. (2001). *Identity and agency in cultural worlds*. Cambridge, MA: Harvard University Press.
- Margolis, M., Estrella, R., Goode, J., Jellison Holme, J., & Nao, K. (2008). *Stuck in the shallow end: Education, race, and computing*. Cambridge, MA: MIT Press.
- National Academy of Engineering. (2010). *Engineering, Social Justice, and Sustainable Community Development*. Washington DC: National Academies Press
- Oakes, J. (2005). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.
- Rahm, J. (2012). Collaborative imaginaries and multi-sited ethnography: Space-time dimensions of engagement in an afterschool science program for girls. *Ethnography and Education*, 7(2), 247-264.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd Ed)*. Thousand Oaks, CA: Sage Publications.
- Triana, E. M., & Rodriguez, M. G. (1993). *United-Unidos: Mathematics and science for Hispanics*. Washington, DC: American Association for the Advancement of Science.
- Warren, B., Ogonowski, M., and Pothier, S. (2005). "Everyday" and "scientific": Rethinking dichotomies in modes of thinking in science learning. In R. Nemirovsky, A. Rosebery, J. Solomon, and B. Warren (Eds.), *Everyday Matters in Mathematics and Science: Studies of Complex Classroom Events* (pp. 119-148). Mahwah, NJ: Lawrence Erlbaum Associates.
- Yoder, B. (2014). *Engineering by the numbers*. *American Society for Engineering Education*. Available at: <http://www.asee.org/papers-and-publications/publications/11-47.pdf>_ Accessed November 5, 2015.

Acknowledgments

We thank participating youth, community members and teachers. This work was funded by [blinded].

Local Versus Global Connection Making in Discourse

Wesley Collier, The University of Wisconsin-Madison, collier.wesley@gmail.com

Andrew Ruis, The University of Wisconsin-Madison, andrewruis@gmail.com

David Williamson Shaffer, The University of Wisconsin-Madison, dws@education.wisc.edu

Abstract: This paper examines techniques for modeling relationships among domain concepts and practices in discourse to assess learning in a CSCL environment. We compare two approaches: a traditional psychometric approach, which models the global correlation structure of student discourse markers across the learning intervention, and a model that accounts for the local correlation structure of discourse markers within activities. We investigate whether: (a) analysis of local correlation structure can identify significant differences between novices and relative experts; (b) these differences reflect meaningful differences in the discourse; and (c) analysis of global correlation structure can identify significant differences between novices and relative experts. We assess whether an approach that models local relationships among concepts in a domain provides useful information beyond what might be extracted from a more traditional modeling approach. Our results indicate that techniques that account for local correlation structure can identify patterns in discourse not reflected in global correlation structure.

Keywords: assessment, epistemological cognition, learning analytics, simulations

Introduction

Digital learning environments capture student discourse as they interact with mentors and peers to solve complex problems (Boulos, 2006). This poses a challenge for assessment of student work. While logged discourse contains concepts from the domain in which the problem is situated, what makes discourse better is typically not just that it contains more domain concepts, or even that it contains different concepts. Chi et. al.'s classic study of expertise in physics problems (1981), for example, found that experts organize their understanding differently than novices do. Similarly Bransford et. al. (1999) showed that experts have acquired a great deal of content knowledge that is organized in ways that reflect a deep understanding of their subject matter.

In other words, the ability to solve complex problems depends not only on having access to domain concepts, but also in understanding the appropriate relationships among them — and being able to mobilize those relationships in the context of real-world problem solving.

To assess complex thinking therefore may require constructing models of the way students use their understanding of the relationships among domain concepts during the problem solving process. In this paper, we look at two classes of mathematical models that can be used to measure this kind of connectivity in discourse. One is a class of models commonly used in educational research that looks at connectivity within the discourse as a whole — that is, the global relationship among concepts within students' discourse. The second approach is a more novel modeling tool that is sensitive to how concepts are related to one another within individual topics or activities within the discourse — that is the local relationship among concepts.

In what follows, we examine these two modeling approaches. In particular we look to see (1) whether the local relationship model can identify statistically significant differences between groups of novices and experts; (2) whether the statistically significant differences are meaningful on a closer qualitative analysis of the data; and (3) whether this same difference is identified in a global relationship model.

That is, we assess whether an approach that models local relationships among concepts in a domain provides useful information beyond what might be extracted from a more traditional modeling approach.

Theory

For several decades, work in the Learning Sciences has examined how complex thinking in a domain involves not only mastering basic skills and concepts, learning how these skills and concepts are systematically linked to one another. Summarizing a broad range of studies, Bransford et. al. (1999) describe the difference between experts and novices as being less about the amount of information that experts have than it is about the way that experts organize the knowledge that they have. diSessa (1988), for example, suggests that while solving physics problems does require understanding basic concepts from the discipline, deep and systematic understanding comes from linking those concepts to one another within a theoretical framework. Novices have what diSessa describes as “knowledge in pieces,” whereas experts understand the connections among different elements of the domain. Shaffer (2004) similarly characterizes learning as the development of an epistemic frame: a pattern of associations

among knowledge, skills, habits of mind, and other cognitive elements that characterizes communities of practice, or groups of people who share similar ways of framing, investigating, and solving complex problems.

In other words, a good model of expertise needs to characterize way in which an individual (or group of individuals) understand the relationships among elements in the domain. To do so, of course, requires some context in which these relationships would be expressed: some record of how individuals approach problems in the domain. This might involve think-aloud protocols (Chi, 1997), problems that require students to “show their work” (McNeil, 2009), or records of group discussions during problem solving (Hmelo-Silver, 2004). That is, characterizing an individual’s understanding of the relationships among concepts and practices in a domain requires some record of work that can be analyzed. This record of work — which is often in the form of a logfile (Peled, 1999) or transcript — then needs to be annotated or coded for evidence of the key concepts and practices of interest. And finally, a model needs to be created that accounts for the relationships among the concepts and practices as reflected in these codes.

Traditional psychometric techniques provide several possible approaches to characterizing relationships of this kind. In theory, one could use correlation matrices to show the correlation structure of the codes based on their frequency in the logfile. However, it is difficult to compare multiple correlation matrices simultaneously due to the large amount of information contained within even a relatively small matrix (Alper, 2013) — which perhaps explains why the review we conducted did not find any examples analyzing concept-concept relations by comparing correlation matrices in the literature. Statistical techniques for analyzing the structure of correlation matrices can reduce the amount of information, making comparison possible, but summary statistics obtained from these techniques are geared towards analysis of single matrices (Cudeck, 1989). Thus, simultaneous comparison of many correlation matrices remains an active topic of research in quantitative methods and data visualization (Alper, 2013; Elmqvist, 2008).

Because it is difficult to compare correlation matrices, education researchers often use dimensional reduction techniques to analyze relationships as linear combinations of observed variables (Hall, 1977). For example, Beishuizen et. al. (2001) analyzed relationships among concepts in an essay writing activity. They identified the concepts of interest and computed the frequencies with which the concepts occurred in each essay. To identify which concept-concept relations were most common, they used a dimensional reduction to group concepts based on the structure of correlations in concept frequencies.

While there are many dimensional reduction techniques in the literature — including principal components analysis, factor analysis, item response theory, multidimensional scaling, and diagnostic classification models — most dimensional reduction techniques are similar in that they model correlation structure across all the data for a given unit of analysis. That is, they characterize the *global correlation structure* (GCS) of the data. In the analysis conducted by Beishuizen and colleagues (2001), for example, all correlations within each essay were considered equally meaningful.

In some cases, however, this approach may not accurately operationalize the relationships learners form among concepts. For example, research on discourse processing suggests that connections between concepts may be made primarily on a topic-by-topic basis rather than across discourse as a whole. Gernsbacher (1991; see also Graesser, Gernsbacher, and Goldman, 1997) argues that meaning is constructed through the hierarchical organization of ideas. A key element of this theory is that coherent discourse is structured by topic, with utterances having clear relationships to other utterances within topics, and few or no relationships across topics. Put another way, meaning is often localized within topics, and thus a model of how learners connect concepts to one another needs to account for this topic-based or *localized correlation structure* (LCS) of discourse.

One example of an approach that measures LCS’s is *Epistemic Network Analysis* (ENA), a suite of tools that can be used for identifying and quantifying connections among elements in coded data and representing them in dynamic network models. A key feature of ENA is that it enables researchers compare networks, both visually and through summary statistics that reflect the weighted structure of connections. ENA can be used to address a wide range of qualitative and quantitative research questions.

In this study, we look at a specific discourse context to explore whether GCS and LCS models suggest different interpretations of the discourse — and if so, which provides a more useful representation of the salient features of the data. That is, we examine the difference between a dimensional reduction technique that is insensitive to topical structure and a technique that is sensitive to topical structure. To do this, we analyze the chat discourse of high school and college students in *Land Science*, a virtual internship in which students work at a fictitious urban planning firm to solve an authentic urban redevelopment problem using two approaches: (1) We use epistemic network analysis (ENA), which models the structure of relationships among domain concepts and practices using correlation structures that account for the topical structure of the discourse; (2) we then analyze the global correlation structures in the frequencies of these domain concepts and practices in the same data using

principal components analysis (PCA). We compare the results of these analyses to determine whether and to what extent the two approaches find different structures of connections in student discourse.

Specifically, we ask:

- RQ1: Are there statistically significant differences between novices' and relative experts' local correlation structures that can be detected using ENA?
- RQ2: Are there meaningful differences between novices' and relative experts' local correlation structures?
- RQ3: Does PCA detect these same differences in local correlation structure by measuring global correlation structures?

Methods

Land Science: A virtual internship in urban planning

In the virtual internship *Land Science* (Shaffer, 2008; Bagley, 2010; Bagley, 2015), students play the role of interns at Regional Design Associates, a fictional urban and regional planning firm. Their task is to develop a rezoning plan for the city of Lowell, Massachusetts that addresses the requests of various stakeholder groups. Students assess stakeholder preferences to understand what community members desire in terms of socio-economic and ecological issues. Not all of the stakeholders' competing concerns can be met, so students must make decisions about which demands to meet and how to meet them. To make these decisions, students discuss options with their project teams via online chat, and they use professional tools, such as a geographic information system model of Lowell and preference surveys, to model the effects of land-use changes and obtain stakeholder feedback. At the end of the internship, students write a proposal in which they present and justify their rezoning plans.

Coding of student chats in the *Land Science* logfile

All actions and interactions that occur during implementations of *Land Science* are recorded a log file. In this analysis, we focus on the chat conversations that students had while solving the rezoning problem. The logfile contains team chat conversations (41,332 lines of chat in total) from 265 students who used *Land Science*, including high school students (novices) (N = 110) and college students (relative experts) enrolled in an introductory urban science course (N = 155). The chat utterances were coded for 17 concepts and practices from the epistemic frame of urban planning, including:

Knowledge of stakeholder representation – knowledge of stakeholders, whose requests pertain to social, economic, and environmental issues

Skills and practices urban planning using tools of the domain – discussion or actions involving the tools – broadly defined – of the urban planning domain, such as a virtual site visit to key regions in the city, a stakeholder preference survey, and *iPlan*, a geographic information system-enabled zoning model

Data-based justifications – justifications using data such as graphs, results tables, numerical values, or research papers

We used an automated coding process based on conjunctive keyword and expression matches that was previously developed and validated (Bagley & Shaffer, 2015; Nash & Shaffer, 2011). Next, we analyzed the coded chats of relative experts and novices using ENA to measure the development of connections among elements of the urban planning epistemic frame. Then, we analyzed correlation structures in the relative frequencies with which students used these elements by PCA.

Epistemic network analysis (ENA)

Epistemic Network Analysis (ENA) is a method of identifying and quantifying connections among elements in coded data and representing them in dynamic network models. ENA enables researchers to compare networks, both visually and through summary statistics that reflect the weighted structure of connections.

To create network models of individual students' discourse, ENA creates a series of adjacency matrices for each student. Each adjacency matrix represents the co-occurrence of codes in one students' discourse during a single activity. The adjacency matrices are binary, meaning that co-occurrences of codes are indicated simply as present or absent: If two codes co-occur in an activity, a 1 is placed in the cell in the adjacency matrix for that activity that corresponds to the intersection of the two codes; cells for codes that do not co-occur in the activity

receive a 0. Binary accumulation of co-occurrences is appropriate in this study because a student who says something twice as much does not necessarily understand it twice as well.

Each adjacency matrix thus represents the relations among the different urban planning concepts or practices made by one student in one activity, which means that each student is represented by a series of adjacency matrices. To identify the structure of connections made by each student, the adjacency matrices are summed into a cumulative adjacency matrix, where each cell represents the number of activities in which the unique pair of codes was present. The data set used for this study contained 17 activities coded for 18 urban planning epistemic frame elements. Thus, the cumulative adjacency matrix for a given student is the summation of 17 adjacency matrices, each with 153 (18 choose 2) possible unique co-occurrences of codes.

Once ENA creates the set of cumulative adjacency matrices for all the students in the data set, each matrix is converted into an adjacency vector by copying the cells from the upper diagonal of the matrix row by row into a single vector. These vectors exist in a high-dimensional space such that each dimension represents a unique pairing of two codes.

ENA spherically normalizes the adjacency vectors to calculate the relative frequencies of co-occurrence. In this high-dimensional ENA space, each adjacency vector represents the pattern of associations of a single unit, and the length of a vector is potentially affected by the number of activities that are contained in the student's discourse. More activities are likely to produce more co-occurrences, which results in longer vectors. This is problematic because two vectors may represent the same patterns of association, and thus point in the same direction, but represent different numbers of activities, and thus have different lengths. ENA solves this problem by spherically normalizing the vectors to unit Euclidean length. The resulting normalized vectors thus quantify for a student the relative frequencies of co-occurrence of codes independent of the number of activities in the model for any given student.

ENA then performs a dimensional reduction via singular value decomposition. (A singular value decomposition is similar to a principal components analysis, but it does not rescale the data.) This provides a rotation of the original high-dimensional ENA space, such that the reduced number of dimensions in the rotated space capture the maximum variance in the data. For every student in the data, ENA creates a point that is the location of the normalized vector under the singular value decomposition.

Finally, ENA positions the network nodes. Normally with dimensional reduction techniques, the basis vectors, or the *loadings*, would tell us how to interpret the positions of points in the space. However, these basis vectors represent connections between codes in the original data. That is, each point represents one of the cells in the cumulative adjacency matrix. That makes it hard to interpret, because if we have 18 codes, as in this study, then we can have up to 153 basis vectors, each of which corresponds to a unique co-occurrence of codes. To interpret the dimensions of this rotated space, ENA uses an optimization routine to position nodes in ENA space such that for any student, the centroid of the student's network model, corresponding to the student's cumulative adjacency matrix, is as close as possible to the location of the projected point. The utility of this correspondence is that we can compare the structures of networks by comparing the locations of their projected points. Projected points are positioned such that their associated networks will have their strongest connections distributed as weightings relative to the positions their projected points. Thus, the position of a point in ENA space summarizes the structure of connections in the networks being modeled.

The result is that: the ENA dimensional reduction models the variance among the different networks being analyzed; the corresponding network graphs allow us to interpret the significance of the locations of points in the ENA model; and we can interpret what aspects of the network structure explain the differences between units in the model.

Principal components analysis (PCA)

We computed the frequencies of each student's codes in the logfile and performed a principal components analysis using the R language's `prcomp` function. The code frequencies were standardized such that the variance in the frequencies of each code was equal to one, and the dimensionality of the code frequency space was reduced by singular value decomposition. Student code frequencies were then projected into the reduced principal component space.

Results

RQ1: Are there statistically significant differences between novices' and relative experts' local correlation structures that can be detected using ENA?

We computed the structure of connections made by each student using ENA and projected them into the space created by the dimensional reduction. Figure 1 shows the projected points of each student's network in ENA space. The projected points showed statistically significant differences between novices (blue) and relative experts (red) on both the first dimension ($\text{mean}_{\text{expert}} = -0.123$, $\text{mean}_{\text{novice}} = 0.088$; $t = 7.446$, $p < 0.001$, Cohen's $d = 0.969$) and the second dimension ($\text{mean}_{\text{expert}} = -0.023$, $\text{mean}_{\text{novice}} = 0.032$; $t = -2.031$, $p = 0.043$, Cohen's $d = -0.258$).

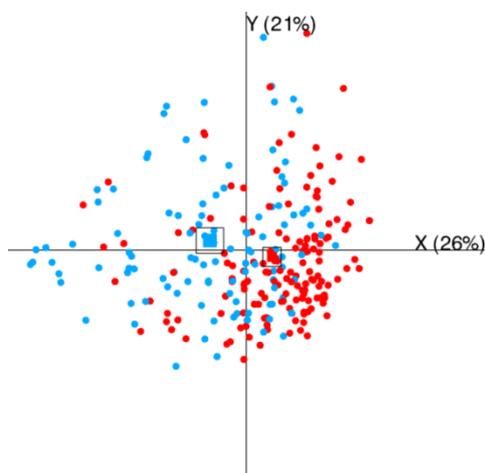


Figure 1. ENA scatter plot showing novice (blue) and relative expert (red). Each point is a single student; the squares are group means; the boxes are 95% confidence intervals (t-distribution) on each dimension; the numbers in parentheses indicate the percentage of variance in the data accounted for by that dimension.

To investigate which connections accounted for the differences between the two groups, we compared their mean epistemic networks. One of the most salient differences between the mean networks was in connections to data-based justifications (lower right in Figure 2; complementary colors indicate connections with data-based justifications). The novice network (blue) showed that data-based justifications were connected only with knowledge elements, while the relative expert network showed that data-based justifications were connected with knowledge elements, skills and actions, and other justification codes. In other words, relative experts made more and more diverse connections to data-based justifications than novices. To understand the meaning of these differences in local correlation structure, we analyzed instances of data-based justification in the student chats.

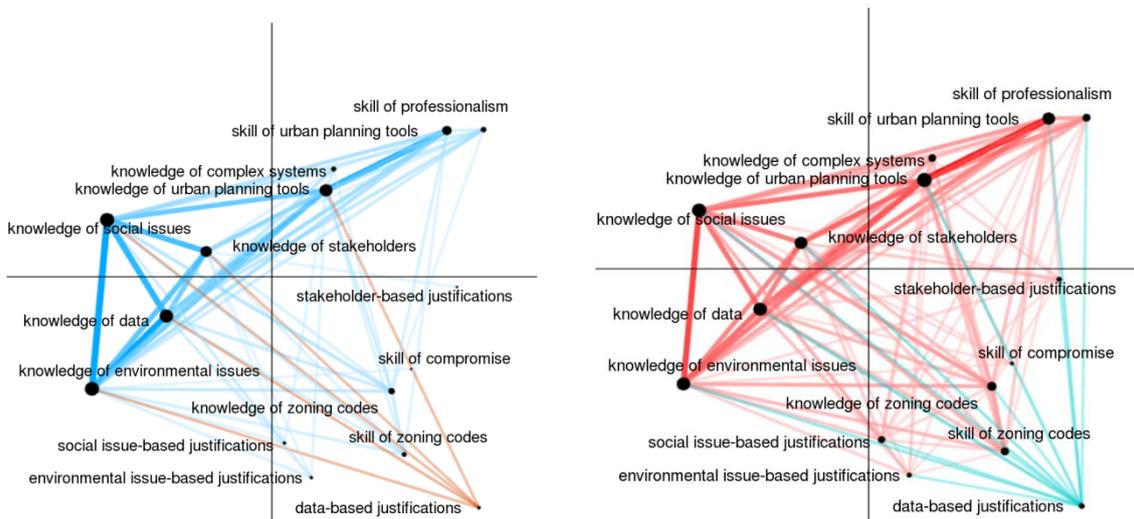


Figure 2. Mean ENA network diagrams showing the connections made by the two groups of students described in Figure 1. Novices (left) connected data-based justifications with knowledge elements; relative experts (right) connected data-based justifications with knowledge elements as well as with skills and other justifications.

RQ2: Are these differences meaningful on closer qualitative analysis of the data?

In one activity in *Land Science*, students are introduced to *iPlan*, the geographic information system tool with which they model how zoning changes affect the socioeconomic and environmental indicators that the stakeholders care about. Students made changes to their plans and observed the effects of changes on the indicator graphs. After finalizing their changes, students used the chat tool to participate in a reflective discussion with their assigned mentors.

In the chat discussions, novices made connections primarily between (1) knowledge of urban planning concepts and practices and (2) data-based justifications. One novice (below) explained the stakeholder Hao’s request (red, knowledge of stakeholder representation) and explained the warrant for this request (blue italics: data, blue underline: the warrant).

Hao says that *the number of Baltimore Orioles is decreasing as a result of the development of the town*. She claims that an environment that is not healthy for birds is not healthy for us.

This connection between stakeholder representation and warranting a stakeholder request on the basis of data obtained in *iPlan*’s graph indicator for animal life population is an example of novices justifying domain-relevant knowledge using data.

While novices connected data-based justifications with domain-relevant knowledge, relative experts connected data-based justifications with a wider range of the domain’s concepts and practices, indicating a more sophisticated grasp of the domain. For example (below), in the same activity, a relative expert asserted that, whatever her team’s next changes in their plan might be, they must increase the amount of housing (red, skills with urban planning tools). The expert justifies her assertion (blue italics: data-based warrant) by appeal to socioeconomic issues (population), environmental issues (runoff), which she can know only by making adjustments with *iPlan* and observing changes in graph indicators.

Going back to Colby’s original question, I think *the plan may be forced to increase housing* anyways, *due to the increasing population, and the runoff is inevitable*.

Thus, in this example, the relative expert built a data-based argument for a design decision in the domain. In contrast, the novice made a data-based argument to justify a specific piece of knowledge he had acquired. The same distinction is seen as well for other types of justification than those based on data. Figure 3 depicts the same mean networks as Figure 2, but the only connections shown are those to justification codes. As was the case with

data-based justifications, relative novices connect other types of justifications primarily with knowledge elements, whereas relative experts connect justifications with knowledge, skills and actions, and other justification codes.

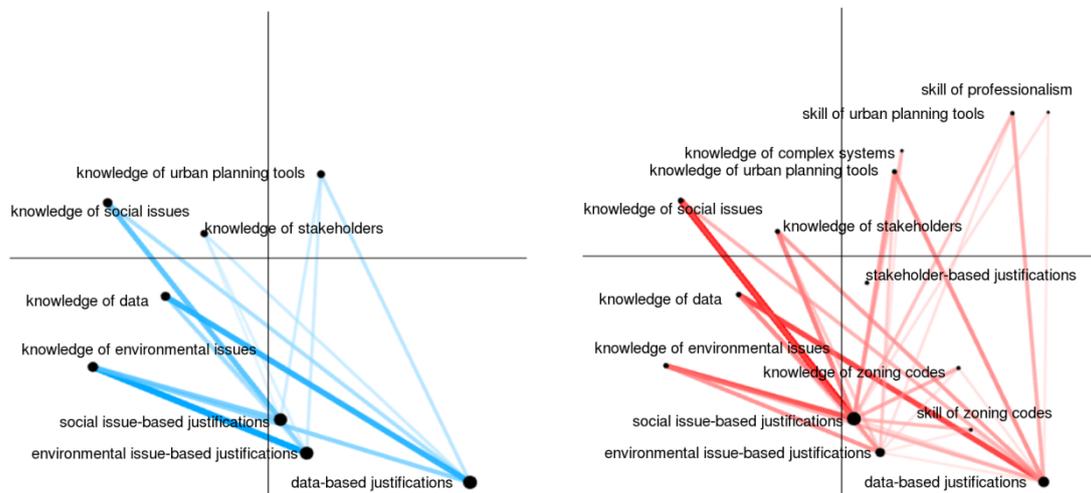


Figure 3. Mean ENA network diagrams of novices (left) and relative experts (right) showing only connections to justification codes. All other connections are hidden.

RQ3: Does PCA detect these same differences in local correlation structure by measuring global correlation structures?

We investigated whether there were differences in the correlation structures of the code frequencies of novices and relative experts using a common dimensional reduction technique, PCA. We computed students' code frequencies, created a reduced space using PCA, and projected the code frequencies into this space. Significance tests did not show significant differences between the novices' and relative experts projections on either PCA component one ($\text{mean}_{\text{novice}} = -0.0814$, $\text{mean}_{\text{expert}} = -0.2252$; $t = 0.6625$, $p = 0.5083$) or PCA component two ($\text{mean}_{\text{novice}} = -0.08983$, $\text{mean}_{\text{expert}} = 0.2529$; $t = -1.5774$, $p = 0.1159$). In other words, the result obtained above using ENA that models using the local nature of students' cognitive connections in discourse is not shown by PCA, which relies on correlation structures in code frequencies overall in each students' data.

Discussion

Our results showed that ENA was able to detect meaningful differences in the logged discourse of *Land Science* where PCA was not. Our PCA approach, which measured the global correlation structures in the frequencies of concepts and practices from the urban planning domain as they appeared in the student discourse, was unable to distinguish the novices from relative experts. ENA, which measured concepts' local correlation structures within activities, was able to make this distinction. Moreover, these distinctions were meaningful in a qualitative analysis of the data. In sum, our results support the claims of diSessa, Shaffer, and others, that student discourse is appropriately analyzed using correlation structures, and more appropriately analyzed using correlation structures that are sensitive to local contexts in which concepts are connected to one another.

This study had several limitations. First, we compared one example of an approach which measures local correlation structures and one example of an approach that measures global correlation structures, although PCA is a very common tool used for analysis of global correlation structure and there are few examples besides ENA of techniques that systematically model local correlation structure in discourse. There is also, of course, the problem that this study is based on the analysis of only one data source. Despite these limitations, however, the work here suggests local correlation structure-based approaches may be more appropriate than traditional global correlation structure-based methods for assessing that assessment student discourse.

References

Alper, B., Bach, B., Riche, N.H., Isenberg, T., & Fekete, J. (2013). Weighted graph comparison techniques for brain connectivity analysis. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13. (pp. 484-492). New York, NY, USA. ACM.

- Bagley, E. (2010). The epistemography of an urban and regional planning practicum: Appropriation in the face of resistance. WCER Working Paper 2010-8. Wisconsin Center for Education Research, University of Wisconsin-Madison.
- Bagley, E. & Shaffer, D. W. (2015). Learning in an urban and regional planning practicum: The view from educational ethnography. *Journal of Interactive Learning Research*, 26(4), 369-393.
- Beishuizen, J. J., Hof, E., van Putten, C. M., Bouwmeester, S. & Asscher, J. J. (2001). Students' and teachers' cognitions about good teachers. *British Journal of Educational Psychology*, 71, 185-201.
- Boulos, M. N. K., Maramba, I. & Wheeler, S. (2006). Wikis, blogs and podcasts: a new generation of web-based tools for virtual collaborative clinical practice and education. *BMC Medical Education*, 41(6).
- Bransford, J. D., Brown, A.L., & Cocking, R.R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press, Washington, D.C.
- Chi, M. T. H., Feltovich, P. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121-152.
- Chi, M.H.T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271-315.
- Cudek, R. (1989) Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, 105(2), 317-327.
- diSessa, A. A. (1988). *Knowledge in Pieces*. In Forman, G. & Pufall (Eds). *Constructivism in the Computer Age*, New Jersey: Lawrence Erlbaum Publishers.
- Elmqvist, N., Do, T. N., Goodell, H., Henry, N., & Fekete, J. D. (2008). ZAME: Interactive large-scale graph visualization. In *Proceedings of IEEE Pacific Visualization Symposium* (pp. 215-222). IEEE.
- Gernsbacher, M.A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 217-263). New York: Academic Press.
- Graesser, A. C., Gernsbacher, M. A., & Goldman, S.R. (1997). Cognition. In T. A. van Dijk (Ed.), *Discourse: A multidisciplinary introduction* (pp. 292-319). London: Sage.
- Hall, C. (1977). Dimensional reduction analysis. *Journal of Experimental Education*, 45(4), 4-8.
- Hmelo-Silver, C.E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16(3), 235-266.
- McNeil, N.M., Uttal, D.H., & Sternberg, R.J. (2009). Should you show me the money? Concrete object both hurt and help performance on mathematics problems. *Learning and Instruction*, 19(2), 171-187.
- Nash, P. & Shaffer, D.W. (2011). Mentor modeling: The internalization of modeled professional thinking in an epistemic game. *Journal of Computer Assisted Learning*, 27(2), 173-189.
- Peled, A., & Rashty, D. (1999). Logging for Success: advancing the use of WWW logs to improve computer mediated distance learning. *Journal of Educational Computing Research*, 21(4), 413-431.
- Shaffer, D. W. (2004). Epistemic frames and islands of expertise: Learning from infusion experiences. In *Proceedings of the 6th international conference on Learning Sciences* (pp. 473-480). Santa Monica, CA: International Society of the Learning Sciences.
- Shaffer, D. W. (2006). Epistemic frames for epistemic games. *Computers and Education*, 46(3), 223-234.
- Shaffer, D. W. (2008). *How computer games help children learn*. New York, NY: Palgrave Macmillan.
- Shaffer, D. W., Hatfield, D. L., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E. A., & Frank, K. (2009). *Epistemic Network Analysis: A prototype for 21st century assessment of learning*. *International Journal of Learning and Media*, 1(1), 1-21.
- Shaffer, D. W. (2014). *User guide for Epistemic Network Analysis web version 3.3*. Madison, WI: Games and Professional Simulations Technical Report 2014-1.

Acknowledgments

This work was funded in part by the National Science Foundation (DRL-0918409, DRL-0946372, DRL-1247262, DRL-1418288, DUE-0919347, DUE-1225885, EEC-1232656, EEC-1340402, REC-0347000), the MacArthur Foundation, the Spencer Foundation, the Wisconsin Alumni Research Foundation, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.

Advancing Learning Visualizations: Situated Action Networks as Scalable Representations of Learning in Social Settings

Alejandro Andrade, Indiana University, laandrad@indiana.edu

Rafi Santo, Indiana University, rsanto@indiana.edu

Abstract: We propose a distinctive method, Situated Action Networks (SANs), rooted in socio-cultural theories of learning that affords visualization and analysis of learning in a way that is theoretically robust yet scalable to large data sets. While visualization is increasingly looked to as a key means of understanding learning, there are few tools at learning scientists' disposal that are simultaneously scalable yet also aligned with sociocultural perspectives. Situated Action Networks attempt to address this by appropriating techniques from social network analysis while aligning them with Cultural Historical Activity Theory. They accomplish this by (1) elevating learning activities to the forefront of learning visualizations, allowing for rich qualitative analyses of learning and (2) creating theoretically aligned indices that afford quantitative analyses within and across learning environments. Using data on collaborative learning dynamics between informal learning organizations as they engage in joint projects, we show the affordances of this method for understanding learning.

Keywords: learning visualization, Cultural-Historical Activity Theory, Social Network Analysis, Situated-Action Networks, sociocultural learning theory, sociocultural learning methodologies

Introduction

Analysis and visualization of large-scale learning data is an increasingly viable and powerful means of understanding the dynamics and environments that support learning activity (Siemens & Baker, 2012). In this paper we propose and show the affordances of a distinctive method, Situated Action Networks (SANs) that can be applied to these ends in a way that is aligned with sociocultural theories of learning. We aim to fill a key gap—many learning visualization techniques that operate at scale are not aligned with sociocultural learning perspectives, and those currently used by sociocultural learning theories cannot operate at scale and have limitations even in their affordances when it comes to qualitative analyses.

SAN appropriates methods from Social Network Analysis (Marin, Wellman, Scott, & Carrington, 2011) and retrofits them in a way that aligns with and solves a number of analytic challenges (Spinuzzi, 2011; Witte & Haas, 2005) associated with Cultural-Historical Activity Theory (CHAT, Cole, 1996; Engeström, 1987; Vygotsky, 1978). To illustrate this technique, we use data on collaborative learning dynamics between informal learning organizations as they engage in joint projects, and we empirically show the affordances of this method for understanding learning. We find that SANs have two key affordances. They (1) elevate learning activities to the forefront of learning visualizations, allowing for rich qualitative analyses of learning and (2) create theoretically aligned indices that afford quantitative analyses within and across learning environments.

Elevating activity and mediational means within CHAT representations

As the study of learning has shifted from looking at learning as a phenomenon specifically concerned with individuals' minds to a process that is fundamentally rooted in social processes (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991), researchers have turned to sociocultural theories of knowledge and learning. CHAT (Cole, 1996; Engeström, 1987; Vygotsky, 1978) provides a sound theoretical framework to ground the interpretation of small-group collaborative learning situations, as well as an expansive lens towards larger transformations of social practices. A central tenet in CHAT is its focus on mediation and object-directed activity. Mediated activity has been conceptualized as a triangle representing a higher-level/cultural path between the subject—i.e., the collective or individual participant—and the object—i.e., the intended goal or motive of the activity. CHAT triangles have been useful in that they have provided a theoretically robust framework that allows rich analyses of complex socio-cultural contexts while, at the same time, provide a common grammar for many sociocultural learning theorists.

However, various researchers (Spinuzzi, 2011; Witte & Haas, 2005) have questioned its methodological utility. For if the object of an activity is not empirically bounded, then it runs away, merging itself with other activity systems or as part of larger activity systems (Spinuzzi, 2011; Witte & Haas, 2005). In a sense, learning activity is invisible as it is everywhere and nowhere. Further, the data is not scalable as it does not afford an easy

way for comparisons. To overcome this limitation, various alternative approaches have been proposed (for instance see Spinuzzi, 2011), but in this paper we align our own with Witte's (2005) proposal to focus on mediational means as the primary methodological tool to leverage our understanding of learning. We can think of two reasons for doing this. First, we want to observe and study mediational means to compare and contrast across activity systems or the change of one activity system over time. Second, the very nature of a human motive cannot be observed directly, but has to be inferred from the participants' actions and/or from intersubjective accounts. Therefore, we believe a promising approach to visualize learning should start by identifying the concrete, observable mediational means, and then test the hypotheses about the possible motives/subjective goals that underlie that activity. By focusing on the mediational means at the outset, without pre-defining an assumption about the relationship between mediational means and goals, we can more accurately model the complexity of activity systems in situ.

Our Situated Action Networks (SAN) approach (Andrade, 2015) tries to strengthen this sociocultural tradition by drawing on Social Network Analysis and retrofitting it to make it more theoretically aligned with CHAT. Social Network Analysis is useful for creating visualizations and quantitative descriptors of social structures. Although other approaches have either used or adapted social network analysis to analyze learners' data (see i.e., Oshima, Oshima, & Matsuzawa, 2012; Shaffer et al., 2009; Suthers, 2011), SAN provides an explicit account of how to integrate social network measures and reframe them to be meaningful within a CHAT framework. In a previous paper, Andrade (2015) introduced the approach and some of the technical details using the example of a pilot apprentice learning how to land an airplane in a computer simulation. Briefly, SAN is intended as a functional model that takes action as the unit of analysis. When action is the unit of analysis, participants and tools (or as we refer here, mediational means) are placed in the model at the same level, that is, as nodes in the graph. The actions represent edges (or links) between one actor node and one mediational mean node. SAN models the activity system by representing the functional link between the social structure and the mediational means. Goals and motives are then hypothesized from the observed actions. SAN aims to take best of both worlds—rich and theoretically rigorous representations of learning on the one hand and scalable techniques for representation and analysis on the other. In doing so, it maintains the ability to look qualitatively at individual or small numbers of representations, in combination with additional qualitative data of the context in question, facilitating rich analyses of learning activity. But it also adds the ability to quantitatively analyze, through use of indices, across multiple representations in ways that allow for powerful pattern analyses.

As our early work focused on short timescales and micro-analysis (see Andrade, 2015), this paper focuses on application of the technique to longer timescales and meso-level analyses with data from informal learning organizations, which illustrates how this method is flexible and scalable. First, we outline the methods.

Methods

As this paper is primarily methodological in nature, we outline the process we went through as we developed the SAN technique, including the rules used to visualize activity and the attempts at creating theoretically meaningful indices that can be used for quantitative analysis. Our technique was one of methodological exploration; we took early experiments conducted by the first author to develop SANs (Andrade, 2015) and actively built on them using a larger data set. We engaged in the following activities as part of that process, each of which we detail in depth within the findings: (1) experimented with a variety of representational forms to see what each afforded in terms of analysis, (2) conducted pilot qualitative analyses on various SANs, (3) brainstormed various quantitative indices that we saw as theoretically aligned with CHAT and (4) conducted interpretation and analysis based on the application of indices to empirical data.

The data we used comes from Hive Research Lab (Santo, Peppler, Ching, & Hoadley, 2015), where we analyzed 94 collaborative project proposals produced by out of school learning organizations, members of the Hive NYC Learning Network (hivenyc.org), focusing on developing educational initiatives that align with Connected Learning approaches (Ito et al., 2013). Supplementary qualitative data included 20 one hour interviews with out of school educators discussing their collaborative work. The proposals were submitted in response requests for proposals by the Hive Digital Media and Learning Fund, a collaborative donor fund, over the course of a four year period, 2012-2015. The projects were always collaborative, with at least two organizations playing roles within a project.

As we will share in more detail in the findings section, the project proposals and interviews were first analyzed in order to develop a coding scheme of the mediational means these organizations used within projects in order to achieve project goals. One researcher coded the full corpus of proposals while a second coded 30%. The researchers discussed disagreements, which helped further refine the coding scheme. Although a second round of discussions to find interrater agreement is still under way, we are confident in the face validity of the coding scheme as we performed member checking and included organization member's feedback. The coded

proposals formed the basis for the generation of SANs which were then subjected to further analysis detailed below.

Constructing situated action network visualizations

We delineate a series of steps involved in creating a SAN. First, tally the occurrence of actions, that is, the use of mediational means by actors. In practice, this means tallying up who is doing what action. For instance, we tallied up which organizations were designing the educational initiative and which were evaluating it (see table 1). Second, bound the activity system. This sometimes is easy, as is our case here because the project submissions defined which organizations were involved in what. Sometimes this requires a little more work, as the object or other elements in the system change. Sometimes, when studying how the system transforms through time, the analyst has to meaningfully decide boundaries. For instance, in previous work when we studied the way an apprentice learned how to land a plane, each attempt was regarded as its own SAN (Andrade, 2015). Third, preprocess the raw data into an appropriate format. This procedural step transforms the raw data frame into a square symmetric matrix in which rows and columns are actors and mediational means, and the entries to the matrix are tallies of the links between actors and means (see Figure 1.a). This matrix looks very much like a regular adjacency matrix, but it is a very sparse matrix with many zeros (on the diagonal, and between actor-actor and means-means entries). Fourth, use a computational package to visualize the graph (see Figure 1.b). We use *igraph* in R for this. Some additional procedural decisions are involved here. First, different node shapes can differentiate between actors and means (e.g., rectangles and circles). Colors can differentiate types of mediational means (e.g., community or design oriented) or kinds of actors (e.g., lead or non-lead). Edges can have different widths according to tallies and/or weights assigned to them. Fifth, try different visualizations to help refine the displayed elements and try different configurations so that particular pieces are highlighted, get peer feedback, and revise the visualization.

Table 1. Raw Data Matrix

Organization	Project Title	Year	Design	Evaluation
Science Exploration Center (SEC)	Hive Awesomeness	2014	1	0
SciMapping (Sci)	Hive Awesomeness	2014	1	1



Figure 1. Left: (a) Adjacency Matrix. Right: (b) Visual Representation.

Findings

Qualitative affordances of situated action networks

In exploring the methodological affordances of SANs, we will first examine how they can support various forms of qualitative analyses and theory building. We see three specific analytic affordances from a qualitative perspective that SANs support: (1) prompting the identification of mediational means within a context, (2) identifying new qualitative patterns by looking *across* SANs, and (3) identifying new qualitative patterns by looking *within* SANs. Below, we explore each of these affordances.

SAN approach as prompt to identify mediational means within a context

Prior to and in order to create SANs, the method asks a researcher to identify the primary mediational means within the context of investigation. To accomplish this, we first began with an in-depth review of the study's qualitative data. We identified particular resources that organizations used as mediational means to achieve specific goals, means that together supported the project's broader object(s). For example, in some projects an organization would utilize design expertise that would shape the curriculum associated with an

afterschool program, another organization would utilize its existing networks to recruit educators that would participate in professional development, and a third might provide the physical space for program. In all, 21 mediational means were identified across the 94 project proposals, falling into the broad categories of expertise-related (leveraging specialized knowledge or competencies), network-related (leveraging access to particular networks of either young people or educators) and resource-related (leveraging provision of material or intellectual property). For each of the codes, an associated description and set of examples from the data sources was compiled into a coding guide.

This process took tacit understandings of the data and formalized them into a consistent list of mediational means that were then clear to the research team. Following this step, the coded data were then used to generate SANs. Additionally, through the full application of the codes to the data, the research team gained an additional ‘felt sense’ of how certain mediational means were used in relation to one another, an understanding that played a critical role in unearthing further qualitative insights that we detail next.

Identifying new qualitative patterns by looking across SANs

Once the full data set was coded and the SANs generated, the research team then sat down and began to look across the SANs to make sense of them. During this process, we utilized qualitative understandings gained both through our fieldwork as well as the process of generating and applying the mediation means codes to the data. When we say ‘looked’ here, we mean that in a very literal sense—the similarities and differences between SANs were brought into sharp relief through the visual representations with certain structural features clearly standing out. As we looked across the SANs, initially tacit hypotheses about common configurations of mediational means were clarified through these empirical representations.

Specifically, we looked to see how hypotheses regarding common objects that might be achieved by certain collaborative configurations might be confirmed, augmented and highlighted through the SANs. Put simply, we sought to see what kinds of broad goals were trying to be achieved through distinctive forms of organizational collaboration, and how these might be characterized into typical models. Upon investigation of the SANs in relation to these hypotheses, six new categories of objects, with associated collaborative partnership configurations, were identified, each representing different objects and utilization of mediational means. Below, we detail two of these categories. For each, we first describe the nature of the distinctive object, highlight how the visual features of that SAN assisted the research team in identifying them, and engage in a brief theoretical treatment of the SAN.

Collaborative design and implementation partnerships

One of the first configurations immediately identified by the research team was that of **collaborative design and implementation**, wherein many or all of the mediational means relating to the creation, development, execution and reflection are shared amongst key actors. The ‘starfish’ visual structure of this kind of configuration was immediately evident across a range of the SANs, and mapped immediately onto qualitative understandings of our team. In the project presented in Figure 2 below, two organizations, the Reese Music Institute and Rhythm Central are all co-engaged in their use of mediational means within a project that involves a series of day-long workshops that engage youth in digital hip hop music production and remix. See Figure 3 below for the legend.

Within the SAN, edges (connecting lines) indicate which actors (circles) are engaged in which mediational means (hexagons), with thicker lines indicating substantive action and thinner lines indicating lighter action. Here, the fact that all mediational means nodes are shared by both actors (indicated by edges going from each actor to all action nodes) shows a high degree of shared object across actors and mutual involvement in transforming all facets of the activity system at hand—both of the organizations are equally involved in the design process, internal training for teaching artists, recruitment of youth, facilitation of the program, etc. The overall structure of visual parallelism in the representation reflects this shared object and joint involvement in transforming the associated activity system.

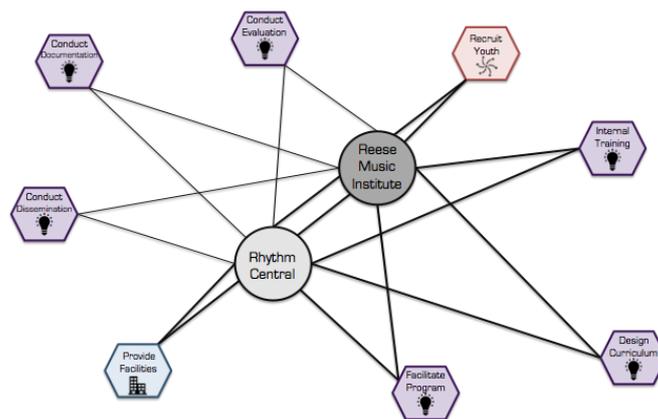


Figure 2. Collaborative Design and Implementation Partnership.

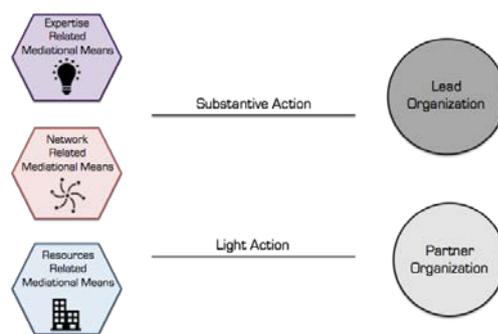


Figure 3. Legend for Situated Action Network visualizations.

Program refinement and spread partnership

As a contrast to the above where mediational means are largely shared across actors, the second configuration we share, **program refinement and spread partnership**, highlights a highly specialized division of labor. Like the first example, this configuration was immediately evident through the highly visible clusters of specialized activity. In this configuration, the lead organization is aiming to circulate a tool and curriculum to new contexts, and collaborating organizations act as sites of adoption. In this example shown in Figure 4, the Science Exploration Center (SEC) is aiming to refine and spread a program it developed that focused on neighborhood-based citizen science that uses sensor technologies to have youth engage in environmental investigations and advocacy. SEC utilizes a number of mediational means relating to supporting other organizations, Children’s Benefit Group and Whitman College, to act as adoption sites. SEC designs curriculum, creates additional materials to support adoption such as rubrics and educator tip sheets, and then trains the adopting organizations through professional development activities. The adopting organizations similarly have a specialized cluster of activities—providing networks of youth and educators, providing the spaces where the program will be implemented, and engaging in the actual facilitation of the program. Two other organizations, SciMapping and the Teaching Capacity Institute (TCI), each play distinct roles around technology development and formative research and refinement activities, respectively. These supporting organizations make viable, along with SEC, the process of model circulation to new activity systems associated with adopting organizations.

From a CHAT perspective, this configuration is characterized by a highly specialized division of labor, as indicated by the various unconnected clusters. While of course each of the actors is working on the same larger project and shared object, each is leveraging specific mediational means it has access to that tackle a distinctive aspect of the work, impacting different parts of the related activity system(s) and indexing different degree of agency and ownership over the learning process. SciMapping is only working on the technological aspect of the initiative, while both SEC and TCI are involved in the larger tool of the curricular model in which this technology is being utilized. TCI plays a role providing feedback loops, through formative research, about how the project is playing out in new activity systems where it is being adopted, supporting refinement of the tool, but is in a technical assistance role here that has somewhat less ownership over the object—it is assisting SEC and SciMapping with their broader shared object of bringing the program to new contexts.

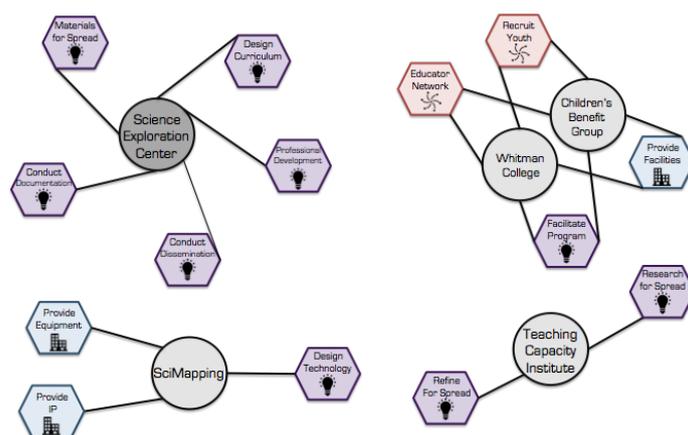


Figure 4. Program Refinement and Spread Partnership.

Each of the configurations shared above, along with others, was identified through a process of drawing on rich qualitative understandings of underlying data in combination with a visual examination across the 94 action networks generated. This sense-making led to the identification of new categories that helped us understand what kinds of problems these collaborative configurations were solving. From a CHAT perspective, we were able to understand what kinds of objects this collective of SANs were addressing.

Identifying new qualitative patterns by looking within SANs

A final qualitative affordance of Situated Action Networks lies in the way they support identification of new qualitative patterns, and associated theoretical implications, by looking closely at a single SAN (as opposed to across multiple SANs). Close attention to specific SANs yielded consequential findings related to the nature of activity systems that spanned multiple organizations within a SAN. Specifically, we found that multiple related but distinct local objects can be at play within an activity system.

The example below took the form of an implementation site partnership, a particular configuration of mediational means noted briefly above. In this instance, the Sync Institute was the central organization, enacting its PhysComp Tinkering program at the Brooklyn Neighborhood Center (BNC) which acted as an implementation site. The Sync Institute engaged in a wide range of critical mediational means, including curricular design, technology design, program implementation, documentation, evaluation, and others. BNC simply provided its space and assisted in recruitment activities. It might be expected that the central learning process involved in such a collaboration would be centered on the Sync Institute's advancement of its prototype physical computing learning tool that it was developing to support beginner level hardware programming, and indeed, in interviews Sync Institute staff members shared that the implementation at BNC was contributing to a larger process of refining this invention. But interviews with BNC's staff revealed that their participation as an implementation site was actually part of a much larger and ongoing learning process the organization was engaged in and that it planned to create a maker space. Represented within this single SAN were both Sync Institute's object of learning in relation to its prototype as well as BNC's object of learning about if and how its young people and staff would be engaged and interested in 'maker' learning activities.

As the example shows and put most simply, organizations can hold differing, but overlapping, local objects as they participate within the same SAN. This finding is supported by the reality that it is rare for two organizations to come to a collaboration with the same prior knowledge and experience, and rarely are the exact same mediational means engaged in by two organizations in a collaboration. Looking closely at the SAN with rich qualitative data in hand allowed the research team to uncover and empirically examine this phenomenon of parallel objects within a SAN, of the unbounded and overlapping nature of activity systems, how an object in one system can be a tool in another, and other important nuances. This all points to the need for careful inquiry into the mediational means as an approach to bounding the activity system so that it can be empirically studied, elaborating on insights by Spinuzzi and Witte that have criticized CHAT methods. Having established multiple qualitative affordances of SANs, we now move to the quantitative affordances of the approach.

Quantitative affordances of SANs: Network indices supporting analysis at scale

Previous adaptations—e.g., Epistemic Network Analysis (Shaffer et al., 2009), Knowledge Building Discourse Analysis (Oshima et al., 2012), Traces (Suthers, 2011)—have made use of traditional measures to describe the characteristics of a network such as centrality, in- and out-degree, or path sizes. We propose new coefficients

because traditional SNA descriptors are not appropriate for SAN—in that adjacency matrices are sparse as there are no connections within the same type of nodes (i.e., actor-actor or mediator-mediator). We propose three kinds of indices that describe different properties of the SAN: (a) global, (b) average egocentric, and (c) individual actor or mediator descriptors. Global indices describe the distribution of actions among participants, and differentiate among types of contributions by particular actors or mediators. By global descriptors we mean overall network composition measures. Average egocentric descriptors do not capture the distribution of actions among actors, instead they reveal the overall number of actors, means, and actions in the network. On the other hand, an egocentric descriptor is an individual node measure, for instance, the number of other connected nodes or the sum of its edges (for weighted networks). The arithmetic average of an egocentric index describes the whole action network. In particular, we include two types of average egocentric descriptors. First, we pay attention to the extent mediational means are used across actors, called *shareness*. We chose this label to indicate that this index represents the average use of mediational means, that is, how much they are shared. Second, we pay attention to the extent to which actors act on various mediational means, called *participation*. This label represents the idea that the more an actor uses mediational means, the greater its participation is in the network. Taking the average, participation produces a general action network descriptor. Individual descriptors acknowledge that not all mediational means affect the course of the activity in the same way. For instance, one may want to distinguish between mediational means that have an effect on the activity object and those that affect other system elements, such as the community. Put another way, different mediational means shape the system in different ways, and therefore individual indices describe different combinations of mediators in the action network. Finally, one can use individual actor node indices. For instance, egocentric descriptors for the leading organization of the project provide valuable information about the way the project is enacted.

Table 2 SAN Descriptors

Descriptor Title	Theoretical Basis	Operationalization
Total # of Actors	Total number of actors in the network	# of actor nodes
Total # of Mediators	Total number of mediators in the network	# of mediators
Total # of Actions	Total number of weighted links in the network.	# of actions
# of Clusters	Number of connected components in the network (as originally intended in SNA). This index can inform whether there is specialization of some actor-mediator relationships.	
Average Shareness	Egocentric index measuring the degree to which a mediator is shared. It ranges between 0, for totally individual mediators, to the total number of actor nodes, for totally shared mediators. Normalized by dividing over the number of actors. A global measure is produced by taking the mean or median.	$(\text{Number of actions for } i^{\text{th}} \text{ Mediator node minus } 1) / \text{Total number of Actor nodes}$
Average Participation	Egocentric index measuring how busy an actor is. It ranges between 0, for totally free actors, to the total number of mediator nodes, for totally proactive actors (i.e., the actor participates in all mediators). Normalized by dividing over the number of mediators. A global measure can be produced by taking the average (mean or median).	$(\text{Number of links for } j^{\text{th}} \text{ Actor node minus } 1) / \text{Total number of Mediator nodes}$
Specific Mediator Leverage	It weights the number of specific mediational means to provide a proxy for the agency an actor has on a specific mediator of the activity.	$\sum [\text{number of } j^{\text{th}} \text{ actor actions} * \text{specific mediator weight}]$
Specific Actor Degree	% of actor actions upon mediational means in the network	$\# \text{ of actor actions} / \text{total} \# \text{ of actions}$

Table 3 shows the results of applying the global and average egocentric descriptors to the two networks described above. It can be seen that these indices capture the greater integration in the Collaborative Design and Implementation Partnership type, where all mediators are acted upon by the actors. It also captures the specialized nature of the Program Refinement and Spread Partnership type, where there are three distinct clusters and therefore less shareness and participation than the first type.

Table 3. SAN indices for two network types

Network Type	Actors	Mediators	Actions	Clusters	Shareness	Participation
Collaborative Design and Implementation Partnership	2	8	16	1	0.50	0.88
Program Refinement and Spread Partnership	5	13	18	3	0.10	0.20

Conclusions and implications

We introduce a visualization technique to explore sociocultural learning activity that is both scalable and theoretically rigorous. Previous CHAT representations, such as Engeström's triangle, while effective in delineating elements of a given activity, have very limited affordances to show distinctive relations within or across activity systems (cf. Witte & Hass, 2005; Spinuzzi, 2011). Essentially, they do not allow the researcher to consistently model the relationships between these elements in a way that is tailored to the particular situation. Situated Action Networks, as we have shown, produce robust visualizations of the observable activity system components—i.e., the actors and the mediational means these actors act upon. Through these visualizations, inferences can be made about the object of the activity (which is not directly observable), and how the mediational means interact with other vertices of the activity system. For instance, some mediational means are directed towards bounding the object of the activity, whereas others interact with the community within which the activity system is embedded. Furthermore, SANs can be described by quantitative indices, which would allow for meaningful scalable comparisons across activity systems or examine longitudinal changes of an activity system. In future work we intend to explore how statistical tools, such as regression and cluster analysis, can help find patterns across the collaborative projects presented here and the ways these indices relate to other key variables such as the size of the organizations involved and the intended learning outcomes.

References

- Andrade, A. (2015). Using Situated-Action Networks to Visualize Complex Learning. In Oskar Lindwall, Päivi Häkkinen, Timothy Koschmann, Pierre Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference* (Vol. 1). Gothenburg, Sweden: The International Society of the Learning Sciences.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*: Harvard University Press.
- Engeström, Y. (1987). *Learning by expanding. An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsulti Oy.
- Ito, M., Gutierrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., . . . Watkins, S. C. (2013). *Connected learning: An agenda for research and design*: Digital Media and Learning Research Hub.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*: Cambridge Univ Pr.
- Marin, A., Wellman, B., Scott, P., & Carrington, P. (2011). Social network analysis: An introduction. *Scott J, Carrington PJ, editors. The SAGE Handbook of Social Network Analysis*, 11-25.
- Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. *Educational technology research and development*, 60(5), 903-921. doi:10.1007/s11423-012-9265-2
- Santo, R., Peppler, K., Ching, D., & Hoadley, C. (2015). *Maybe a Maker Space? Organizational Learning about Maker Education within a Regional Out-of-School Network*. Paper presented at the FabLearn: Conference on Creativity and Fabrication in Education, Stanford, CA.
- Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., . . . Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning.
- Siemens, G., & Baker, R. (2012). *Learning analytics and educational data mining: towards communication and collaboration*. Paper presented at the Proceedings of the 2nd international conference on learning analytics and knowledge.
- Spinuzzi, C. (2011). Losing by Expanding Corraling the Runaway Object. *Journal of Business and Technical Communication*, 25(4), 449-486.
- Suthers, D. D. (2011). *Interaction, mediation, and ties: An analytic hierarchy for socio-technical systems*. Paper presented at the Hawaii International Conference on the System Sciences (HICSS-44), Hawaii
- Vygotsky, L. S. (1978). *Mind in society* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.): Cambridge, MA: Harvard University Press.
- Witte, S. P., & Haas, C. (2005). Research in activity: An analysis of speed bumps as mediational means. *Written Communication*, 22(2), 127-165.

Acknowledgments

This work was made possible through the participation of Hive NYC Learning Network member organizations and the generous funding of the Hive Digital Media and Learning Fund at the New York Community Trust.

Fighting for Desired Versions of a Future Self: Young African American Women's STEM-related Identity Negotiations in High School

Carrie D. Allen, University of Colorado Boulder, carrie.allen@colorado.edu
Margaret Eisenhart, University of Colorado Boulder, margaret.eisenhart@colorado.edu

Abstract: In this paper, we investigate how the national narrative of increasing opportunities for and broadening participation of young women of color in STEM was taken up locally at one racially-diverse, urban high school. Using ethnographic and longitudinal data, we focus on two young women of color as they negotiated and maintained STEM-related identities in the discursive and practice contexts of their lives at school. Using Holland and Lave's concept of history-in-person (2001), we view the young women as fighting for particular versions of a future self, while entangled in discursive and social relations that threatened to position them differently than they wished to be. Our findings suggest a need for an explicit naming and examination of the "double bind" that young women of color experience as they move through school environments and special support to prepare them for the challenges they may face in STEM-related college programs or workplaces.

Keywords: youth identities, STEM education, urban education, social practice

Introduction

Katie sat across from me (Allen), her face bright with enthusiasm, hands pressed against the tabletop as she leaned forward into the conversation. It was the summer after Katie's senior year in high school. We were sitting together outside a Starbucks near her home, and she was in mid-story explaining why she had decided to pursue a degree in engineering:

[In AP Calculus] we started doing these revolutions about y and x axes. If you look at a graph, and you take a three dimensional shape - like a cone or a cup - and you want to know the volume, you can put it on a graph and you can calculate the volume from revolving it around an x or y axis...it was a really hard math concept, but it was also one of the only exciting ones...it was just so exciting for me! So I started thinking, Where can I do this [kind of math] all the time?...I started talking to my dad and he was like, "Well engineering is somewhere where you could do this...if you want to do the math all the time, then go into engineering."

As an African American female, Katie's tenor as an excited and interested young woman in pursuit of math knowledge and an engineering career tells a story that contradicts the dominant historical narrative about women of color in STEM (science, technology, engineering, and mathematics) in the United States. This narrative positions females of color (Black, Latina, and Native American) as *disinterested*, *underprepared*, and unlikely to pursue STEM or succeed in these fields (Ong, Wright, Espinosa, & Orfield, 2010). Throughout high school, however, Katie took advanced STEM courses, achieved high grades in these classes, and in college she pursued a STEM-related degree. In fact, in our five-year longitudinal, ethnographic study, we observed a trend similar to Katie's across the majority (18) of our 23 female participants. These young women expressed interest in STEM-related fields at the start of high school and maintained that interest into their first year of college.

One way we could interpret the successes of these young women is to say that they were aware of and intentionally resisting the historical narrative regarding women of color in STEM that positions them as something other than what they were or wished to be. After all, their actions and self-descriptions reflect alternatives to the ways that national narratives position them. This interpretation is in fact how we approached our analysis initially. However, as we began to look more closely at the lives of these young women throughout high school, what we found instead was that their fight was not with this national narrative, but with local school narratives that negatively positioned students of color more broadly and remained silent on issues of gender, the intersection of gender and race, and the implications for STEM.

In this article, we illustrate the local struggles these young women engaged in to construct and maintain STEM-related identities in the context of their high school lives. In particular, we focus on the local discourses

and practices of one school learning environment within and against which two of the young women in the larger study engaged in STEM identity work.

Conceptual framework

Following Holland and Lave's development of the concept of "history in person" (2001), we view the young women in our study as engaged in a struggle with historical narratives about the trajectories of women and people of color within STEM. That is, when young women of color consider possible futures in STEM, their efforts can be viewed as a struggle or fight "for particular versions of the future" (Holland & Lave, 2001, p. 28), and for particular versions of a *future self*. When young women of color construct science, engineering, or math identities in school, they must do so in the presence of robust and enduring narratives and practices that may define them in ways they do not wish to be. Similar to other identity work, these self-negotiations are continually contested in local practice "as history is constituted in the space that encompasses both social participation and self-authoring" (Holland & Lave, 2001, p. 29). In this article, we examine how social participation and self-authoring in school can be 'pulled through time' into STEM interest, identity, and pursuit.

As one example of a historical structure of privilege, schools "infuse and restrain local practice" (p. 5) in ways that have consequences for how students come to view themselves and author selves within STEM. In particular, research points to academic, social, structural aspects of the school learning environment that negatively impact the experiences of women of color within STEM. For example, schooling practices can make it seem that science is the province of people who are privileged in society—upper middle class, White, and male (Bang, Warren, Rosebery, et al., 2013; Carlone, et al., 2014). School science is often treated as neutral with respect to cultural and social experiences (Bang et al., 2013), and instruction tends to focus on achieving established answers in uniform ways (Calabrese Barton et al., 2012; Carlone et al., 2014). Such practices shape students' views of *who* or *what* is "scientific" and whether they fit such depictions.

As young women negotiate STEM identities in their schools, they can also carve out spaces for new or different storylines (Holland et al., 1998). In their analysis of middle school girls' identity work across multiple social contexts (such as the classroom and an after-school program), Calabrese Barton and colleagues (2012) describe the case of Chantelle, who, by the end of the study had developed a self-identity as someone who might pursue a career in science. An African American girl and someone who aspired to be a dancer and singer, Chantelle was positioned by her teacher as "a student in the middle," "easy" because of her quiet and compliant behavior in class, and as someone who "struggled" with content, particularly in math. However, her active membership in an after-school science club supported Chantelle in successfully positioning herself as someone who was "hard working" and "engaged in science," as she presented to her class what she was learning about energy efficiency in the after-school club. Further, in their analysis of the identity work of successful women of color in science, Carlone and Johnson (2007) found that undergraduate women with an "altruistic" science identity broke from the historical script of those who pursue and succeed in science. By emphasizing their commitment to serving others over more familiar characteristics of a scientist, they created spaces for constructing a different kind of science identity.

Examining students' instantiations and negotiations of STEM selves within their learning environments offers one way to theorize how to better support females of color in STEM fields and in their pursuit of STEM careers, while highlighting the work these students must engage in to answer hegemonic institutional practices and the historical narratives that allow these practices to persist. This is our primary aim here. Further, by following young women's STEM identity negotiations through high school and into college, we have a view of STEM identity development during a crucial moment in young people's lives.

Methods

This article employs an embedded multiple-case methodology (Yin, 2013) to explore (1) the ways that the historical narrative regarding STEM and STEM pathways gets taken up in local discourse and practices at the young women's high schools; (2) the kinds of identities these women construct over the course of high school and into college; and (3) how these young women instantiated and negotiated STEM identities in the context of their schools.

Research context

The study described in this paper comes from a larger longitudinal and ethnographic study examining high school STEM opportunity structures and students' figured worlds of STEM (see Eisenhart et al, 2015) in Denver and Buffalo. The ethnographic study took place between 2010-2013 in eight high schools (4 in Denver, 4 in Buffalo) that served mostly students of color and mostly students on free or reduced-price lunch. Both authors participated

actively in collecting and analyzing these data. In 2014 and 2015, we conducted follow-up surveys with students to learn about their post-high school graduation experiences (college choice, major in college, work experience).

Near the start of the larger study, in late 2010, the research team recruited approximately 12 focal students (ages 15-16; 6 girls and 6 boys) from each of four high schools in each city. The intent was to follow closely the focal students' experiences in STEM from 10th through 12th grade. The focal students were selected from volunteers among the largely minority population of students who were in the top 20% of their high school class in math and science based on their grade point averages (GPA) and scores on state standardized proficiency tests after their first year of high school (9th grade). The research team chose high-achieving students in math and science because they seemed the most likely to be interested and to participate in STEM education opportunities offered by the high schools.

Participants and school site

We selected the two students highlighted in this article through a multi-stage process aimed at identifying representative cases of young women of color who succeeded in high-level STEM courses and anticipated pursuing a STEM-related college major. First, we created a data display for all female participants in the larger study; the data display included students' demographic information as well as characteristics we identified as important to students' STEM authoring and positioning in school. These included STEM courses taken, years of STEM courses taken, grades in STEM courses, anticipated college major from 10th-12th grade, post-secondary plans (as of grade 12), and actual post-secondary decisions (during their first year post-high school). From this table, we identified students who took STEM courses all four years of high school and who maintained interest in a STEM major through their senior year of high school. From this list, we then selected underrepresented students of color (Black or Latino students), excluding those who were White or Asian. This list included seven students from three of the four Denver schools. We selected for comparison cases 4 students (2 Latinas, 2 African Americans) from the two high schools that were most comparable in terms of graduation requirements, Advanced Placement (AP) and advanced course offerings, and strong school leadership: Capital and Southside. In this paper, we present the case of Katie and Naomi, two African American females from Capital who elected to study engineering in college.

Data collection and sources

Data for this analysis were collected over a 5-year period: students' second year of high school (grade 10, 2010-11) through their second year of college (2014-15). During the students' high school years (Years 1-3 of the study), the research team conducted face-to-face student, parent, and school personnel semi-structured interviews, observed selected math and science classrooms, and collected students' high school transcripts. Interviews typically lasted 45-60 minutes in length and were recorded and transcribed. During the students' first two years after high school, online surveys were administered to all student participants. These data sources are described in more detail below.

Students in the study were interviewed 5 times over the course of the three-year ethnographic study: once during the spring of 2011 (Year 1), and then in the fall and spring of following years (Years 2 & 3; 2011-2013). The Year 1 (students' 10th grade year) spring interview focused on establishing a baseline of students' perceptions of their schools and themselves as students at the school. Students were asked to describe their school, their academic strengths, the courses they were taking and how they decided on those, and what they intended to study in college. During Year 2 (students' 11th grade year), interview questions asked students specifically about their views of math and science, characteristics of those who are good at math and science, and students' views of themselves in relation to these ideas. In the spring of Year 2, students were additionally asked about their plans for college applications and college majors. Interviews in the final year shifted to college plans and preparation. Students were asked if they still intended to study a STEM-related major, what colleges they had applied to, and, in spring, what their final post-high school plans were. Because of their involvement in an earlier analysis, one of the participants discussed in this paper - Naomi - additionally participated in a follow-up phone interview during the spring of their freshman year in college (Year 4 of our study). This interview was intended to clarify their decision-making about what colleges to attend and majors to pursue; we additionally asked about what courses they were taking at the time.

Further, we conducted school-personnel interviews, which included math and science teachers, counselors, and school administrators (principals or assistant principals). Math and science teachers who taught one or more focal students at each school were interviewed once per year (Years 1-3), as were the focal students' counselors and school administrators. Interview questions asked about the recent history of the school, views of the school and its students, academic proficiency indicators, opportunities for high achieving students, post-high

school preparation, and graduation and college enrollment rates. We drew on these interviews as a way to identify and later substantiate claims about school practices, challenges, and influential discourses.

In addition to school personnel interviews, we conducted interviews with at least one parent of each student participant each year of the ethnographic study. Parent interview questions asked about parents' views of the school, the students at the school, and their child's experiences at the school; additional questions focused on parents' views of how well the school was preparing students for college and work. These data served as an important additional data source regarding school practices and discourses about the school.

Analysis

Because we were interested in both the ways historical narratives about women of color in STEM were taken up locally at students' schools as well as students' identity authoring and negotiation within these contexts, we conducted our analysis through a multi-stage process. First, we coded student, parent, and school-personnel interviews for views of the student with regard to academics, STEM, future plans, college, how positioned socially and academically, decisions about courses to take, relationships to family, and decisions about major. Compiling these coded data and students' transcript, we developed case summaries for each student that aimed to trace students' positioning, authoring, and interests within school and related to STEM over the course of their high school years. We then created data displays for each student of *critical moments*, *STEM/academic instantiations*, and *college/future instantiations*, organized temporally and based on all 5 years of student data (including student surveys). "Critical moments" were defined as moments that students name as being significant in shaping their interests in, pursuits of, and views of self within STEM. These displays allowed us to identify themes regarding (1) each student's authoring of academic and STEM identities over time; and (2) the schooling practices that students engaged in. After discussing and finding agreement regarding student authoring, we then looked for the schooling practices themes that emerged in the student data within the school and parent interviews to confirm or disconfirm our emerging claims. We discussed each claim, interrogated them for alternative explanations.

Findings

Although we set out to understand how narratives about young women of color were taken up locally in the school's practices and discourses, we found, instead, a dominant school narrative predominantly about race but not gender. The narrative at Capital focused on "decreasing the gap" through recruiting Black students to take more advanced courses and through motivating these students to challenge themselves more academically. Further, Capital's discourse and practices frame the low academic achievement of Black students as a result of students missing or, at times, intentionally skipping opportunities. Additionally, as Katie and Naomi authored themselves as good students, those capable in STEM, and those intending to pursue STEM careers, they had to do this in direct response to and in contrast with the local discourses and practices that positioned them otherwise.

Motivating students to "challenge" themselves

Capital High School serves a diverse population of students (23 percent Latino/a, 45 percent White, 25 percent Black) and was generally known as a "good" school within the surrounding community and the state. Students from Capital performed well, overall, on standardized exams, the school had a longstanding and successful Advanced Placement program, and many students who graduated from Capital continued on to competitive or highly selective colleges.

Although considered academically "successful" and highly ranked, Capital's narrative about students of color manifested in pervasive discourses about the "achievement gap" between students of color and White students. Capital's principal stated:

...if you look at how the kids are scoring on state standardized tests, our Latino and African American students are not testing nearly as high as our White students. And so...we want to fix the gaps here and help kids catch up and be on par with their other classmates...I think it's egregious that our kids...that there are those gaps. (130109_INT)

Naomi's mother (culturally Jewish with children who identify as African American) referenced the achievement "gulf" as a reason she and her husband elected to send their two older sons to a private school instead of Capital: "When we were looking at different high schools, what we kept hearing about [Capital] was that there was this gulf in achievement between the Caucasian students and the minority students" (110727_INT). Further, when Katie was asked why she thought she was often times one of the few Black students in her advanced courses, she responded, "Um, I think that there's definitely...we need to close the achievement gap, because it, like, it's not

because [Black students] are not smart enough to be in these classes... I feel like it, it, has a bit more to do with like, if someone is generally looking out for you" (121130_INT).

The gap was described by teachers, parents, and students as something "egregious," that should be "closed," and that primarily existed because of the low academic expectations students of color, and others, set for them. For example, students of color were described as not wanting to "challenge themselves" and "assuming" they "could not be successful" in advanced courses. As Katie stated in the above text, "most [Black] kids just assume... 'I probably can't [take harder courses].'" And, so, they don't. As we discuss later, Katie's comment points to the important role of social supports in addressing and resisting this historical narrative regarding Black students' academic ability. This depiction of Black students not challenging themselves is further described by Capital's principal:

[We] found that...a lot of kids, especially the minority students were saying, 'I don't want to challenge myself. I'm a B student in middle school,' you know, blah blah blah. And we said, 'you know what, in order to get these kids to try to achieve higher and try to get them in the [AP] class when they're juniors and seniors, we've got to get them to challenge themselves, go into the honors classes, find out that they can be successful and then move up.

Further, Naomi's mother stated, "I wouldn't say that the cultural norm is for there to be high achievement among [Black females at Capital]. It starts with something as simple [as] what classes they sign up for freshman year and whether [the students] are taking advantage of the honors and AP classes" (130109_INT).

However, student course-taking patterns were not simply viewed as an individual choice. Rather, shared ideas about what courses a student should take and the kinds of students who generally take them were viewed as also shaping these decisions. As Katie suggested, students of color hear from "everyone" that they "probably can't" be successful in advanced courses. And, to Katie, students of color hear this "you-can't" message "everywhere":

Allen: So "everyone" in that instance [is who]? Do you feel like it's, like a way that people talk at the school, or-?

Katie: I think definitely [it's the way people talk]; it's also peers, it's...maybe even [students'] parents. It's literally *everywhere*. Like in my speech and debate class, it's been...tons of upper middle class [kids] and...it's been predominantly White. And the teacher last year...she was like, "I'm sick of AP, because if you ask a lot of kids, 'Would you wanna [take AP speech]?', they're like, 'Oh, you gotta have a lot of money to do that.'" (121130_INT)

What's described here by Katie are shared ideas about the kinds of students - White, upper middle class - who take advanced or higher-level courses at the school, and the consequences this has for the Black and Latino students at the school.

One institutional practice that worked to maintain racial segregation at Capital while reinforcing the notion that White students were more academically capable and motivated (i.e., choosing to "challenge themselves") was what the principal called "blockers". Several years before our study, students at Capital had to receive a teacher recommendation in order to register for advanced-level courses: "I remember the first year [I worked at Capital]...the students only got into the [advanced-level] class with a teacher recommendation. Well what happens? Only White kids were getting into the honors classes and [AP] classes" (120120_INT).

At the time of our study, eradicating these "blockers" and encouraging teachers to "recruit" students of color into honors and advanced placed courses had become a school-wide emphasis, at least rhetorically:

So [because teacher recommendation resulted in only White students being in advanced courses], we [the school administration] said, 'No more blockers. If you [the student] feel you are strong enough to [take an advanced course] then [do that].' And we encourage teachers: 'If you see someone, especially a minority student, that you think has some potential, that you think could maybe be successful, to try to encourage them to take [Advanced Placement courses].'
Or encourage them to at least take an honors class. (120120_INT)

And although Capital staff had been working on this issue, Naomi and Katie were still navigating an advanced-course landscape that at least appeared glaringly White. During her sophomore year, Katie stated, "There's a lot

of ...well in some of my classes there are only like 5 Black kids, and I am in all honors classes except for Spanish. In my AP [Geography] class and my advanced Algebra II, I am the only Black kid" (110421_INT_Katie).

Discourses at Capital about the "achievement gap" favoring White students, the shared beliefs about the kinds of students who take advanced courses and eventually attend college, and the institutional practices that tracked students of color into lower-level courses seemed to be consistent with the ways most students of color thought about and navigated academic, and particularly advanced course-taking, decisions. As we will demonstrate further, these local discourses and practices, mediate their practice with a broader historical narrative about the personal shortcomings of students of color at Capital, played a critical role in the self-authoring Katie and Naomi engaged in.

Authoring identities within and against locally-embodied historical narratives

As Naomi and Katie, two young Black women, authored identities as "good students" at Capital, they did so in seeming contrast to the prevailing discourse about the underachievement of students of color and the pattern of racial segregation in advanced courses at the school. Katie and Naomi described themselves as "good students," emphasizing their hard work, enjoyment of learning, and their ability to succeed in their courses. They took honors courses (in various subjects) throughout their four years of high school, and AP courses their junior and senior years. By their senior year they had taken AP or honors Chemistry, AP Biology, or Honors Physics and AP Calculus.

However, in constructing themselves as good students, they drew from the school's discourse about students of color. Katie, for example, authored herself as one of the few Black students to take advanced courses, and as one who was able to overcome institutional barriers and discursive practices that her peers could not. She described other Black students as "believing that they can't do things and [believing] that they aren't smart." In contrast, Katie described herself as capable and not deterred by low grades or challenging courses: "...in some of my classes, like in math, sometimes I don't get the grades that I want, but because I like it so much and really want to understand it, that's what keeps me coming in everyday to my algebra teacher asking for help (110421_INT_Katie). Although Katie characterized other students of color as those who often were not motivated in school and did not challenge themselves, she authored herself as someone who enjoyed learning, wanted to learn more, and was willing to put in the extra effort (e.g. coming in everyday for help).

Naomi similarly drew on the prevailing discourse at Capital to distinguish herself: "I still think [doing math] is fun, and people are like, 'You have *fun* doing your math homework?' And I'm like, 'Yes I do.' People were like, 'The AP Calc[ulus] test sucks so much.' And I was like, 'I kind of enjoyed it.' I don't mind doing three hours of math" (130528_INT_Naomi). In both cases, Katie and Naomi took up the prevailing discourse about the low achievement and motivation of students of color at their school, and they did so as a means of distinguishing themselves from others who shared their racial characteristics. They accepted the denigration of students of color and used it to single themselves out as special.

As Katie and Naomi constructed their identities in contrast to the embodied narratives at their school, they simultaneously did work to support their identities as those good at math and science and capable of going into STEM fields. That is, as these young women authored selves who were good students they aligned themselves with culturally constructed characteristics (and those adopted by these students) of those individuals who navigated both school and STEM successfully. Students who took advanced courses, knew the material well, got good grades, worked hard, and pursued opportunities to learn more were ways that these two described both *themselves* as good students and *others* who were good at math and science.

For these young women, part of authoring a STEM-related identity was instantiating their ability to do well in math and science classes. For both students, math prowess was an integral part of their STEM identities. Consistently the girls authored selves who had natural math ability, being able to "get" math content, and math being a subject that "always makes sense" to them. Naomi, for example, stated that math "comes really easy" for her (121121_INT). Additionally, Katie described herself as "really good in math, the logic [of it]. [I'm good at] looking at a problem and seeing if it makes sense. She also described herself as "liking [math] so much" that she "really wanted to understand it" and so, "that's what keeps [her] coming in everyday to [her] algebra teacher asking for help" (110421_INT). Further, Katie's ability to do well in AP Calculus during her senior year prompted her exploration into (and eventually decision to pursue) a career in engineering. With the support of her father she was able to see the ways that solutions in AP Calculus related to engineering practices. She described working on a problem for class that involved a "really hard math concept" that was "exciting" to her: "...it was just so exciting for me [to work on this problem], and so I started thinking, 'Where can I do this all the time? If I want to do this kind of math, where can I do it?'" Math ability seemed to provide students with an academic leverage that bolstered their engineering identities.

Apparent in both Katie and Naomi's descriptions is an authoring of someone who takes pride in performing well academically and who enjoys learning in STEM, particularly math. These depictions stand in stark contrast to the characterizations of students of color at Capital as those who do not want to challenge themselves or those who take "easy classes" to graduate with higher grades.

Conclusions and implications

Our initial frame of reference for analyzing the two cases described in this article was to view them as examples of young women of color who successfully resisted the historical discourse that positions them as disinterested in STEM or unprepared to pursue STEM in college and beyond. We found that these young women were encouraged to pursue advanced math and science by many of their teachers, administrators, and friends; they took advantage of these advanced course opportunities; and they excelled in them. By the end of high school, both were proud of their accomplishments and still interested in math and science and the possibility of a future in a STEM field. Both had chosen a STEM major in college and were pursuing it at least through their second year of college.

However, from the perspective of Holland and Lave's concept of history-in-person (2001) and in light of our ethnographic data, these young people did not portray themselves as engaged in a struggle with a historical discourse about women in STEM. Rather, their struggle was with racial discourses at school that positioned them differently than they were or wished to be. A discourse about encouraging young women or students of color to pursue STEM was not present at Capital. The women's success in school, particularly in math, and their consequent positioning (qualifications, interests, identities) to go on in STEM were byproducts of their struggles and cultural productions regarding race.

Discourses of racial underachievement, nonetheless, were pervasive at the girls' school. As many other researchers have shown, these discourses affected the students academically and racially (e.g., Davidson, 1996). With a relatively diverse population, students of color at Capital were positioned as underachieving in comparison to Whites: lacking motivation, doubting their academic ability, and missing the social supports needed to navigate high school successfully. Teachers, administrators, parents and students engaged in discourses about the social and academic segregation, de facto tracking, and motivational shortcomings of students of color compared to Whites. The imperative for the school to "fix the achievement gap" was pronounced, but the onus was primarily on the students of color to step up to opportunities provided at school.

In the discursive context of racial underachievement, these young women fought to define themselves as different, i.e., to author themselves as "good students" in contrast to peers who shared their racial designation. They successfully challenged their expected positioning as historically racialized subjects, and in doing so, they gained valuable academic capital. As they enacted good student identities, they simultaneously came to identify as good in math or science or both and to be identified by others as "good students" in part because of their prowess in math and science. But they achieved this identity by accepting the validity of their schools' (and society's) negative representations of students of color as a group. They utilized features of the negative representation to distinguish themselves from their group—to make their racial peers the "other," and to author themselves as better than those others.

In authoring selves in contrast to others at their schools and authoring identities as those capable of going into STEM, these young women did create opportunities for their continued pursuit of STEM. However, their authoring of identities as good students in STEM did not engage with national narratives regarding women of color in STEM or the very consequential reality of the marginalization, isolation, and, at times, outright abuse that women of color often face as they pursue STEM in higher education and careers. We worry about the limited resources these women will have to support them as they try and navigate these spaces in the future. High STEM drop-out or push-out rates among women of color may be one consequence.

Further, with regard to our original aim in this analysis, these findings suggest a need to address the intersections of race, gender, and STEM more explicitly in schools. We wonder what it might have meant for young women in our study to be equipped with an awareness of the "double bind" women of color face (discrimination as a "double" minority) in STEM (Malcolm, Hall, & Brown, 1976; Williams, Phillips, & Hall, 2014); and what it might look like to equip students with strategies that enable them to challenge not just their positioning within local discourses, but to challenge the discourses themselves.

References

- Eisenhart, M., Weis, L., Allen, C. D., Cipollone, K., Stich, A., & Dominguez, R. (2015). High school opportunities for STEM: Comparing inclusive STEM-focused and comprehensive high schools in two US cities. *Journal of Research in Science Teaching*, 52(6), 763–789. <http://doi.org/10.1002/tea.21213>

- Bang, M., Warren, B., Rosebery, A. S., & Medin, D. (2013). Desettling Expectations in Science Education. *Human Development*, 55(5-6), 302–318. <http://doi.org/10.1159/000345322>
- Calabrese Barton, A., Kang, H., Tan, E., O'Neill, T. B., Bautista-Guerra, J., & Brecklin, C. (2012). Crafting a Future in Science: Tracing Middle School Girls' Identity Work Over Time and Space. *American Educational Research Journal*, 50(1), 37–75. <http://doi.org/10.3102/0002831212458142>
- Carlone, H. B., & Johnson, A. (2007). Understanding the Science Experiences of Successful Women of Color : Science Identity as an Analytic Lens. *Journal of Research in Science Teaching*, 44(8), 1187–1218. <http://doi.org/10.1002/tea>
- Carlone, H. B., Scott, C. M., & Lowder, C. (2014). Becoming (less) scientific: A longitudinal study of students' identity work from elementary to middle school science. *Journal of Research in Science Teaching*, 51(7), 836–869. <http://doi.org/10.1002/tea.21150>
- Davidson, A. L. (1996). *Making and molding identity in schools: Student narratives on race, gender and academic engagement*. Albany, NY: SUNY Press.
- Holland, D. & Lave, J. (2001). *History in person: Enduring struggles, contentious practice, intimate Identities*. Santa Fe, NM: School of American Research.
- Holland, D., Lachicotte, W. J., Skinner, D., & Cain, C. (1998). *Figured Worlds. Identity and Agency in Cultural Worlds*. Harvard University Press.
- Malcolm, S.M., Hall, P.Q., & Brown, J.W. (1976). *The double bind: The price of being a minority woman in science*. Washington, DC: American Association for the Advancement of Science.
- Ong, M., Wright, C., Espinosa, L. L., & Orfield, G. (2010). *Inside the Double Bind: A Synthesis of Empirical Research on Women of Color in Science, Technology, Engineering, and Mathematics*. National Science Foundation.
- Williams, J. C., Phillips, K. W., & Hall, E. V. (2014). *Double Jeopardy?: Gender Bias Against Women of Color in Science. Tools for Change*.
- Yin, R. K. (2013). *Case study research: Design and methods*. Sage publications.

Acknowledgments

This paper is based on research supported by the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Blending Play and Inquiry in Augmented Reality: A Comparison of Playing a Video Game to Playing Within a Participatory Model

David DeLiema, Christine Lee, Noel Enyedy, Randy Illum, and Maggie Dahn
deliema@g.ucla.edu, christine.lee.ucla@gmail.com, enyedy@gseis.ucla.edu, randy@remap.ucla.edu,
maggiedahn@gmail.com
University of California, Los Angeles

Asmalina Saleh, Joshua Danish, Megan Humburg, and Charles Mahoney
asmsaleh@indiana.edu, jdanish@indiana.edu, mahumbur@umail.iu.edu, charmaho@umail.iu.edu
Indiana University

Abstract: Researchers have increasingly demonstrated how technologies such as augmented reality (AR) can leverage embodiment within play to help students use physical movement to explore complex concepts. Using Vygotsky's (1978) notion of play, we examine how two distinct AR environments—rule-based game play and open-ended modeling play—support 1st and 2nd graders' inquiry (N=122) into how matter changes state at the level of microscopic particles. We further use the notion of keys (Goffman, 1974) to examine how the students construct distinct participation frameworks (Goodwin, 1993) within the two activity designs, and how this organization of activity may impact their learning experience. Our analyses show that students within a game-play environment were more oriented towards accomplishing a goal rather than understanding how a system works whereas those in the modeling-play group focused more explicitly on understanding mechanism and process.

Keywords: embodiment, augmented reality, keys, conceptual blending, interaction analysis

Introduction

The exploration of games to support learning has garnered much attention in recent years (Clark, Nelson, Sengupta, Angelo, 2009). Part of the attraction of games for learning is the playfulness, engagement, and risk-taking they engender (Gee, 2007). In particular, researchers have increasingly demonstrated how new technologies such as augmented reality (AR) can leverage embodiment within games to help students use their physical experiences to explore complex concepts (Enyedy, Danish, Delacruz, & Kumar, 2012; Lee, 2015). However, it is not yet clear how the intersection of embodiment and games effectively supports learning (Lindgren, 2015).

Our study aims to examine how two different game-like AR designs that leverage embodiment can support student learning of states of matter and its particulate nature. We examine evidence from an ongoing design-based research project: Science through Technology Enhanced Play (STEP) to highlight distinctions in two related play environments. In both designs, our overall goal is to shape students' experiences of learning about particles through role-play by varying the nature of the rules underlying the activities, the short-term aims of the activity, and the role of narrative in supporting learning. In STEP (Danish, Enyedy, Saleh, Lee, & Andrade, 2015), first- and second-grade students "shrink" down to the size of microscopic particles to investigate liquid, solid, and gas phase changes through collective body movements. First and second graders (6-8 years old, N=122) are assigned to either one of two AR environments: 1) a design that is oriented towards students producing their own models, implicit rules, and a more fluid narrative (**modeling-play**), or 2) one focused on winning as an end state, overt rules, and a fixed narrative (**game-play**). We believe that both formats are equally effective in supporting learning; indeed, pre-post results indicate that students in both groups perform equally well on assessments of content knowledge. However, our goal is to develop empirically grounded theory that teases apart *how* features of game-play and modeling-play affect the character of student inquiry in ways that impact learning gains and other forms of success that do not appear on typical assessments.

Theoretical framework: Game-play and modeling-play

Vygotsky (1978) observes that there is an imaginary situation and a set of rules in both socio-dramatic play and game play. However, he argues that rule-based games tend to have thin imaginative contexts (e.g. in chess, the pieces do not graphically attack one another), while socio-dramatic play tends to have rich imaginary narratives but more flexible and negotiable rules (e.g. playing at superheroes entails freedom to decide who has specific superpowers). In both cases, Vygotsky notes that the playful context affords children an opportunity to look upon

the rules which govern it in a new light, thus exploring those rules and learning about them in ways that are not as intuitive in other contexts. Similarly, Caillois (1955) describes early childhood socio-dramatic play and the playing of games in later childhood as occupying two points along one spectrum. *Paidia* refers to early pretend play without structured rules and *ludus* refers to games with enforced rules.

In our design, we leverage the distinctions highlighted by Vygotsky (1978) and Caillois (1955) to better understand how these instantiate in learning environments. Specifically, we contrast how a win-oriented, more rule-based, and fixed narrative (game-play) engenders qualitatively different approaches to scientific inquiry for students than a more open-ended design in terms of rules, goals, and narrative (modeling-play). We refer to the second design as modeling play because students are engaged in predicting scientific phenomenon based on our computational representation of particle behavior (Schwarz et al., 2009). Our hypothesis is that different levels of structure within these two designs mirror the distinctions between science and engineering mindsets that Schauble and colleagues (1991) noted. Specifically, engineering mindsets (e.g., game-play) are oriented towards accomplishing a goal rather than understanding how a system works. In contrast, scientific mindsets (e.g., modeling-play) focus effort more explicitly on understanding mechanism and process.

Liminal blends

The STEP simulation software, which is the heart of both activities, is designed to support students in constructing liminal blends (Enyedy, Danish, & DeLiema, 2015). Liminal blends incorporate conceptual blending theory (Fauconnier & Turner, 1998) into a larger distributed cognition framework to highlight how individual mental processes intersect with the material and social world. Briefly, conceptual blends occur when learners integrate disparate cognitive resources to reach new conclusions. Liminal blends are a type of conceptual blend where a student fuses her first-person experience with that of an imagined or an inanimate object (e.g. imagining/gesturing being a point on a graph) and moves back and forth between perspectives without making a clear distinction between them. Integrating resources in this way obscures the division between the physical and conceptual, and has been referred to as semiotic fusion (Nemirovsky, Tierney, & Wright, 1998) or liminal worlds (Ochs, Jacoby, & Gonzales, 1996). In such cases, a “blurring” takes place as participants appear to move fluidly between the physical world in one moment and the symbolic world in the next. The emergent properties of the blend support or hinder learning. In the STEP environment, we therefore intentionally develop new virtual resources (e.g., graphical representations) to support this kind of blending.

Research design and methods

The STEP environment (see Figure 1) was designed to support students as they explored and reflected on science content through embodied play. Microsoft Kinect cameras were placed around the classroom to capture student movement, and the STEP software used students’ movement to control aspects of a computer simulation of water particles assembling in different states of matter. As 6-12 students moved about the space, each one was assigned a representation in the shape of a particle, and these particles interacted with one another to create solid, liquid, and gas. As the students moved around, they saw the lines connecting each particle to its nearest neighbor change color, with each line color representing a different type of bond (white for solid, blue for liquid, and red for gas). Three state meters on the side of the screen also showed students what percentage of the bonds within the current simulation were currently representing solid, liquid, and gas at any given time.

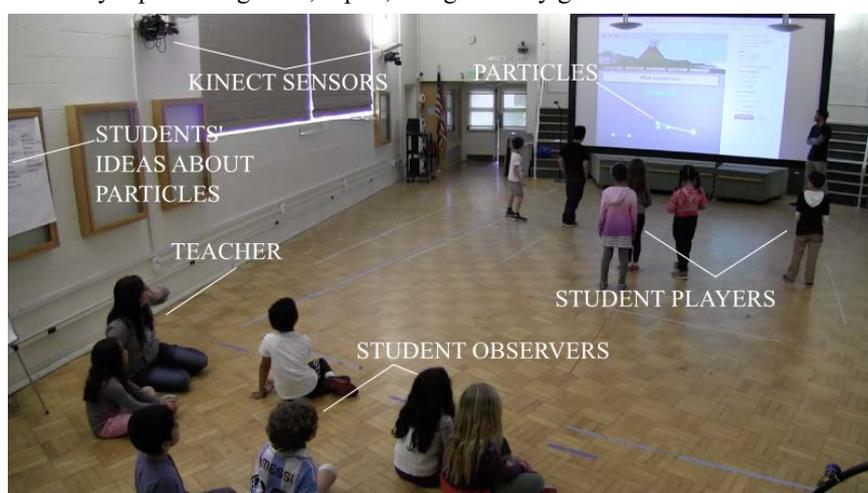


Figure 1: The technological and activity set-up of the STEP system.

STEP activities

Students collaboratively used their physical movements to explore particle behavior, using the feedback displayed on a projected screen to adapt their actions in line with the goals of each activity. Those students who were not participating in the simulation were cast as observers whose job was to reflect on connections between their classmates' actions and particle behavior occurring in the simulation.

We extended a previous application of the STEP technology (Danish, Enyedy, Saleh, Lee, & Andrade, 2015) to include two modes of play. In modeling-play, the onscreen simulation depicted particles moving in an empty container. No specific goals, explicit narratives, or rules were built into the software, though the teacher and students developed goals such as constructing a specific state of matter (see Figure 2, at left). In the game-play version of STEP, students were tasked with helping a robot escape a volcanic island to “win” the game. The core narrative was that particles were inside the robot’s “engine” and students could control the robot’s behavior by maintaining specific rules, as instantiated by states of water (e.g. creating gas at the particle level moves the robot, solid ice protects it, and water puts out fire). The simulation screen was split, with the top half showing the actions of the robot and the bottom half displaying the movement of the students’ particles (see Figure 2, at right). Across the two modes, students engaged in similar activities. First, the students began by exploring the effects of hot and cold environments on the macro level properties of matter. Next, students transitioned to a micro level view of matter in which each student controlled one particle and the class reflected on the particle behavior. Finally, we switched focus to the impact of energy on particle behavior, with each student controlling an energy wand that heated up or cooled down any particle it touched.



Figure 2: Screen shots of what the students see on the screen in the play and game conditions.

Methods

Participants were from five mixed-age classrooms with 1st graders ($n=58$) and 2nd graders ($n=66$). There were a total of 122 participants (6-8 years old), 48 in the game-play condition (52% girls) and 72 in the modeling-play condition (54% girls). Four teachers participated and each had more than six years of teaching experience.

We first examined the extent to which the different groups learned after participating in the two game modes by conducting a mixed ANOVA. Content understanding was assessed in pre-post tests using content interviews. Content understanding was operationalized as descriptions of particle behavior in the different states of matter (matter-type codes) and the mechanisms behind state changes (change-type codes). These codes were derived from our earlier study (Danish et al., 2015) and Paik, Kim, Cho, and Park (2004). Students were asked about how particles behaved in different states as well as the mechanisms behind state changes. A total of three coders analyzed and categorized the pre-post video data. Interrater agreement between pairs of coders with Pearson’s correlation ranged from .798 to .849, whereas the intraclass correlation for individual raters was .851.

Participation frameworks and keys

To describe how these mindsets or forms of inquiry differ between game-play and modeling-play, we build on the notion of *participation frameworks*—the tacit and explicit organization of participants’ interactional rights and responsibilities (Goodwin, 1993)—within AR learning environments. Participation frameworks are known to vary across contexts (Erickson & Schultz, 1981), but in this case we examine how distinct participation frameworks are blended together in such a way that social dynamics shape which conceptual resources are brought to bear

during learning. For example, in comparing play contexts, we would ask: Who has the right to generate a new goal or a new narrative in play and what are the participants responsible for accomplishing? Within our two designed play contexts, these questions result in different forms of inquiry and use of different conceptual resources. One way of understanding the impact of participation frameworks, particularly when the goal is to understand how they are built up from multiple, potentially contradictory contextual factors, is to focus on *keys*: “the set of conventions by which a given activity, one already meaningful in terms of some primary framework, is transformed by the participants to be something quite else” (Goffman, 1974 p. 43-44). In social interaction, people can perform more than one transformation to an activity by embedding keys within keys, in effect creating harmonies from the layering of multiple participation frameworks. In the following section, we elaborate on the construct of layered keys in the context of STEP.

Keys in the STEP activity

We focus on those instances where two or more keys are blended together to create a scenario greater than the sum of their parts in the same way that harmonies in music are perceived as more than independent notes. In STEP, students’ learning experiences take place within a layering of keys. The primary framework on top of which the technology and curriculum unfold is science class. This primary event becomes keyed in three ways: as a simulation, as play, and as inquiry (see Figure 3). The technology in the classroom invites students to key particle movement into a human-scale, slowed-down enactment of the real thing, what learning scientists have labeled participatory simulations (Colella, 2000). The teachers and researchers explicitly invite students to key the participatory simulation toward a playful stance, taking what could otherwise be a serious orientation to the simulation and giving students leeway to have fun and pursue what piques their interest. The experience of being inside the particle simulation and taking a playful stance both become further keyed in an inquiry context, where students are asked to observe, experiment, and record patterns in particle movement.

In summary, science class becomes keyed to a human-scale enactment of particles, which becomes keyed in a playful tone, which becomes keyed in an inquiry mindset. This layering of keys makes it so that the status of students’ first-person experiences inside the particle simulation cannot be separated from the playful key or from the inquiry key. In game-play, the middle-layer play key is substituted for a goal-directed narrative that contains criteria for success and failure (e.g. a scoring system). In our analysis, we explore how the students and teachers establish each layered key and blend them together. We point out variations in goal setting, idea generation, and experimentation throughout inquiry that result from these different configurations of keys.

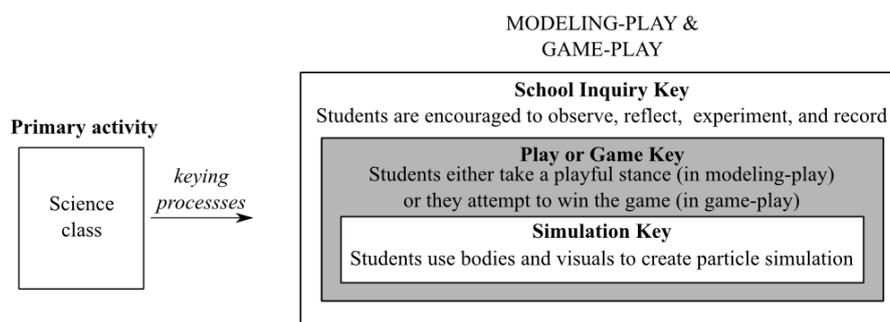


Figure 3: Layers of keys in the STEP activity.

We examine how the different game modes supported learning by combining narrative description with close interaction analysis (Jordan & Henderson, 1995). Consistent with the early stage of our research, we selected two episodes for our analysis, one from modeling-play and one from game-play, to provide a grounded description of students’ inquiry in each condition and to advance our theoretical understanding of blended keys. These episodes consist of different inquiry patterns that allow us to explore the variation in the two play modes and provide exemplars of layered keys, which we can refine empirically and theoretically in subsequent work.

Results

A mixed ANOVA was conducted to assess how the two play modes influenced students’ pre-post test scores. There was no significant interaction between play mode and time, Wilks Lambda = .996, $F(1, 118) = 0.305$, $p = .513$, partial eta squared = .004. There was a substantial main effect for time, Wilks Lambda = .305, $F(1, 118) = 269.286$, $p < .0001$, partial eta squared = .695, with both play modes showing an improvement of scores between the pre and post-interviews. As expected, the main effect comparing the two play modes was not significant, $F(1,$

118) = .0557, $p = .457$, partial eta squared = .005, suggesting no difference in how the modeling-play and game-play modes impacted students' scores. Given that the learning outcomes were equivalent in both groups, this fueled our interest to examine the different stances toward inquiry in each mode.

Inquiry processes in the modeling-play condition

Inquiry in the modeling-play condition takes place within the context of three layered keys: simulation, play, and inquiry (see Figure 3 above). The teacher, Ms. Jones, unfolds the first key (simulation) when she explains to the students, "Today, you guys are going to be particles." Moments later, Ms. Jones keys the simulation in a playful direction: "We are going to shrink you down with a special magic shrink machine," in reference to a hula-hoop decorated with spray-painted styrofoam balls (depicting particles). Ms. Jones then introduces the third key to students seated outside the simulation space: "The rest of you, your job is to notice what happens. If you have observations, say them out." These keys establish a space in which the class is not solely creating a particle *simulation*; they are creating a *playful* particle simulation and one engaged for the purpose of *scientific inquiry*.

Soon, all 12 of the students have moved through the hula-hoop portal and are playing as particles. Three students initiate a sidebar conversation with the teacher (not audible in our recording). Ms. Jones says, "Okay, I've heard a couple people say that." Ms. Jones looks up at all of the students, and in a loud voice, counts down: "5, 4, 3, 2, 1, freeze." The countdown marks a process of backgrounding the simulation and play keys, and foregrounding the inquiry key. The students then begin proposing ideas for testing whether location in the space determines particle color (state) and they decide to have boys stand on one side and girls on another. With the class split in half, the girls request that the boys stop running around because their movement is interfering with the test, the boys stop moving, the students observe the screen, the teacher asks if the idea is correct, and students' respond in chorus: "No!" The class has ruled out one hypothesis.

A student, Marie, proposes a new idea: Hugging can turn two people into one particle. The students form hugs in groups of two or three and shuffle around laughing. At this point, the students and teacher have blended three keys: The students are testing an idea from a peer's observation (inquiry key), and they are doing so with movement in the tech space (simulation key) and with a light-hearted stance (play key). Amidst the laughter, a new inquiry key, again in the form of a side bar, forms between the teacher and a student named Bethany. The teacher quiets the class down and Bethany explains: "Um, if we can make a caterpillar, and um, we can see if it's one particle or many particles." The students assemble into a caterpillar formation, standing in one straight line, holding on tightly to one another. They observe that the caterpillar formation does not amount to a single particle. And yet as they stand together in their caterpillar formation, one student at the back of the line, Carl, breaks free (4.1) and dances with a smile on his face to the opposite side of the room (see Figure 4).

Ms. Jones: Wh-wh-whoa, what happens when — what's Carl doing?

Students: [(laughter)]

Carl: [(walks back to caterpillar line)]

Ms. Jones: No. Go back! Go back!

Carl: (again moves away from the caterpillar cluster)

Brandon: Look, it's red! It's red! When you're close to each other it's yellow and when you're far apart it's red.



Figure 4: Carl breaks free from the "caterpillar" and Ms. Jones uses the deviation to key inquiry.

This moment marks an opportunistic harmony between three active keys. Carl, keying the activity with a playful stance, breaks free from the caterpillar formation and dances his way across the room. When Ms. Jones publicly calls out Carl's action, Carl seems to assume that he has violated a norm, and begins to walk back to the caterpillar. However, seizing the value of the play-based deviation, Ms. Jones re-shapes the moment into inquiry,

telling Carl to return. Another student, wedged between his peers at the center of the caterpillar, also pursues the inquiry key, pointing to the screen and providing a crisp statement about how distance shapes color (4.2, 4.3, and 4.4). With respect to conceptual understanding, the keying of play, inquiry, and simulation amount to an incipient recognition of how distance between particles determines state of matter.

Discussion: Emergent goals and experimental ideas in modeling-play

We make two points about the inquiry process in the above episode of modeling play that we argue can be understood as a product of the layering of keys in the activity: Students generate emergent goals for their inquiry and they experimentally and opportunistically test new ideas. On the first point, the students select goals for inquiry that come from observations made in the play key. The students explore three ideas—side of the room determines color, hugging produces a single particle, and caterpillar produces a single particle. When a student proposes a new idea to the whole class, the students background the play key and foreground the inquiry and simulation keys. At any point, the students can attempt to re-invoke the play key (as we see with Carl in the caterpillar episode) and trigger another cycle of goal setting and idea generation.

However playfully students select goals for inquiry, they also formulate intelligible ideas and experimentally test them. The students think together about productive ways to test the idea and they execute the test with careful observations. Even when the students disprove their ideas, they make opportunistic progress uncovering salient dynamics in the simulation. For example, even though the side of the room is irrelevant to particle states of matter, the students testing this idea recognize that movement on the boys' side has added a confounding variable to their experiment. Similarly, even though clumping together into a caterpillar does not produce a single particle, one student's playful deviation from the caterpillar line leads to another student noticing that the deviation changed the particle color, and thus that distance might be in play. These two dimensions, speed and distance, are the two parameters that determine the state of matter in the STEP particle simulation. Despite that the students have not yet formulated the rules for speed and distance, they uncover both as relevant properties of state change.

Inquiry processes in the game-play condition

In the game-play condition, inquiry takes place within the context of three layers of keys: the particle simulation, the game narrative, and the inquiry mode. Ms. Lopez initiates the simulation key: "Are you ready to become particles?" Before she selects six students to enter the simulation space, Ms. Lopez keys the inquiry layer: "I'm going to ask those friends who are staying here, you have a very important job as scientists. You have to observe what it is your friends are doing, and together we have to decide do we need to change these rules to help them?" The teacher then invites six students into the simulation space after which both of the researchers trigger the game key, explaining that the game is hard and requires teamwork.

About half way into the game, the students have created gas and liquid to move the robot forward and put out fire. Half of the students are playing as particles (players), and half of the students and the teacher are watching from outside the simulation space (observers). The players start working to create liquid when the teacher notices that the players instead need to create an ice force field. She tells the players to look at the screen and the players catch on. In this exchange (see Figure 5), we see the players abruptly shift their goal to wanting to create an ice shield for the robot. The shift marks a foregrounding of the game key, where the players cater to what the protagonist in the game narrative needs to survive. Three different students then suggest three different strategies in rapid suggestion: "don't touch," "spread out," and "move a little to form a circle." The players never test the second two suggestions. After Researcher 1 pauses the game, both researchers comment on the players' actions relative to the game (not "forming anything" and "not working"), foregrounding the robot.

With the game paused, two observers and the teacher re-state strategies earlier suggested: "spread out" and "form a circle while moving slowly," the latter of which an observer and the teacher read from the paper that has the students' particle rules from the day before. The players ignore both suggestions. Instead, Horace says, "Wait, everybody get on an X" (referring to pieces of blue tape on the floor from an unrelated activity). Sandra offers: "Guys, get in the corner." A player asks Researcher 1 to unpause the game, which he does. Sandra again says the players should move to the corner and Horace tells a player to move to a specific X. Ms. Lopez adds on: "Where Sandra was standing, something happened." All the while, some players are standing motionless on Xs, other players are moving around, and the clock is counting down. The Researcher pauses the game again. Ms. Lopez again attempts to draw attention to the "something" that happened in Sandra's prior location, but a player asks for the game to be timed in again. This request marks the tension between the inquiry and game keys: The "something" Ms. Lopez notices becomes backgrounded when a player re-initiates the game key. Horace yells for everyone to spread to a corner, no ice forms, and the students lose the game.

Henry (observer): FORCE FIELD!
 Horace (player): Force field. Get. Okay. Don't touch.
(players stop holding hands and freeze)
 Claudia (observer): SPREAD OUT!
 Ms. Lopez: Do you want to [pause it?]
 Horace (player): [We have] to go
 Jeffrey (observer): (turning to read from the white board)
 Move a little from a circle while
[moving slowly]

Researcher 1: [I'll pause it so]
 you guys can think a little bit.

Researcher 2: Were you forming anything [right there?]

Researcher 1: [This was not] working here, huh?

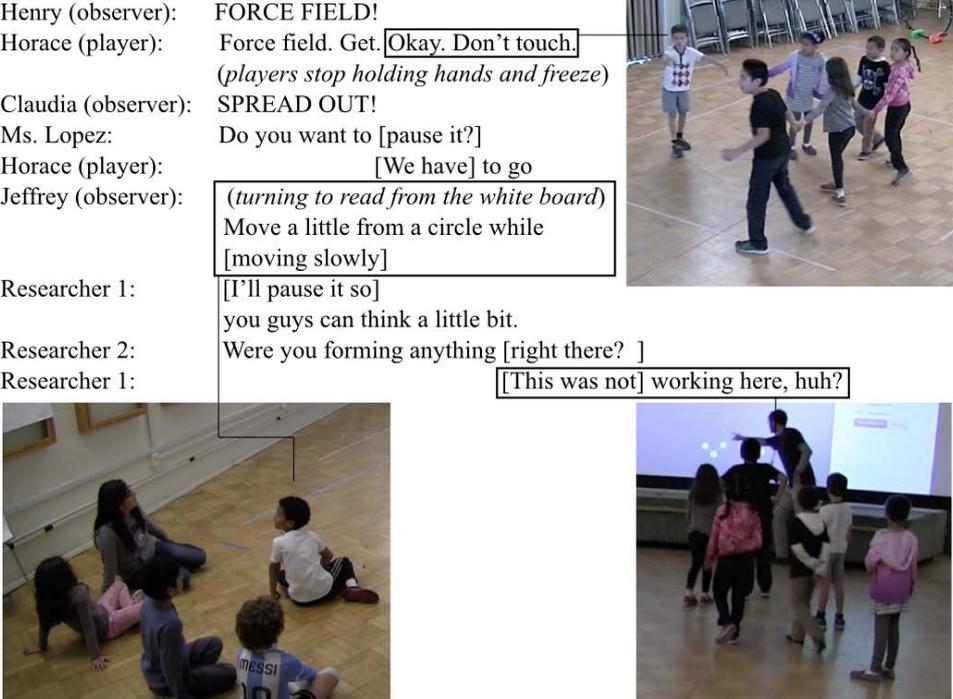


Figure 5: The students shift their goal to serve the needs of the robot and suggest a variety of strategies.

Discussion: Fixed goals and deploying strategies

In contrast to the emergent goals that formed in the modeling-play activity, the game-play activity represents an orientation to the fixed goals available in the robot key. Henry announces this goal at the start of the episode: “ICE SHIELD!” Even though students can in principle move any way they please in the simulation, the robot narrative becomes the deterministic anchor for their goals. The entire community shifts its attention to the goal of creating ice, and that goal persists up until the point that the clock runs out and the robot dies. This marks the foregrounding of the robot game key over the inquiry key.

In contrast to the experimental and opportunistic testing and observation of ideas in the modeling play activity, students in the game activity produce a nearly endless stream of action-based strategies with a focus on whether those strategies create an ice shield. In total, four different students produce five strategies: “don’t touch,” “spread out,” “move a little from a circle,” “stand on an X pattern,” and “get in the corner.” Three of those ideas are never tested in full. The inquiry key, described by the teacher at the head of the lesson as entailing observing the simulation and re-examining the particle rules, becomes adapted to accommodate the pace and structure of the game. When an opportunity arises to investigate the “something” interesting that Ms. Lopez notices, the game key is immediately re-activated and the players attempt yet another strategy. With respect to conceptual understanding, the students’ throw-it-at-the-wall-and-see-what-sticks approach eventually (in subsequent attempts at the game) leads to a successful strategy for helping the robot.

Conclusion

Despite that our current assessment shows that students learned the science content equally well in the modeling-play and game-play activities, our qualitative analysis suggests that different features of the two activities—the open versus structured narrative and the lack of assessment versus the built-in assessment of success/failure—drove students to set goals and explore ideas in substantively different ways. These distinct inquiry patterns occur at the intersection between the structure of the play/game experience and the learning community’s own prioritization and blending of different keys (simulation, game, play, inquiry). Given the small number of episodes analyzed qualitatively, and the idiosyncratic design of the game (e.g. the game is collaborative, has a timer, and usually involves students taking turns), we cannot make broad conclusions about the value of games versus play for science learning. However, our theoretical framework around blended keys points to a methodology for tracking how features of game design and modeling play may afford more or less productivity at different stages of inquiry. From our analysis, we can consider how teachers and designers can foreground, background, and blend simulation, play/game, and inquiry keys at different stages of ongoing activity to promote idea generation,

experimentation, rehearsal, etc. We can also consider how different configurations of simulation, play, and games make it harder or easier to sustain inquiry and argumentation. How keys are established, backgrounded, and foregrounded in negotiation between students and teachers, and the value these different configurations provide for inquiry, are open questions. As learning scientists grapple with the nuanced spectrum between play and games, we see blended keys as a viable construct for teasing apart the influences on inquiry that may have implications for design and instructor education.

References

- Caillois, R. (1955). The structure and classification of games. *Diogenes*, 3, 62-75.
- Clark, D., Nelson, B., Sengupta, P., & D'Angelo, C. (2009). *Rethinking science learning through digital games and simulations: Genres, examples, and evidence*. Washington DC: National Research Council.
- Colella, V. (2000). Participatory simulations: Building collaborative understanding through immersive dynamic modeling. *The Journal of the Learning Sciences*, 9(4), 471-500.
- Danish, J. A., Enyedy, N., Saleh, A., Lee, C., Andrade, A. (2015). Science through Technology Enhanced Play: Designing to support reflection through play and embodiment. In Lindwall, O., Häkkinen, P., Koschman, T. Tchounikine, P. & Ludvigsen, S. (Eds.) "*Exploring the material conditions of learning: opportunities and challenges for CSCL*," the Proceedings of the Computer Supported Collaborative Learning (CSCL) Conference, (Vol. 1, pp. 332 – 339). Gothenburg, Sweden: The International Society of the Learning Sciences.
- Enyedy, N., Danish, J. A., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *International Journal of Computer-Supported Collaborative Learning*, 7(3), 347-378.
- Enyedy, N., Danish, J. A., DeLiema, D. (2015). Constructing liminal blends in a collaborative augmented-reality learning environment. *Int'l. Journal of Computer Supported Collaborative Learning*, 10(1), 7-34.
- Erickson, F., & Schultz, J. (1981). When is a context? Some issues and methods in the analysis of social competence. In J. Green & C. Wallat (Eds.), *Ethnography and language in educational settings* (pp. 147–160). Norwood, NJ: Ablex.
- Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science*, 22, 133–187.
- Gee, J. P. (2007). *Good videogames and good learning: Collected essays on video games*. New York, NY: Peter Lang Publishing.
- Goodwin, M. H. (1993). Tactical uses of stories: Participation frameworks within girls' and boys' disputes. In D. Tannen (Eds.) *Gender and conversational interaction* (pp. 110-142). New York: Oxford University Press.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge, MA: Harvard University Press.
- Jordan, B. & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39-103.
- Lee, V. (Ed.) (2015). *Learning technologies and the body: Integration and implementation in formal and informal learning environments*. New York: Routledge.
- Lindgren, R. (2015). Getting into the cue: Embracing technology-facilitated body movements as a starting point for learning. In V. Lee (Ed.), *Learning Technologies and the Body: Integration and Implementation in Formal and Informal Learning Environments* (Vol. 135). New York: Routledge.
- Nemirovsky, R., Tierney, C., & Wright, T. (1998). Body motion and graphing. *Cognition and Instruction*, 16(2), 119–172.
- Ochs, E., Jacoby, S., & Gonzales, P. (1994). Interpretive journeys: How physicists talk and travel through graphic space. *Configurations*, 2, 151–171.
- Paik S. H., Kim, H. N., Cho, B. K., Park, J. W. (2004). K-8th grade Korean students' conceptions of 'changes of state' and 'conditions for changes of state.' *Int J Sci Educ*, 26(2), 207–224.
- Schauble, L., Klopfer, L. E., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching*, 28(9), 859-882.
- Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y, Hug, B. & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632-654.
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Qualitative and Quantitative Information in Cognitive Group Awareness Tools: Impact on Collaborative Learning

Melanie Erkens, Patrick Schlottbom, Daniel Bodemer
melanie.erkens@uni-due.de, patrick.schlottbom@stud.uni-due.de, daniel.bodemer@uni-due.de
University of Duisburg-Essen

Abstract: A large body of research covers the positive impact of cognitive group awareness tools on collaborative learning processes. These tools provide learners with visualizations of relevant cognitive information on their learning partners, triggering the resolution of cognitive conflicts, improving accurate partner modeling and, associated therewith, implicitly guiding learners' activities instead of prescribing certain ways of behaving. Said visualizations can be differentiated by the type of awareness information they provide: qualitative or quantitative. We systematically compared the impact of both types of information on collaborative learning in an experimental study ($N = 51$). The results suggest that the availability of both qualitative and quantitative information can evoke cognitive conflicts that guide learners' communication behavior in a goal-oriented way. Furthermore, visualizations combining both types of information appear to facilitate accurate partner modeling.

Introduction

A large body of research covers the positive impact of cognitive group awareness tools on collaborative learning processes (cf. Janssen & Bodemer, 2013). Such tools support learners by providing them with relevant cognitive information on their teammates' knowledge (Janssen & Bodemer, 2013) and implicitly guiding them through the learning process, the latter meaning that the tools suggest certain ways of thinking, communicating and behaving instead of prescribing them (Bodemer, 2011). This enables the learners to (self-)regulate their learning with reference to learning mechanisms being central to collaborative learning, e.g., conflict and partner modeling (Dillenbourg, 1999). Although research on existing tools suggests a strong effect of such tools' visualizations on learning behavior and outcomes, the various types of visualized information, namely qualitative or quantitative information, have not been investigated separately so far. Thus, we aim to systematically investigate the impact of both types of visualizations on collaborative learning to better understand their degree of efficiency and get clues on how to design even better cognitive group awareness tools in the future.

Cognitive group awareness tools are characterized by its main function of transforming cognitive information, e.g., by categorizing or aggregating ratings on knowledge, and visualizing it (cf. Bodemer & Buder, 2006). In so doing, they differ in the types of cognitive information they provide: Some tools visualize *what* a learner knows about various concepts in comparison to another learner (e.g., Schnaubert & Bodemer, 2015; Dehler, Bodemer, Buder, & Hesse, 2011; Engelmann & Hesse, 2010), and other tools visualize *how much* the learning partner knows or rather combine both information in one knowledge representation (e.g., Erkens & Bodemer, 2015; Sangin, Molinari, Nüssli, & Dillenbourg, 2011). By focusing on the functions of the aforementioned cognitive group awareness tools, we can differentiate between qualitative and quantitative information. Qualitative cognitive information provides binary categories of knowledge distributed across concepts, i.e., whether or not knowledge on a specific concept is present. In contrast, quantitative information describes quantitative values that represent the extent of knowledge. Both can also be combined so that the learners can be aware of how much they know about each concept belonging to a topic. Such information can be visualized for one or several learning partner(s), usually including the learners themselves to simplify the comparison between teammates, allowing goal-oriented collaboration (Bromme, Hesse, & Spada, 2005).

Comparability between learning partners is an important factor in this context, since it is associated with relevant learning mechanisms such as partner modeling and conflict. The concept of conflict goes back to Piaget (1977) and is based on his idea that interactions with physical or social environments can lead to a disequilibrium. In the case of social environments contradicting the knowledge of a learner, intrapersonal (Piaget & Inhelder, 1969) or interpersonal cognitive conflicts (Doise & Mugny, 1984) might occur. Cognitive development is making progress when learners find cognitive balance through assimilation or accommodation (Piaget, 1977). Both types of cognitive conflict are regarded as important antecedents to individual learning (Buder, Schwind, Rudat, & Bodemer, 2015). Furthermore, it is assumed that learners maintain a model of their learning partners from which they can infer and internalize their knowledge (Dillenbourg, 1999). From a Vygotskian perspective, internalization of socially regulated and mediated knowledge is the motor of cognitive development taking place when a layperson interacts with a more capable partner (Tudge & Rogoff, 1989).

Cognitive group awareness tools aim to facilitate the comparison between learning partners by visualizing or highlighting conflicts between them and making conflictual situations more salient (Bodemer, 2011). From a Piagetian perspective, comparative visualizations draw one's attention to knowledge gaps triggering intrapersonal conflicts or to contradicting knowledge that appears to trigger interpersonal conflicts—both of which need to be resolved. Intrapersonal conflicts can support collaborative learning, since knowing about shared and unshared knowledge resources triggers discussions about topics, with which only one learner in a group is familiar (Schittekatte & Van Hiel, 1996). Dehler and colleagues (2011) further found that the visualization of binary qualitative information on topical knowledge (present or not) guides the communication between learning partners. Learners who were aware that they did not comprehend a text paragraph, independent of the learning partner's knowledge of the paragraph, asked more questions than learners with visualized knowledge on the topic. When deficits of partners were visualized, learners also gave more explanations, irrespective of whether a learner knew something about the topic or not (Dehler et al., 2011). Thus, the group awareness tool used in this study guided learning behavior in so far as learners asked their partners for help or offered their help to ignorant peers. Furthermore, Sangin and colleagues (2011) showed that visualized differences of knowledge can guide students not only in their behavior, but also improve their learning outcomes. Discovering gaps in their knowledge through the comparison of quantity of knowledge (percentage) per learning module (providing peers' prior knowledge in a collaborative concept mapping phase) motivated students to converge to their partner's better knowledge status as the students achieved better learning outcomes than unaware learners after the collaboration. Since the different tools mentioned previously have used qualitative information (Dehler et al., 2011) or combined qualitative with quantitative information (Sangin et al., 2011), we assume that the availability of both types of information can increase the number of content-related questions and explanations (including an explorative investigation of possible interaction effects).

Hypothesis 1: Learners supported by qualitative and/or quantitative cognitive information ask more questions than learners without such support.

Hypothesis 2: Learners supported by qualitative and/or quantitative cognitive information explain more than learners without such support.

Concerning partner modeling, the relevance of reducing the effort for grounding processes in collaborative learning was highlighted (Bodemer, 2011). Bodemer (2011) found that learners with group awareness support spent less time for grounding and modeling processes and were more involved in meaningful discussions than those without support. Sangin and colleagues (2011) further confirmed that the aforementioned increase of learning outcome depends on the positive impact of partner modeling. With the visualization of objective cues on their peers' prior knowledge, learners became more accurate in estimating their partner's knowledge, and this accuracy predicted higher learning outcomes (Sangin et al., 2011). Since the latter tool provides qualitative and quantitative awareness information, or rather more comprehensive information than the other visualizations, this leads us finally to the following assumption:

Hypothesis 3a: Learners supported by a combination of qualitative and quantitative information estimate their partners' knowledge more accurately than learners with availability to solely qualitative information.

Hypothesis 3b: Learners supported by a combination of qualitative and quantitative information estimate their partners' knowledge more accurately than learners with availability to solely quantitative information.

Methods

We conducted an empirical study to investigate the impact of different types of awareness information on learning behavior. After an individual phase of text reading to learn about biogas plants, students were asked to communicate with a bogus learning partner, of whose non genuine nature they were not aware. Therefore, the learning environment offered the participants to write down questions and/or explanations. During this 'collaboration', the learners were supported (or not) by a knowledge profile that allowed them to compare their self-assessed knowledge on eight concepts about biogas plants with their partner's knowledge. The respective bogus partners' knowledge was assigned by an algorithm aiming to cover all possible knowledge combinations with random values. As seen in Figure 1, we offered four different versions of visualizations: (1) no profile (qual-/quan-), (2) a profile displaying the qualitative information *what* concepts both know about or not (qual+/quan-), (3) one displaying the quantitative information *how much* learners know on biogas plants in total (qual-/qual+),

and (4) another displaying a combination of qualitative and quantitative information on *how much* learners know about each concept (qual+/quan+). A list of concepts was available in each condition.

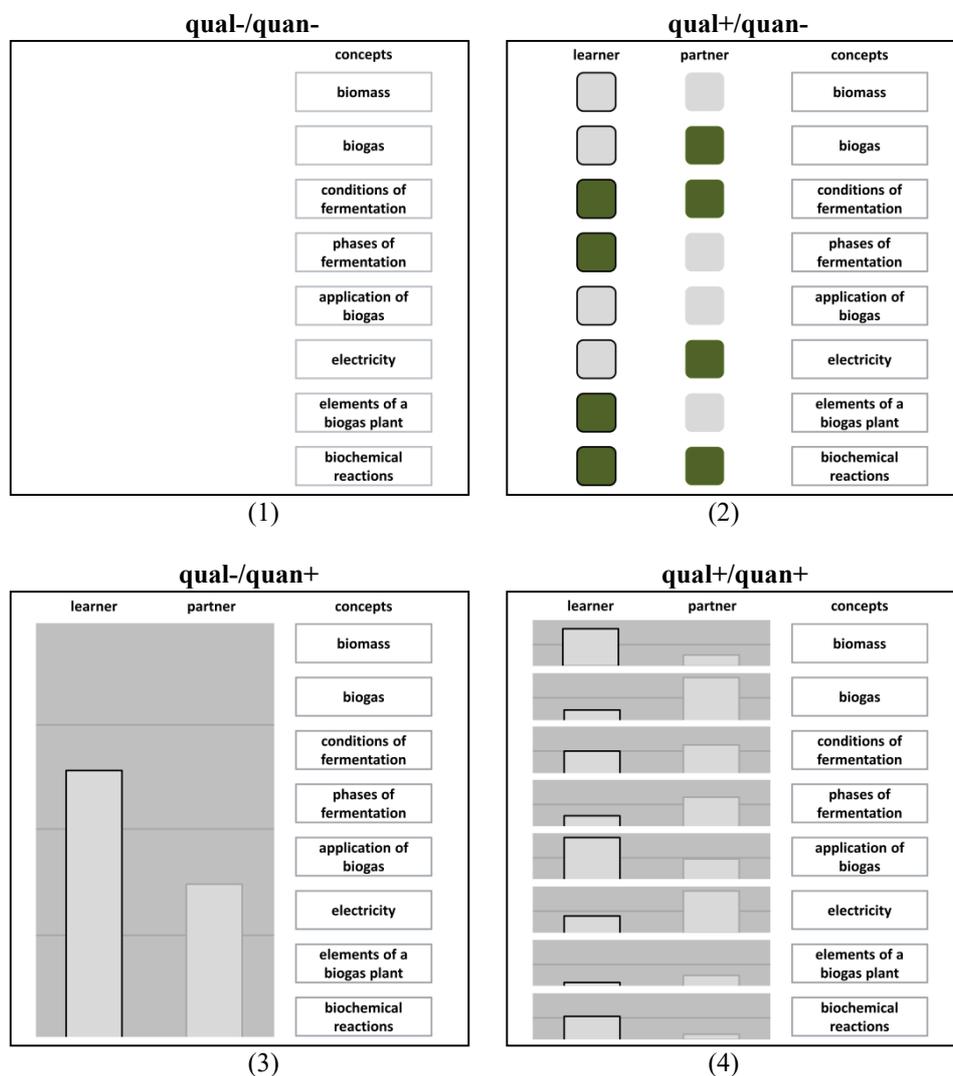


Figure 1. Awareness conditions: no information (1); qualitative information on present (dark boxes) and missing (light boxes) knowledge (2); quantitative information as bar chart (3); combination of both (4).

Sample and design

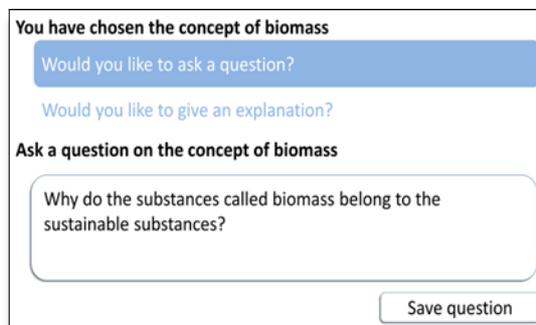
The sample consisted of 51 persons, mainly university students, 18 to 30 years old ($M = 23.18$, $SD = 3.06$). The participants were randomly assigned to one of the four experimental conditions in a 2x2 factorial design. As you can see from Figure 1, we varied the graphical knowledge representation concerning its qualitative dimension (no visualization of qualitative information vs. visualization of qualitative information) and its quantitative information (no visualization of quantitative information vs. visualization of quantitative information). As dependent variables, we captured the number of questions (hypothesis 1), explanations (hypothesis 2), and accurate estimations concerning the partner modeling distance (hypothesis 3). All three variables are further described in the next section.

Instruments and dependent variables

Number of questions and explanations

The visualization of a concept list in all conditions enabled us to assign learners' contributions to the respective concepts. The learning environment allowed subjects to select a concept and to choose between "Would you like to ask a question?" or "Would you like to give an explanation?". Depending on the selection, a text window

prompted the participant to either formulate a question or an explanation on the concept and to save the text afterwards. Figure 2 illustrates the choice to ask a question (featured by a dark background) and shows the input window with a learner's question. Participants could decide to contribute either a question or an explanation, both, or none at all for each concept. Subsequent to this bogus collaboration, we counted the numbers of questions (DV in hypothesis 1) and explanations (DV in hypothesis 2) per person.



The screenshot shows a web-based interface for asking a question. At the top, it says "You have chosen the concept of biomass". Below this, there are two buttons: "Would you like to ask a question?" (highlighted with a dark blue background) and "Would you like to give an explanation?" (light blue background). Underneath, the text "Ask a question on the concept of biomass" is displayed. A text input field contains the question: "Why do the substances called biomass belong to the sustainable substances?". A "Save question" button is located at the bottom right of the input area.

Figure 2. Learner's view when asking a question.

Estimations of knowledge, visualization of constellations, and calculation of distances

Learners' estimations of own knowledge were needed for two reasons: (1) for the visualization of the learners' 'real' knowledge about biogas-related concepts as a first component of their knowledge profile, and (2) as a basis to apply the algorithm that generates values to complement said profile with the visualization of the bogus learning partner's knowledge (if required by the test condition). Concerning (1), the learners had to rate their own degree of knowledge (with regard to their assumed competence to explain it to a learning partner) prior to the bogus collaboration via a six-point scale (reaching from 1 = 0% knowledge to 6 = 100%, to be rated separately for each out of eight concepts about biogas plants). To visualize the qual-/quan+ profiles, we created a total of all values given and visualized it in one aggregated bar chart per profile. The qual+/quan- profiles were realized by classifying the values from 0 to 40% as 'no knowledge given' of a concept and the values from 60 to 100% as 'knowledge given'; the former minor knowledge status was visualized as a light grey box and the latter major knowledge status as a darker green box. To create the qual+/quan+ profiles, we finally visualized each value by a bar chart for each concept. Referring to (2), we designed the algorithm to build on the values described before. The algorithm generated additional random values within a reasonable range that were included into the visualizations, taking into account that in both qual+ conditions each type of knowledge combinations should ideally occur in equal numbers. Furthermore, the learners were asked to additionally estimate their partner's knowledge after the collaboration. Therefore, they had to rate on a six-point scale reaching from 1 (0% knowledge) to 6 (100% knowledge) how much knowledge they award their partner concerning each concept about biogas plants. These estimations of bogus partner's knowledge were needed for the calculation of the partner modeling accuracy. Therefore, the values were compared to the visualized knowledge of the partner in order to determine the distance between both (DV in hypothesis 3).

Procedure

The experimental study was conducted in our research laboratory under controlled conditions. Each participant worked on the learning environment on a single computer. After welcoming and declaration of consent, we informed the participants that they have to read and memorize a learning text and collaborate with another participant in a computer-supported scenario. Then, we invited them to start the experiment and fill out a demographics questionnaire querying their age, sex, and level of education. Having completed these questions, the subjects were given 15 minutes to read the learning material on biogas plants. Following this learning phase, they were requested to estimate their own knowledge on each given subtopic of biogas plants, just as it was described in the last chapter. The subsequent bogus collaboration took ten minutes. In this phase, the knowledge profile of the respective condition was presented to the participants (cf. Figure 1), and they could make their contributions in the form of questions or explanations (cf. Figure 2). Finally, they estimated their bogus partner's knowledge. A graphical overview of the whole procedure can be found in Figure 3.

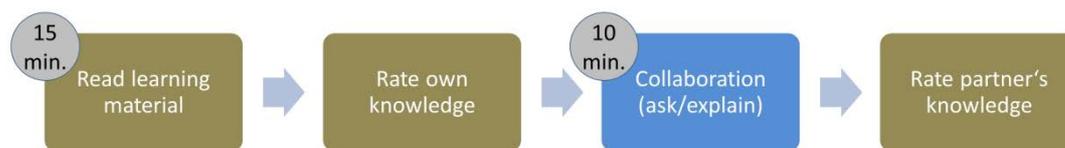


Figure 3. Graphical overview of the experiment's procedure.

Findings

For answering the question of how awareness information affects learning behavior, we observed the impact of qualitative and quantitative cognitive information on the number of contributions to a bogus communication and on the modeling of the learning partner. Two-factorial ANOVAs were used to investigate each single hypothesis, with qualitative information (no visualization of qualitative information vs visualization of qualitative information) and quantitative dimension (no visualization of quantitative information vs visualization of quantitative information) as between-subject factors. All effects are reported as significant at $p < .05$.

Impact of awareness information on the number of questions (hypothesis 1)

We hypothesized that learners supported by the visualization of qualitative information ask more about subject matters than learners without qualitative information. To test this, we observed the number of total questions per person as dependent variable. This total included solely complete sentences, since we excluded contributions of less than three words from the calculation. There was a significant main effect of available qualitative information (qual+/quan- and qual+/quan+) on the number of questions, $F(1, 47) = 4.12, p = .048, \eta_p^2 = .081$. Regarding the assumption that quantitative group awareness triggers questions, there was also a significant main effect of quantitative information (qual+/quan+ and qual-/quan+) on the number of questions, $F(1, 47) = 4.64, p = .036, \eta_p^2 = .090$. Finally, there was no interaction effect of qualitative and quantitative information on the number of questions, $F(1, 47) = 0.29, p = .590, \eta_p^2 = .006$. Table 1 shows the related descriptive statistics.

Table 1: Number of average questions in each condition.

awareness information	qual- ($n = 26$)		qual+ ($n = 25$)		both ($N = 51$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
quan- ($n = 26$)	2.15	2.34	3.08	1.80	2.62	2.09
quan+ ($n = 25$)	3.15	2.54	4.75	2.09	3.95	2.43
both ($N = 51$)	2.65	2.45	3.88	2.09	3.25	2.34

Since we wanted to identify which visualization of knowledge results in the most questions, we further used a two-factorial repeated measure ANOVA with visualized qualitative information only (qual+/quan- vs. qual+/quan+) as between subject factor and knowledge constellation as within-subject factor ((1) learner's deficit with partner's knowledge vs. (2) same knowledge (both have deficit or both have knowledge) vs. (3) partner's deficit with learners knowledge). As dependent variable, we observed the number of questions in each of said knowledge constellations divided by the total number of occurrences of this constellation per person (which was varying between subjects due to their self-assessment).

There was a significant main effect of knowledge constellation on the percentage of questions asked in each category, $F(2, 34) = 12.11, p < .001, \eta_p^2 = .416$. The pairwise comparison for the main effect of knowledge constellation indicates a significant difference ($p < .05$) between levels 1 and 2 (learner's deficit and same knowledge) and 1 and 3 (learner's deficit and partner's deficit) but not between levels 2 and 3 (same knowledge and partner's deficit). In 72.8% ($SD = 29.6\%$) of cases in which the comparison to their bogus collaboration partner enclosed missing knowledge, learners asked a question on the belonging concept. In contrast, only in 49.9% ($SD = 37.8\%$) of cases with same knowledge, and in 28% ($SD = 34.7\%$) of cases with better expertise than the partner triggered questions. This seems to indicate that learners asked a question if they recognized own missing knowledge compared to partner's knowledge. Furthermore, there was a significant effect of the knowledge profile on the weighted number of questions asked in each of the three categories of knowledge constellation, $F(1, 17) = 6.63, p = .020, \eta_p^2 = .281$. As you can see from Figure 4, the combination of qualitative and quantitative information appears to outperform solely qualitative cognitive information. Finally, there was no significant interaction effect of knowledge constellation and qualitative information on the weighted number of questions, $F(2, 34) = 0.83, p = .447, \eta_p^2 = .046$.

Impact of awareness information on the number of explanations (hypothesis 2)

We tested the assumption that learners supported by a profile with qualitative information explain more about subject matters than learners without access to qualitative information by investigating the number of total explanations as dependent variable. This total included solely complete sentences again. There was a significant main effect of qualitative information on the number of explanations, $F(1, 47) = 4.44$, $p = .041$, $\eta_p^2 = .086$. Regarding the further assumption that the presence of quantitative information also triggers explanations, there was no significant main effect of the quantitative dimension on the number of explanations, $F(1, 47) = 3.45$, $p = .069$, $\eta_p^2 = .068$. Furthermore, there was no interaction effect between qualitative and quantitative

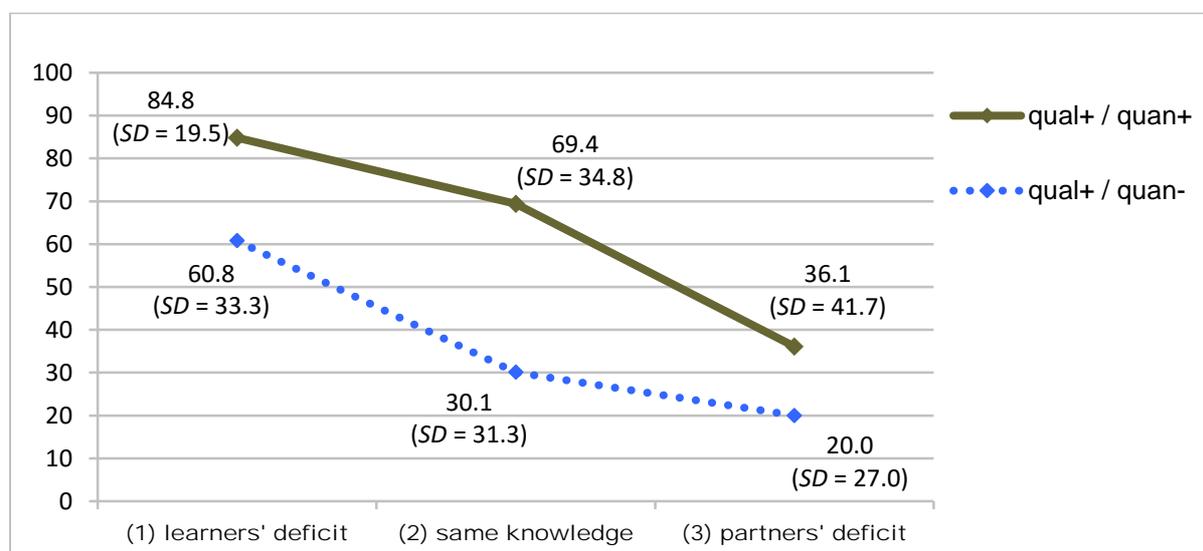


Figure 4. Percentage of questions asked per knowledge constellation weighted by the total number of occurrences of this constellation per person.

information on the number of questions, $F(1, 47) = .85$, $p = .179$, $\eta_p^2 = .038$. Table 2 shows the descriptive statistics, disclosing that, in contrast to our first assumption, most explanations were given, if no qualitative information was available.

Table 2: Number of average explanations in each condition.

awareness information	qual- ($n = 26$)		qual+ ($n = 25$)		both ($N = 51$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
quan- ($n = 26$)	4.61	1.45	3.00	1.35	3.81	1.60
quan+ ($n = 25$)	4.85	1.67	4.50	2.11	4.68	1.86
both ($N = 51$)	4.73	1.54	3.72	1.88	4.24	1.77

Impact of awareness information on partner modeling (hypothesis 3)

We used independent t -tests to test the assumption that learners supported by a combination of qualitative and quantitative information estimate their learning partners' knowledge more accurately than learners with visualizations of either qualitative or quantitative information. This derogation from the main research design was required to separately compare (a) both qual+ conditions (qual+/quan- vs. qual+/quan+) with each other, and (b) both quan+ conditions (qual-/quan+ vs. qual+/quan+). The control group (qual-/quan-) was excluded from both calculations, since there was no partner information available in this condition. To observe the accuracy of partner modeling as dependent variable, we calculated the distance between the sums of estimated partner's knowledge (which were additionally divided by 8 in the case of the quan+ conditions) and algorithm-generated values visualized in the knowledge profile. The smaller the value, the better was the accuracy.

Concerning hypothesis 3a, participants supported by the combination of qualitative and quantitative information ($M = 1.33$, $SD = 0.65$) estimated their partners more accurately than participants with solely qualitative information ($M = 1.62$, $SD = 0.89$). However, this difference was not statistically significant $t(15) = 0.50$, $p = .621$, $r = .13$ (hypothesis 3a). Regarding hypothesis 3b, participants with available qualitative and quantitative information differed significantly from participants with support of quantitative information,

$t(23) = 2.83, p = .010, r = .51$. Learners estimated their partners more accurately if both types of information were available ($M = .48, SD = 0.32$) than without the combined visualization ($M = 1.34, SD = 1.01$).

Conclusions and implications

We conducted this empirical study to systematically investigate the impact of visualizations of qualitative and quantitative cognitive information on collaborative learning. Both types of information are systematically gathered and made available by cognitive group awareness tools, allowing learners for comparing to others, shaping learning mechanisms, and facilitating the control and (self-)regulation of their learning. Our assumption that the visualization of both qualitative and quantitative information moves learners to ask more questions in general was confirmed. With an additional exploration of the two conditions offering qualitative information, we further showed that visualizations supporting the awareness of one's own deficit (and partner's knowledge) had a significantly higher impact on questioning behavior than visualizations of the same or (binary / discrete) visualizations of more knowledge than the partner. Furthermore, we could show that this effect was significantly stronger if learners were supported by the combination of qualitative and quantitative information. This only partially confirms the results of Dehler and colleagues (2011), since they found that the awareness of one's knowledge gaps (independent of partner's knowledge) is sufficient to trigger questions, meaning that visualized gaps encouraged learners to ask partners for help. Our results indicate that besides the awareness of own gaps the concurrent awareness of the partner being knowledgeable about the concept might be relevant to shape the questioning behavior. Indeed, the bogus scenario requires some caution regarding the interpretation so that further research is needed in order to clarify the question what specific role is played by qualitative and quantitative information.

Concerning the number of explanations, it could have been expected that the visualization of partner deficits guides learners to give more explanations (Dehler et al., 2011). In contrast to the assumption that learners supported by qualitative information explain more than learners without such support, we could find that especially the absence of qualitative information significantly increased the number of explanations. Even available quantitative information seemed to lead to more explanations, if they were presented without additional qualitative information. One explanation as to why learners give fewer explanations with increasing level of informational detail might be that they can better choose relevant cases and interact in a more goal-oriented way. Furthermore, another learning mechanism central to collaborative learning could be of relevance, namely (self-)explanation (cf. Dillenbourg, 1999). The absence of information may drive learners to remember the learning text by paraphrasing it in the text windows provided during the collaboration phase of the study. However, further research is needed here to investigate this concern more deeply, not least because of the artificial scenario without a real learning partner.

Finally, we investigated whether learners supported by a visualization that combines qualitative and quantitative information estimate their partners' knowledge more accurately than learners with availability of single cognitive information. It is suggested by a large body of research that the absence of cognitive information results in matching own knowledge with estimations of partner (cf. Sangin et al., 2011). Small accuracy values in our study indicated that the estimations of learners were largely based on the visualizations presented to them. Although especially the availability of combined qualitative and quantitative information appeared to be useful in this context, it needs to be clarified, if the presentation of qualitative information is sufficient to support partner modeling. Taking cognitive load (cf. van Merriënboer & Sweller, 2005) into account, it might be that visualizations combining qualitative and quantitative information are associated with additional mental effort.

Overall, we can derive from these results that qualitative and quantitative information visualized by cognitive group awareness tools is suitable to elicit a desired collaborative learning behavior. It is often mentioned that prescribing collaborative activities might carry the risk of demotivating the learners (Hesse, 2007). Group awareness tools offer an opportunity to satisfy the resulting need for implicit guidance and learners' empowerment associated therewith. Tracing back to Piaget's approach, it is recommended to use knowledge profiles that combine qualitative and quantitative information where the intention is to make learners aware of cognitive conflicts that elicit a meaningful exchange between them. The same applies to the support of partner modeling which is needed in a Vygotskian sense to maintain an accurate model of the learning partners and optimize internalization processes.

References

- Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior*, 27(3), 1079–1086. <http://doi.org/10.1016/j.chb.2010.05.016>
- Bodemer, D., & Buder, J. (2006). Supporting Collaborative Learning with Augmented Group Awareness Tools, 3(2), 123–139. <http://doi.org/10.1007/s11412-008-9037-5>

- Bromme, R., Hesse, F. W., & Spada, H. (2005). Barriers, Biases and Opportunities of Communication and Cooperation with Computers: Introduction and Overview. In R. Bromme, F. W. Hesse, & H. Spada (Eds.), *Barriers and Biases in Computer-Mediated Knowledge Communication* (Vol. 5, pp. 1–14). New York: Springer-Verlag. http://www.springerlink.com/index/10.1007/0-387-24319-4_1
- Buder, J., Schwind, C., Rudat, A., & Bodemer, D. (2015). Selective reading of large online forum discussions: The impact of rating visualizations on navigation and learning. *Computers in Human Behavior*, *44*, 191–201. <http://doi.org/10.1016/j.chb.2014.11.043>
- Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2009). Providing group knowledge awareness in computer-supported collaborative learning: Insights into learning mechanisms. *Research and Practice in Technology Enhanced Learning: RPTEL*, *4*(2), 111–132.
- Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, *27*(3), 1068–1078. <http://doi.org/10.1016/j.chb.2010.05.018>
- Dillenbourg, P. (1999). What do you mean by “collaborative learning”? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp. 1–19). Oxford, UK: Elsevier.
- Doise, W., & Mugny, G. (1984). *The Social Development of the Intellect*. Amsterdam: Pergamon Press.
- Erkens, M., & Bodemer, D. (2015). Potentials of Text Mining based Technologies for Improving Cognitive Group Awareness. In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015* (Vol. 2). Gothenburg, Sweden.
- Hesse, F. (2007). Being told to do something or just being aware of something? An alternative approach to scripting in CSCL. In F. Fischer, I. Kollar, H. Mandl, & J. Haake (Eds.), *Scripting computer-supported communication of knowledge-cognitive, computational and educational perspectives* (pp. 91–98). New York: Springer.
- Janssen, J., & Bodemer, D. (2013). Coordinated Computer-Supported Collaborative Learning: Awareness and Awareness Tools. *Educational Psychologist*, *48*(1), 40–55. <http://doi.org/10.1080/00461520.2012.749153>
- Piaget, J. (1977). *The development of thought: equilibration of cognitive structures*. Viking Press.
- Piaget, J., & Inhelder, B. (2008). *The Psychology Of The Child*. Basic Books.
- Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior*, *27*(3), 1059–1067. <http://doi.org/10.1016/j.chb.2010.05.032>
- Schittekatte, M., & Van Hiel, A. (1996). Effects of partially shared information and awareness of unshared information on information sampling. *Small Group Research*, *27*, 431–449. <http://doi.org/10.1177/1046496496273006>
- Schnaubert, L., & Bodemer, D. (2015). Subjective Validity Ratings to Support Shared Knowledge Construction in CSCL. In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015* (Vol. 2). Gothenburg, Sweden.
- Tudge, J., & Rogoff, B. (1989). Peer influences on Cognitive Development: Piagetian and Vygotskian Perspectives. In M. H. Bornstein & J. S. Bruner (Eds.), *Interaction in Human Development*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, *17*(2), 147–177. <http://doi.org/10.1007/s10648-005-3951-0>

Exposing Piaget's Scheme: Empirical Evidence for the Ontogenesis of Coordination in Learning a Mathematical Concept

Dor Abrahamson, University of California, Berkeley, dor@berkeley.edu
Shakila Shayan, Utrecht University, s.shayan@uu.nl
Arthur Bakker, Utrecht University, A.Bakker4@uu.nl
Marieke F. van der Schaaf, Utrecht University, M.F.vanderSchaaf@uu.nl

Abstract: The combination of two methodological resources—natural-user interfaces (NUI) and multimodal learning analytics (MMLA)—is creating opportunities for educational researchers to empirically evaluate seminal models for the hypothetical emergence of concepts from situated sensorimotor activity. 76 participants (9-14 yo) solved tablet-based non-symbolic manipulation tasks designed to foster grounded meanings for the mathematical concept of proportional equivalence. Data gathered in task-based semi-structured clinical interviews included action logging, eye-gaze tracking, and videography. Successful task performance coincided with spontaneous appearance of stable dynamical gaze-path patterns soon followed by multimodal articulation of strategy. Significantly, gaze patterns included uncued non-salient screen locations. We present cumulative results to argue that these ‘attentional anchors’ mediated participants’ problem solving. We interpret the findings as enabling us to revisit, support, refine, and elaborate on central claims of Piaget’s theory of genetic epistemology and in particular his insistence on the role of situated motor-action coordination in the process of reflective abstraction.

Keywords: attentional anchor, coordination, eye-tracking, genetic epistemology, NUI, proportion

Background and objectives: Revitalizing LS interest in genetic epistemology as complementary to sociocultural models of conceptual development

The eminent cognitive-developmental psychologist Jean Piaget, who would be celebrating his 120th birthday this August 2016, has had a rocky career in the Learning Sciences. Despite a near-centennial stretch of prodigious, paradigm-changing academic oeuvre, despite the omnipresence of constructivist educational parlance in preK-12 STEM rhetoric, and despite his indirect yet formative and enduring mark on the design of commercial pedagogical products for discovery-based learning, Piaget’s groundbreaking construct of a schema has received some bad press. The construct suffers, perhaps, via its too-convenient association with Piaget’s oft-critiqued yet oft-misunderstood Stage Theory or the indefatigable attacks on the validity of his clinical methodologies. But whereas Piaget-bashing has generated many a dissertation and built entire research programs, his theoretical constructs and model of conceptual schemata rising from sensorimotor operatory schemes, we posit, have yet to find their match as explanantia for meaningful situated learning. At the very least, we concede, the waning of empirical Piagetian research is hampering our field’s intellectual progress and increasingly vitiating its relevance to the changing terrain of educational media (Abrahamson, 2015; Abrahamson & Sánchez-García, in press).

We are calling to renew Piagetian discourse specifically on mathematical learning, and more specifically, mathematics learning with state-of-the-art interactive media (Forman, 1988; Lindgren & Johnson-Glenberg, 2013; Marshall, Antle, van den Hoven, & Rogers, 2013; Moreno-Armella, Hegedus, & Kaput, 2008; Sarama & Clements, 2009). Even more specifically, we are looking for forms of empirical research in environments where both student and researcher respective activities avail of multimodality, with the student engaging in explorative activity that the researcher monitors, documents, measures, and analyzes, even in real time (Martin & Sherin, 2013; Schneider, Bumbacher, & Blikstein, 2015; Worsley & Blikstein, 2014). This brave new world of multimodality in design, instruction, and research demands theoretical infrastructure for thinking seriously, anew, about situated motor-action skill acquisition as it relates to conceptual development. In turn, we are thus also looking to draw on a century of progress in the somatic–kinesiological disciplines (Bernstein, 1996; Kelso & Engström, 2006; Newell & Ranganathan, 2010; Thelen & Smith, 1994) as these bear on the action-to-concept learning process (Bamberger, 2013). In fact, we will argue that Piaget’s genetic epistemology is key to populating learning-sciences discourse with this diversity of fresh, pertinent, and resonant perspectives. In a sense, we are stepping back to jump forward.

Up front, we wish to clarify that our call is to build on, rather than supplant, a research tradition of treating

Piagetian themes through qualitative analysis (Abrahamson, 2012; Dubinsky, 1991; Gray & Tall, 1994; Norton, 2008). In fact, it is precisely these types of investigations that we wish further to pursue by introducing new constructs and methodological techniques.

To contextualize and substantiate this call, we present and discuss empirical data collected during the implementation of experimental educational interventions, in which young study participants were engaged in technologically enabled embodied-interaction activities designed to foster presymbolic proportional reasoning. We will argue for the unique and pivotal traction of Piaget's thesis on our research by way of explaining the critical role that his constructs of sensorimotor scheme and reflective abstraction served in making sense of our data. Namely, the Piagetian perspective enabled us to posit the significance of nuanced changes in children's sensorimotor activity for their conceptual ontogenesis as well as the implications of these findings for both theory and practice of mathematics education. Please note: This proceedings paper is a précis for a full-length peer-reviewed journal article (Abrahamson et al., in press), where we are able to provide far more information on theory, methods, and results.

Theoretical framework

Piaget (1896-1980) was fascinated by children's opportunities for personal development through engaging in the social enactment of cultural practice, such as moral development through game play. However he viewed culture, along with its social agents, practices, norms, and material artifacts, moreso as the setting and playing field of ontogenesis than as its very fabric and constitution.

In the latter 20th century, a slew of monographs inspired by the cultural-historical psychology of Lev Vygotsky (1896-1934) impressed upon our intellectual community a set of views not readily perceived as concordant with Piaget's epistemological theory. Instead of foregrounding the child's piecemeal construction of cognitive structures, these views underscored the critical role of sociocultural activity structures as shaping individual disciplinary enskilment, such as mathematical competence. These alternative views include the theorization of: (a) artifact appropriation and contextual adaptation as the *sine qua non* of mediated maturation into communal techno-scientific practice, including visualizations, orientations, and discourse (Newman, Griffin, & Cole, 1989; Saxe, 2012; Wertsch, 1979); (b) learning as legitimate peripheral participation in the social co-enactment of purposeful cultural practice (Lave & Wenger, 1991; Rogoff, 1990); and (c) discourse as the vehicle and substance of knowing (Sfard, 2010).

Scholars holding constructivist views of cognition have retaliated that, notwithstanding, meaning must be grounded in tacit, presymbolic sensorimotor routines and innate/early cognitive capacity (Allen & Bickhard, 2013; Denison & Xu, 2014; Harnad, 1990) and concepts are built painstakingly by coordinating multiple personal and situated resources for *ad hoc* productive engagement (Case & Okamoto, 1996; Noss & Hoyles, 1996; Smith, diSessa, & Roschelle, 1993). In the fray of this grand altercation some scholars are looking to forge dialogue between these die-hard entrenched camps (Abrahamson, 2012; Cole & Wertsch, 1996; diSessa, Levin, & Brown, 2015). By and large, though, the field is at a stalemate, with each faction chiding the other, "Show me!"

We have something new to show that might jostle the field out of its stalemate. We reasoned that if only we could demonstrate empirically student behaviors that are better accounted for by constructivist than sociocultural theory, then we would be in a better position to argue for a dialectical view of mathematical learning—a view of learning as action-based ontogenesis in facilitated settings—settings that we regard as culturally-historically evolved instrumented fields of promoted action (Abrahamson & Trninic, 2015). Moreover, by way of reemphasizing the critical role of sensorimotor activity in conceptual development, we could justify an introduction, into learning-sciences discourse, of evolving models of teaching and learning imported from disciplines focused on motor-action skill development and methodology (Abrahamson & Sánchez-García, in press).

Our renewed interest in Piaget's theory of learning, and in particular his model of reflective abstraction, emerged from unexpected quarters. Namely, philosophers of radical enactive cognition (Chemero, 2009; Hutto & Myin, 2013), who reject exclusively representationalist epistemologies, have been seeking corroboration via partnerships with social scientists engaged in the empiricism of skill acquisition. For example, Hutto and Sánchez-García (2015), respectively a philosopher of cognition and a sociologist of sport, collaborated in articulating a radical-enactivist interpretation of athletic performance. In particular they developed the construct of an attentional anchor, which then became central to our own work (Abrahamson & Sánchez-García, in press), as we now explain.

An *attentional anchor* (AA) is a phenomenological aspect of an agent's implicit or explicit goal-oriented interaction with the environment. AAs may be a specific object (real or imagined), area, or other pattern or behavior of the perceptual manifold that an agent detects, invokes, selects, and uses to enact the activity at hand. For example, a juggler struggling to coordinate the trajectories of many balls simultaneously flying through the

air might imagine a tall rectangle rising in front of her and aim the balls to its vertices. Whether discovered or taught, the AA interpolates itself into the agent–environment relation to serve as an enabling task constraint—it becomes a new systemic element that hones and channels attention during perception–action couplings. The AA reduces operational complexity, rendering ergonomic and feasible otherwise overwhelming tasks. An agent acting on an AA experiences it as a “steering wheel” overlaid upon the perceptual field—the attentional anchor becomes the mediating proxy both for operating on the environment and interpreting feedback from the environment. Specifically, the AA brings forth to the agent new latent affordances by objectifying, specifying, and foregrounding the environment’s task-oriented, invariant goal information structures (Chow et al., 2007; Kelso & Engström, 2006; Newell & Ranganathan, 2010). As such, the AA thesis resonates well also with central tenets of Enactivism (Varela, Thompson, & Rosch, 1991).

We are intrigued and motivated by an apparent resemblance of the constructs *motor-action coordination* and *attentional anchor* from ecological-dynamics theory to those of *coordination* and *category* in Piaget’s model of reflective abstraction. Reflective abstraction is the highest of three abstraction levels that Piaget distinguished. It concerns a learner’s coordination of actions. Piaget viewed this type of abstraction as ‘constructive’ in the sense that new syntheses emerge that bring forth interaction regularities toward encapsulating and generalizing new capacity. In short, reflective abstraction is “the construction of mental objects and of mental actions on these objects” (Dubinsky, 1991, p. 101). It should come as no surprise that Piaget’s work resonates with dynamical-systems theory, given his deep commitment to anti-representationalist, situated structuralism (Piaget, 1970; Turner, 1973). In the remainder of the paper we will discuss an empirical study, in which we have tracked what appear to be children’s sensorimotor behaviors that mark a new coordination focused on an attentional anchor; a coordination leading to the reflective abstraction of a higher-order functional structure and its conscious articulation as a new phenomenal category. As we explain, this situated cognitive process is pivotal to a designed activity on proportional relations.

Methods

In total, 76 volunteering students from the Netherlands participated in two studies. Study 1 included 30 students of 5th/6th grade (mean age = 11[3] years; 13 male, 17 female) from five elementary schools. They all worked on the Parallel Bars activity (see Figure 1a). Study 2 included 46 students of 7th/8th grade (29 male, 17 female; mean age = 13[5]) from two prevocational schools. Of these, one group (26 students) only worked on “parallel tasks”—Parallel Pluses (Figure 1a) followed by Parallel Bars (Figure 1b); the other group (20 students) only worked on “orthogonal tasks”—Orthogonal Pluses (see Figure 1c) followed by Orthogonal Bars (see Figure 1d).

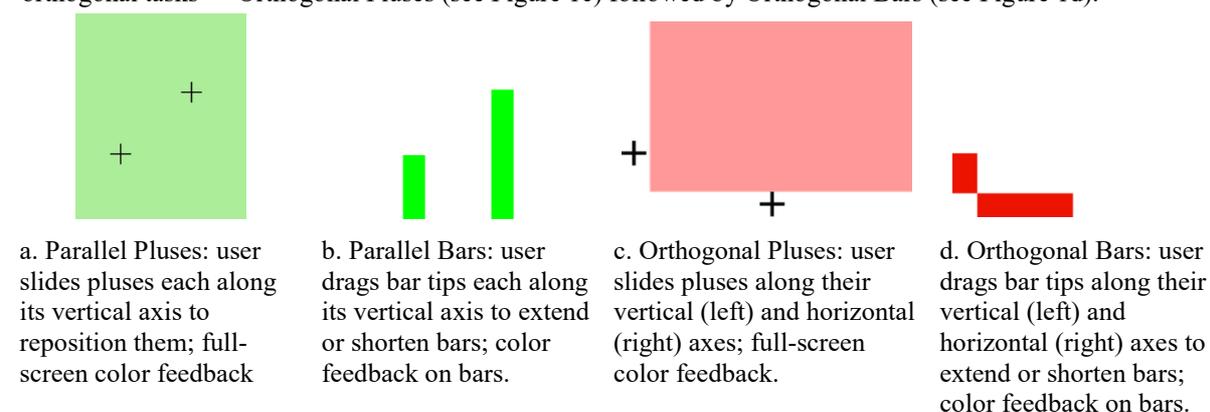


Figure 1. Sample screenshots from enacting four activity modules in the touch-screen tablet application. To make the screen green, participants had to manipulate either the positions of cursors (1a, 1c) or the extension of bars (1b, 1d) either along parallel (1a, 1b) or Cartesian axes (1c, 1d).

The tasks were variations on the Mathematical Imagery Trainer for Proportion (Reinholz, Trninic, Howison, & Abrahamson, 2010), an interactive technological device designed for students first to develop new operatory schemes underlying mathematical concepts and then mathematize these schemes using standard frames of references (e.g., a grid, numerals). The task is implemented in a multi-touch tablet, with each hand (or each index finger) controlling one element on the screen, either a plus-shaped cursor (the “plus” task conditions) or the edge of a stretch/shrink rectangle (the “bars” task conditions). The task objective is to move these elements on the screen so as to achieve a specified goal state: keeping green either the whole screen (“plus”) or elements thereof (“bars”). The software mediating between user-action input and screen-color output instantiates mathematical

datum point (e.g., 10 and 20 cm, respectively, above the screen base) then calculates their quotient (e.g., 10/20). A match with a preset ratio (e.g., 1:2) makes for green, otherwise red (Figure 1). Thus in the case of a 1:2 ratio, users might move their index fingers along the screen constantly keeping the right-hand double as high as the left (“parallel” conditions, Figures 1a, 1b) or double as far from the origin (“orthogonal” conditions, Figs. 1c, 1d), or they might attend to other properties of the performance, such as the distance between their hands or their speeds (Abrahamson et al., 2014).

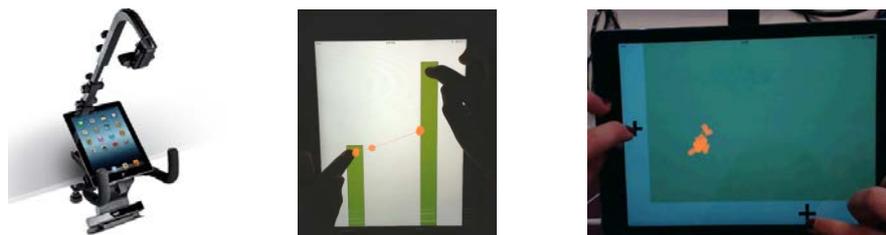


Figure 2. On left: Tobii Mobile Device Stand for X2. The stand is attached to the edge of a desk. The iPad is positioned in the center. The eye-tracker is placed on the stand base, with the camera on the top. Center: sample integrated eye-tracking and video data from the Parallel Bars condition. On right: Orthogonal Pluses data sample (left hand moves the plus up/down; right hand moves the plus right/left).

The intervention and analysis followed principles of task-based semi-structured clinical interviews (e.g., Ginsburg, 1997). Our data set comprises videography (of student actions and multimodal student–tutor discourse), streaming logs of touchscreen activity, and eye-gaze tracking (see Figure 2). This complex data constellation was designed so as to serve us in developing a more detailed and comprehensive theoretical model for the spontaneous emergence of new sensorimotor coordinations grounding mathematical conceptions. We used visualization software that superimposes the eye-tracking paths onto the videography, so that we could see which particular locations on the screen were in the users’ foveal vision as they were manipulating the virtual objects in dialogue with the researcher. Computational analyses of users’ visual pathways on the screen fed into micro-analyses of their concurrent actions and multimodal utterance (Siegler, 2006). The analyses enabled us to discern general patterns in students’ search for, and articulation of, effective bimanual manipulation strategies. We were particularly interested in implicating the emergence of attentional anchors that first support the bimanual motor-action then come forth into dyadic discourse as new mathematical objects and solution procedures. Furthermore, we evaluated whether Piaget’s four phases of reflective abstraction—interiorization, coordination, encapsulation, and generalization—could be genuinely implicated and differentiated in the data as depictive markers parsing students’ activity flow. For reliability, two researchers independently analyzed part of the data corpus, then shared their findings, and finally watched the videos repeatedly until reaching agreement over all their observations.

Results

Cumulative findings: A Piagetian analysis of learning proportion as reflective abstraction

Both within and across age and condition groups, students differed along several dimensions relevant to the study, including: (a) duration of time elapsed until discovery of a first effective interaction routine; (b) time to complete the whole task; and (c) pace of finger movement (fast or slow) at the initial exploration phase. Participants also differed in the incorrect rules they initially posited, their eye-gaze patterns accompanying successful hand coordination strategies, and their lines of reasoning toward effective solutions. These individual differences notwithstanding, the progress of *all* participants through the activity bore the pattern presented in Table 1. The discoveries students made en route to figuring out “green” interaction rules replicate our earlier findings (Reinholz et al., 2010). However adding eye-tracking visualization into the data manifold now enables us better to model the emergence of these discoveries from students’ interactions and characterize the discoveries in terms of reflective abstraction phases.

Table 1: Participants’ cross-condition behavioral sequence follows Piaget’s reflective abstraction phases

1. Interiorization: Exploring task environment	Students: (a) explored the task environment without any clearly discernable plan or strategy; (b) found greens haphazardly; (c) could not replicate green positions; (d) attempted strategies that did not bear out, e.g., moving fingers in equal pace; (e) realized there should be a spatial relation between
--	--

	the hands; (f) attempted to coordinate actions. Concurrently, eye gaze shifted between the moving fingertips.
2. Coordination: stable sensory patterns emerge concurrent with effective motor-action performance	In the course of attempting to develop an effective bimanual dynamical motor-action scheme for keeping green, gaze patterns emerged (see Figure 3) that: (a) followed tentative localized discovery of effective positions and constraints on action; (b) manifested as iterated rapid shifts among specific interface elements; (c) included at least one un-manipulated point; (d) settled on consistent, stable, and reoccurring forms; (e) coincided with significant improvement in overall performance; (d) coincided with more continuous as opposed to abrupt motor action; (f) enabled to reconstruct/replicate/repair previous green locations; and (g) preceded logical-mathematical reflective reasoning, discovery, or articulation of rules.
3. Encapsulation: Articulating sensorimotor patterns results in objectifying tacit elements, enhanced performance	Probed to articulate their strategy, students objectified an attentional anchor and then elaborated on it, forming new conjectures. Initially, though, their conjectures tended to belie their actions, such as speaking of a fixed distance between the moving fingers in the Parallel conditions, whereas in fact they had been changing the distance covariate with height. As they enacted their thoughts, however, they gradually came to appreciate the error, such as noticing that the distance in fact increases with height. After several replications they expressed their inference, such as saying, “No I was wrong.” At times, the experimenter guided this process by either challenging students or orienting them on critical features in the visual display. In turn, process of articulating and evaluating effective strategies resulted in better performance.
4. Generalization: From iterated, qualitative-process rule to explicit functional rule: articulating a latent mathematical relation as a constant property	Once students had validated an effective strategy, their actions were no longer explorative. Their gaze pattern intensified, e.g., more consistent and more rapid eye-gaze shifts along the triangular attentional anchor in Parallel Bars. Concurrently, their utterance included qualitative properties of objects and prospective actions. Introducing the grid precipitated a shift toward quantitative reasoning, e.g., “When they are lower they are one line apart, when they are in the middle they are more lines apart, and when the right hand is at the very top they are most apart.” Supplementing the numerals resulted in students unpacking the bimanual composite into ordered pairs, e.g., left at 1, right at 2; left at 2, right at 4; etc. Eventually they recognized a constant intra-pair quantitative (multiplicative) relation, e.g., “Oh wait it’s a half... I know it’s a half, the left is always half of the right.” They thus shifted from a scalar, inter-position process rule for iterated enactment (the higher you go, the bigger the distance) to an explicit intra-position functional rule with predictive power (wherever right is, left is half). That is, they articulated the notion of a constant ratio that underlies proportional equivalence.

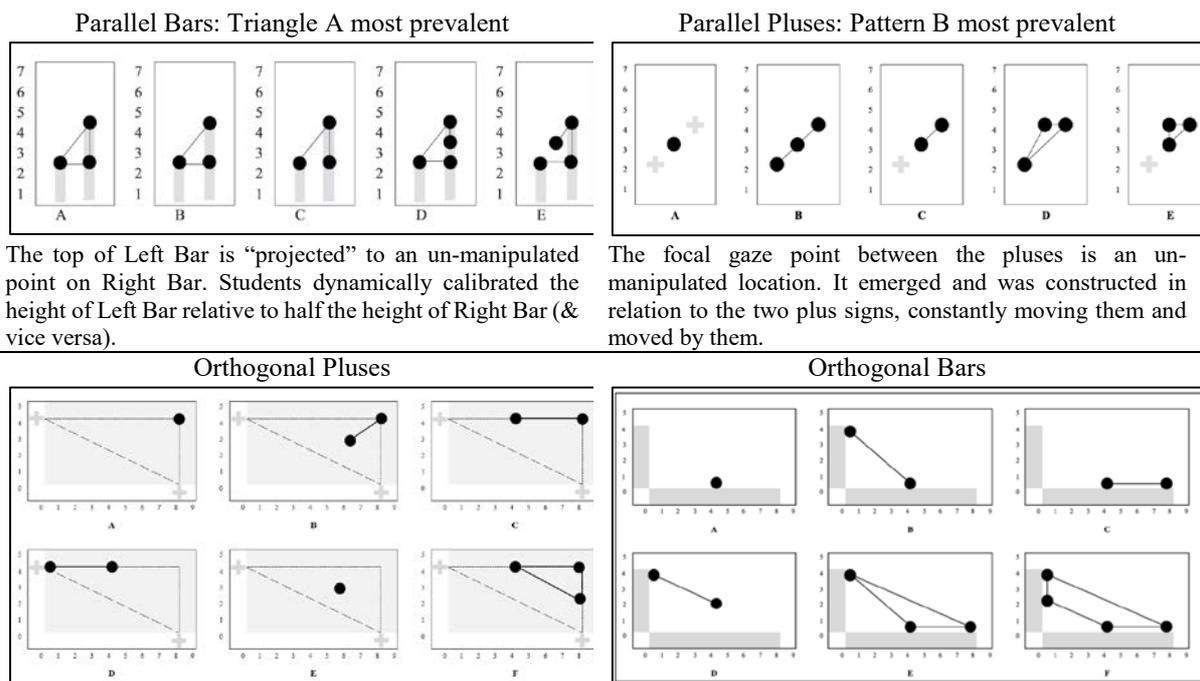


Figure 3. Schematic overview of the variety of emergent dynamical gaze patterns reveals attentional anchors.

In Figure 3, circles represent focal gaze points, lines are gaze paths. *Triangulated with the tablet action-logging and clinical data, we interpret these gaze patterns as evidence for ecologically coupled sensorimotor attentional anchors mediating effective enactment of problem solutions for the embodied-interaction task.* In our

collaborative-analysis sessions, as we watched the superimposed gaze/video data, we have been compelled by the dynamical evolution of these forms, and in particular when we played these movies in fast motion: It is as if bits and pieces of a would-be instrument—a handle or steering wheel—assemble in the task environment as solution means; as actionable media “between” student and objective. These media, the attentional anchors, emerge via co-evolving dialectical process: attentional anchors are invented *for and by* the sensorimotor scheme that yields it as a means of accomplishing the situated task objective. As the interview advances, the attentional anchors ascend: from latent aspects of the task environment; to tacit, dynamical, ecologically coupling patterns; to bonafide articles of discourse and reasoning.

Summary: Beyond representations—appreciating Piaget as a non-cognitivist

The structure that the child constructs through goal-oriented engagement in the task environment is not a representation in the sense of some accessible mental content in her head. Rather, the structure is a cognitive construct, a tacit relation that emerges between the subject and the objective world through adaptive efforts toward equilibrating effective engagement. The structure functions as a dynamical systemic dialectic, by which are formed both the subject’s schematized action routines of engagement with the world and, reciprocally, those worldly categories being engaged—aspects of the world toward which this schematized sensorimotor activity is oriented and transforming; categories by which the child is effecting aspects of the environment. In the particularities of the child’s engagement with the MIT-P technological system, the emergent operatory schemes are correlational. For example, in the case of the Parallel Pluses the reciprocally emergent category upon which these schemes are operating is often the interval between the hands. The emergent correlational manipulation of the interval coordinates two operations upon it—transforming its elevation, transforming its size—so that the higher the interval is (or the farther it is along the screen), the bigger it should be so as to effect and maintain the desired worldly state (making and keeping the screen green). This covariational coordination is created through a process Piaget called reflective abstraction, that is, the construction of a higher-order operational structure—the organization of a new phenomenal invariance that breaks away from, yet contains and coordinates, existing routinized operations that hitherto had been sufficient for productive engagement with simpler categories yet hence prove insufficient. To iterate, this coordination is centered on the new category, the interval between the hands.

Looking at results from implementing the MIT-P system in an eye-tracking study, we have attempted to make sense of our data from this Piagetian perspective. In particular, we have been curious about shifts in students’ visual attention toward the objects they are manipulating—shifts that co-occur with, or briefly anticipate, an apparent organization of new action patterns as well as the multimodal discursive articulation of these patterns into proto-mathematical propositions. Emblematic of these pattern shifts is that students will incorporate into their new routine a visual attention toward a location on the screen that is not a constituent part of the objects being manipulated. For example, they may stare at a point, a blank locus between two objects that they are manipulating—a point that apparently is strategic for constructing and applying the new coordination, such as the “higher-bigger” dynamical correlation discussed above. Whenever these new coordinations constitute schemes that we evaluate as proto-conceptual, such as schemes leading to proportional reasoning, it is very tempting to state that the children are re-inventing mathematical concepts within our designed fields of promoted action. That is, we seem to be witnessing the process of reflective abstraction, and this process is mediated by the children’s participation in the discovery and enactment of a cultural practice of our design, a sensorimotor practice they are never shown but are steered toward.

We are thus offering an explication of mathematical learning as a Piagetian constructivist process embedded in a Vygotskian cultural–historical framework. In so doing, we are also endeavoring to redress a lacuna in Piaget’s theoretical thesis, namely his little concern for sociocultural enframings of children’s logico-mathematical ontogenesis. As Turner (1973) writes:

Piaget’s model of psychogenesis is formulated in an artificial sociological vacuum; he has never confronted the question of the socio-cultural components of the mind at the level of the basic structure of the psychogenetic process itself. (p. 364) [Piaget] has, in other words, not yet come to grips with the problem of the specific social and cultural mechanisms through which cultures and societies participate in and control the genetic development of the individual psyches of their members. (p. 369)

Conclusions and implications

When Piaget began publishing on cognitive developmental psychology a whole century ago, clinical interviews were cutting-edge scientific method. By detecting systematic patterns in children’s action and utterance during

interviews, as they attempted to respond to his questions and solve his puzzles, and building on a colossal battery of cross-sectional studies, Piaget put forth a cognitive theory of genetic epistemology. Central to this theory was a painstaking explanation for individuals' subjective construction of psychological objects—new phenomenal categories that come forth to enhance, mediate, and regulate effective worldly transactions. These new categories and attendant sensorimotor schemes coalesce as the child's cognitive adaptations—emergent interaction routines enabled yet constrained by innate cognitive architecture. That is, the mind constructs a new category and, whenever doing so, extends and tightens its grip on the world.

Though much water has since flowed under Geneva's Mont-Blanc bridge, the Learning Sciences have not advanced much in evaluating Piaget's central claims respecting the child's construction of new psychological objects as solutions to problems of sensorimotor interaction. To be sure replication, qualification, and elaboration have been offered aplenty, and yet abstraction itself—the construct and process—have not been validated via independent measures.

These are early days in our quest to witness the psychological construction of new objects as it occurs. And yet our findings to date embolden and impel us to submit that we are literally seeing reflective abstraction. Empirical data from our task-based interviews, and in particular children's eye-gaze patterns triangulated against their tablet action-logs and audio–video recordings, are aligning remarkably well with Piaget's constructivist model of cognitive development. What more, by seeing what the child is looking at and manipulating we now understand far better our own successes and failures as educational designers in guiding the children to mathematize these tacit constructions.

More broadly, we have demonstrated alignment between a core construct from Piaget's theory of genetic epistemology—reflective abstraction—and tenets of Enactivism, dynamical-systems theory, ecological psychology, and socio-kinesiology. We thus join Allen and Bickhard (2013) in challenging and encouraging our colleagues to revisit Piaget's seminal contributions; to see for themselves the emergence of conceptual categories; to understand what this might all mean in practice; and make that practice a reality.

The ICLS 2016 conference theme “directs our gaze to the commitment of the Learning Sciences to provide a more insightful understanding of how people learn.” By directing our gaze to students' gaze, we hope to do just so.

References

- Abrahamson, D. (2012). Discovery reconceived: Product before process. *For the Learning of Mathematics*, 32(1), 8-15.
- Abrahamson, D. (2015). The monster in the machine, or why educational technology needs embodied design. In V. R. Lee (Ed.), *Learning technologies and the body* (pp. 21-38). New York: Routledge.
- Abrahamson, D., Lee, R. G., Negrete, A. G., & Gutiérrez, J. F. (2014). Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM Math. Education*, 46(1), 79-93.
- Abrahamson, D., & Sánchez-García, R. (in press). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences*. doi:10.1080/10508406.2016.1143370.
- Abrahamson, D., Shayan, S., Bakker, A., & Van der Schaaf, M. F. (in press). Eye-tracking Piaget: Capturing the emergence of attentional anchors in the coordination of proportional motor action. *Human Development*.
- Abrahamson, D., & Trninic, D. (2015). Bringing forth mathematical concepts: Signifying sensorimotor enactment in fields of promoted action. *ZDM Mathematics Education*, 47(2), 295–306.
- Allen, J. W. P., & Bickhard, M. H. (2013). Stepping off the pendulum: Why only an action-based approach can transcend the nativist–empiricist debate. *Cognitive Development*, 28(2), 96-133.
- Bamberger, J. (2013). *Discovering the musical mind: A view of creativity as learning*. NYC: Oxford U. Press.
- Bernstein, N. A. (1996). Dexterity and its development (M. L. Latash & M. T. Turvey, Eds.). Mahwah, NJ: LEA.
- Case, R., & Okamoto, Y. (Eds.). (1996). *The role of central conceptual structures in the development of children's thought* (Vol. 61[1-2], Serial No. 246). Chicago: University of Chicago Press.
- Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- Chow, J. Y., Davids, K., Button, C., Shuttleworth, R., Renshaw, I., & Araújo, D. (2007). The role of nonlinear pedagogy in physical education. *Review of Educational Research*, 77(3), 251-278.
- Cole, M., & Wertsch, J. V. (1996). Beyond the individual-social antinomy in discussions of Piaget and Vygotsky. *Human Development*, 39(5), 250-256.
- Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition*, 130(3), 335-347.
- Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 95-126). Dordrecht, The Netherlands: Kluwer Academic.
- diSessa, A. A., Levin, M., & Brown, N. J. S. (Eds.). (2015). *Knowledge and interaction*. New York: Routledge.

- Forman, G. (1988). Making intuitive knowledge explicit through future technology. In G. Forman & P. B. Pufall (Eds.), *Constructivism in the computer age* (pp. 83-101). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ginsburg, H. P. (1997). *Entering the child's mind*. New York: Cambridge University Press.
- Gray, E., & Tall, D. (1994). Duality, ambiguity, and flexibility: A "proceptual" view of simple arithmetic. *Journal for Research in Mathematics Education*, 25(2), 116-140.
- Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42, 335-346.
- Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge: MIT Press.
- Hutto, D. D., & Sánchez-García, R. (2015). Choking RECTified: Embodied expertise beyond Dreyfus. *Phenomenology and the Cognitive Sciences*, 14(2), 309-331.
- Kelso, J. A. S., & Engström, D. A. (2006). *The complementary nature*. Cambridge, MA: M.I.T. Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. NYC: CUP.
- Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment. *Ed. Researcher*, 42(8), 445-452.
- Marshall, P., Antle, A. N., van den Hoven, E., & Rogers, Y. (Eds.). (2013). The theory and practice of embodied interaction in HCI and interaction design [Special issue]. *ACM Transactions on Human-Computer Interaction*, 20(1).
- Martin, T., & Sherin, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences*, 22(4), 511-520.
- Moreno-Armella, L., Hegedus, S., & Kaput, J. (2008). From static to dynamic mathematics: Historical and representational perspectives. *Educational Studies in Mathematics*, 68(2), 99-111.
- Newell, K. M., & Ranganathan, R. (2010). Instructions as constraints in motor skill acquisition. In I. Renshaw, K. Davids, & G. J. P. Savelsbergh (Eds.), *Motor learning in practice* (pp. 17-32). Florence, KY: Routledge.
- Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone*. New York: Cambridge University Press.
- Norton, A. (2008). Josh's operational conjectures. *J. for Research in Mathematics Education*, 39(4), 401-430.
- Noss, R., & Hoyles, C. (1996). *Windows on mathematical meanings*. Dordrecht: Kluwer.
- Piaget, J. (1968). *Genetic epistemology* (E. Duckworth, Trans.). New York: Columbia University Press.
- Piaget, J. (1970). *Structuralism* (C. Maschler, Trans.). New York: Basic Books.
- Reinholz, D., Trninic, D., Howison, M., & Abrahamson, D. (2010). It's not easy being green: Embodied artifacts and the guided emergence of mathematical meaning. In P. Brosnan, D. Erchick, & L. Flevaris (Eds.), *Proc. of PME-NA 32* (Vol. VI, Ch. 18: "Technology," pp. 1488-1496). Columbus, OH: PME-NA.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. NYC: Oxford U. Press.
- Sarama, J., & Clements, D. H. (2009). "Concrete" computer manipulatives in mathematics education. *Child Development Perspectives*, 3(3), 145-150.
- Saxe, G. B. (2012). *Cultural development of mathematical ideas*. Cambridge, UK: Cambridge University Press.
- Schneider, B., Bumbacher, E., & Blikstein, P. (2015). Discovery versus direct instruction. In T. Koschmann, P. Häkkinen, & P. Tchounikine (Eds.), *Proc. of CSCL* (Vol. 1, pp. 364-371). Gothenburg, Sweden: ISLS.
- Sfard, A. (2010). *Thinking as communicating*. NYC: Cambridge University Press.
- Siegler, R. S. (2006). Microgenetic analyses of learning. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology* (6th ed., Vol. 2, Cognition, perception, and language, pp. 464-510). Hoboken, NJ: Wiley.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3(2), 115-163.
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Turner, T. (1973). Piaget's Structuralism (review article). *American Anthropologist*, 75(2), 351-373.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind*. Cambridge, MA: M.I.T. Press.
- Wertsch, J. V. (1979). From social interaction to higher psychological processes. *Human Dev.*, 22(1), 1-22.
- Worsley, M., & Blikstein, P. (2014). Using multimodal learning analytics to study learning mechanisms. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *EDM* (pp. 431-432). London: Inst. of Ed.

Acknowledgements

This work was funded by an Education and Learning grant, Utrecht University, January 2014.

Teacher Epistemic Learning in the Innovation Diffusion

Longkai Wu and David Hung

longkai.wu@nie.edu.sg, david.hung@nie.edu.sg

National Institute of Education, Nanyang Technological University, Singapore

Abstract: In this paper, a teacher epistemic learning model for implementation is proposed with the aim to prepare seeded teachers in reflection-for-action and to build their capacities in designing and enacting the curriculum on their own virtue. Fundamentally, epistemic learning is proposed because changing classroom practices is more than a surface or shallow change phenomena, but requires teachers to fundamentally shift in perspective or way of seeing how learning is to be practiced. We describe this epistemic learning model and explore its effectiveness. Various types of data were collected, including surveys, interviews, field notes, and teachers' lesson design artifacts. It is found that the seeded teachers from the five schools, regardless of their own teaching profiles and school contexts, have obtained high degrees of buy-in of the innovation and developed their readiness towards implementation for the future enactments. It is envisioned that such an epistemic learning model will inform the design for teacher professional development, in the pursuit of innovation diffusion, serving a wider community of the teaching practice.

Introduction

Teachers play an important role in the process of diffusion of any curricular innovation (Urhahne, Schanze, Bell, Mansfield, & Holmes, 2010). How teachers perceive the innovation at hand and build their capacity to implement the particular curriculum will determine the success of the innovation to a large extent. At the same time, the characteristics of teachers, together with the characteristics of the innovation and features of environmental context, account for the outcome of diffusion of an innovation (Rogers, 2003). In fact, teachers benefited more when they participated in professional learning in a collaborative form (Cloonan, 2009; Kopcha, 2012) and their school contexts and needs had to be considered (Stein, Smith, & Silver, 1999). In many of the literature, teachers' learning community and network building were highlighted (Scribner, Cockrell, Cockrell, & Valentine, 1999; Sun, Penuel, Frank, Gallagher, & Youngs, 2013). In the context of diffusion of innovation, Dearing, Greene, Stewart, & Williams (2011) proposed the idea of an "implementation registry" in the domain of healthcare, which is an online resource for practitioners within or across different institutions for sharing knowledge about solutions to challenges during dissemination, diffusion and implementation of an innovation. It is more than a normal registry in that it links whether an innovation works with why and how it works. It "fosters innovation and implementation success" by building informal, virtual communities of practice (vCoPs) across institutions. How-to knowledge was identified and captured in vCoPs for practitioners from different working sites to obtain understanding about the innovation so as to decide whether to take up the innovation to address their own challenges and what adaption should be done for their own context.

In our endeavor to achieve the diffusion of a curricular innovation from one seeding school to five more seeded schools, we have investigated the learning of a group of seeded teachers through an across-schools collaborative seeding journey, so as to develop their buy-in and readiness for implementation in their own respective schools. Roger (2003) described this process as an innovation-decision process, which involves five steps: (1) knowledge, (2) persuasion, (3) decision, (4) implementation, and (5) confirmation. The adoption of an innovation is determined by the interplay of factors from multiple levels of the school system such as macro-level ones like the national educational policies and socio-cultural factors of the school's learning ecology, meso-level ones like school-researcher partnerships, and micro-level ones like classroom-based work and interactions (Looi, 2011). In these different steps of the innovation-decision process, teachers would have different experiences and evolve their own understanding of the innovation.

In this paper, we first introduce the background of the innovation diffusion, and then propose a model of teacher's epistemic learning where different types of activities are designed for seeding teachers' comprehensive understanding of the innovation. Different parties, including the school administrators, the early adopter teachers (EATs), school administrators and the seeded teachers (STs) from the five schools, and researchers, were engaged in the learning journey. By examining the learning process and teachers' perception of the innovation, we want to explore the effectiveness of such a model, and hence contribute to the literature about the mechanism of teacher learning in the innovation-decision process, especially in the phases before and after decision of trying out the innovation is made.

Model of teacher epistemic learning

Building on the lessons learned from prior technology-based educational improvement research, we clearly recognize the importance of empowering teachers and building capacity to effect deeper changes in teachers' beliefs, knowledge, and practices (Fishman, 2005). Hence, we propose the model of teacher's epistemic learning in the stages of diffusion of innovation, as shown in Figure 1.

There are several design principles we put forward for the professional development sessions: 1) Sharing of the EATs to the STs from other schools should be participatory; 2) STs will also have a chance to have a lived in or embodied experience of what it means to teach such lessons; 3) STs have the flexibility to personalize the curricular innovation considering the local needs of their own schools; 4) EATs also benefit from a reflective practitioner stance of re-looking and adapting their own innovation and innovation approaches through their active participation and sharing with STs; 5) Communities of practice for EATs, STs and non-seeded teachers to share experiences, challenges, tips and constraints of how to enact a classroom innovation (with researchers as meso-level catalysts but to eventually fade away).

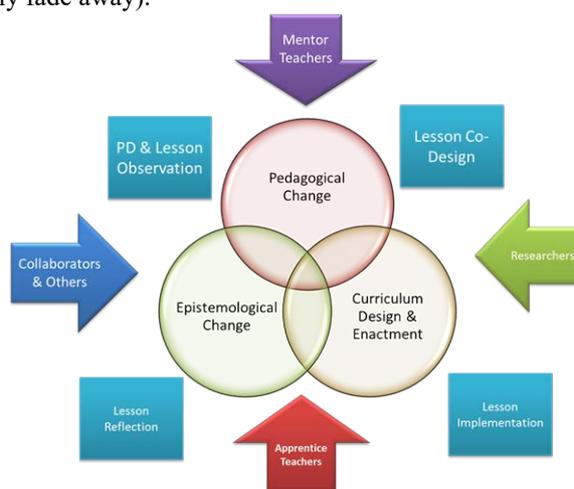


Figure 1. Model of Teacher's Epistemic Learning in Innovation Diffusion.

Based on the design principles, a model is proposed as shown in Figure 1. It consists of four types of activities, and involves different levels of learning agents. 1) In the first activity which we call "infusion", or the kick-off meeting, different parties in the scaling-up project, including all the teachers and school administrators from the seeding schools, as well as the MOE officers, gather for initial understanding of the innovation about "what it is" and "does it work" from perspectives of both researchers and pioneer practitioners. In the meeting, the effectiveness of MSL on students learning was presented to teachers, especially students' improvement on semester assessments in answering multiple choice questions (MCQ), open-ended (OE) questions and the total scores. The objective of the project was made clear to all the parties, as well as the responsibility and key performance indicators. 2) In the "lesson observation" activity, the STs have opportunities to have embodied experiences of the real classroom. The school administrators from each school also observed a few lessons to have a sense of what their own students might be experiencing. 3) The lesson co-design forms a teacher professional learning community where the EAT and STs exchanging ideas on lesson designs for the innovative curriculum, as well as other issues regarding innovation diffusion.

Context and participants

The innovation to be spread in this study is mobile seamless learning (MSL in short) which refers to the synergistic integration of the learning experiences across various dimensions such as formal and informal learning contexts, individual and social learning, and physical world and cyberspace (Chan, Roschelle, Hsi, Kinshuk, Sharples, Brown, et al. 2006).

The five schools were identified by the cluster superintendent based on a few criteria. Some of these criteria are whether there is some basic level of commitment by the school leaders towards using ICT in teaching and learning, how ready are they to embark on such an ICT project, and the support and buy-in by the principals especially whether they will stay in their school long enough to see through the project. The choice of schools is also influenced by the desire to spread the opportunity across a diversity of schools in the cluster. The teachers involved in this project were chosen by the principals.

Twelve STs from the five seeding schools were involved. Nine out of twelve teachers have relatively low level experiences in teaching (equal to or less than six years). Most of them thought that they currently taught in a teacher-centred way, which is about focusing more on knowledge delivery and students passively receiving knowledge, whereas four teachers thought they taught in a somewhat student-centered way by incorporating inquiry-based activities in their teaching.

Findings

Pedagogical change

From the lesson observations, the STs were impressed with the learning culture in the classroom: students were doing inquiries, they were not afraid to ask questions, tended to find evidence to support their scientific claims, explained well to the teacher and their peers about what they were thinking, and collaborated somewhat orderly with good division of labor. As Wilson expressed in the interview, he was very impressed that the students behaved naturally like scientists, which could not be trained and achieved in the traditional way of teaching:

The way I see Jane's student answers (the questions that Jane asked in the classroom) right, it's very encouraging because that is how a scientist, a researcher, a person who is into doing science (answers). That's how you ask questions and how you answer questions. And that's how she does with her class. She expects them to be (a) mini-scientist.

STs also saw that students were also very skillful at using the mobile devices for learning, whether individually or collaboratively, as Kabir said:

I think the kids (are) very comfortable using the devices, and they are able to do the collaboration. I think we all saw it in Socratic and all that (activities), (that) they are able to collaborate, answer Jane's questions and then move on from there. So I think that's very important value I see for the 21st century, communications, confidence, and all these things are embedded to do this (MSL).

Thus, from the lesson observation, the STs were able to see that the benefits of MSL not only included the exam performance, but also promoting a cultural change for learning in which students were doing self-directed learning.

The STs had opportunities to see how the teaching in MSL would be different from their current practices. The lesson observation provided STs a chance to see lessons from a different perspective. As Jane reflected from her point of view, when teachers teach in the class most of the time they are actually "blinded by the things that they have to accomplish", but when as an observer, they get "brighter eyes on what is happening in the classroom". As Jane said, some STs did reflect the things they might not be so conscious about when teaching, such as the questioning. In the lesson observation, Rohana paid special attention to the questions asked by Jane in classroom and summarized her questioning styles:

I think Jane plays a very important role, in the types of questions she asked her pupils. When I stepped into her lessons, I can see that her questions are scaffolds. Er, she started from very simple questions, and she is very dynamic. She will respond accordingly to students' responses. So if pupils are able to show higher-order thinking, she will streamline the questions to ask more complex kind of questions to trigger their learning

When Jane reflected about the elements that led to students' formation of the spirit of inquiry, she talked about the importance of the questions and students' perception change in answering the questions. In Jane's class, she made it a point to let the students know that they could not find the so-called right answers in the textbook, and there were usually more than one answer to a question. She stressed that science is about interpretations and finding evidence to support one's interpretations or predictions. Jane held this belief and she walked the talk in her teaching. Consequently, the STs observed that from the lesson observations, Wilson alluded to one difference in his teaching practices comparing his teaching and Jane's:

In our class it's very much (about) what they (the students) observe, (what I do in my classroom is) that I get them to observe and to find out and then give them a right or wrong answer. You know. Whereas in Jane's class you see the children really go in depth into each and every animal. Each and every particular group of animals, they go in depth into looking at it. When

talking about fur, (the students in Jane's class really did research to see) what is fur and what is hair you know. (They) Talk about giving births to young alive, what is the opposite or how other ways do animals give birth or reproduce and further on and further on. That is something that I seldom do in my class.

As we can see from Wilson's example, he began to reflect the difference between his teaching and Jane's, which affects the depth of knowledge gained by students. In Jane's MSL lessons, students were encouraged to research in depth and they were supported in doing so. In this type of learning, teachers played an even more important role by asking appropriate questions to guide students' inquiry than by directly telling them the answer.

Besides learning from Jane about the questioning techniques, STs also reflected that they learnt about the skills of providing guidance to students in this type of classroom teaching, such as how to guide students to get useful information from the vast information online, and "how to gear students towards the position of a scientist" (from Amber).

The impact of the lesson observation on the STs was also reflected by STs' follow-up practices. Rohana already had more than ten years' teaching experiences, but she tried to change her teaching practices after she observed Jane's lessons. Thereafter when she taught back in her own school, she started by asking questions in a different way, and she gave one example of her change:

For (the topic of) digestion, usually for P3 science, for this kinds of pupils I have not actively using questioning techniques like (those) employed by Jane, usually I will ask what are the different parts of the digestive system. Now I ask something like how could the digestion in stomach help in supplying energy, if I remember correctly. After that I bring in something like how the circulatory system, and the respiratory system also play a part in digestion for example. So more questions drawing linkages or inter-topical linkages, because the theme is about systems, so I try to make as much connections with other different types of system so that it make more sense, and it's more meaningful for them to learn other types of system.

Thus Rohana became more self-conscious about what questions to ask and began to try it out for her students, such as questions that drawing linkages between different topics to make learning more meaningful. She admitted before the lesson observation she was inclined more towards lecture-type teaching so that her students were not so responsive when she changed to ask more questions, but after a few more weeks, students were used to it and became more active in answering and responding.

Curriculum design and enactment

Teachers co-designed the theme of "Diversity" in the nine sessions. For each chapter within the theme, teachers went through the process as described in Figure 2. As not all schools has the same scope and sequence of content, the teachers first standardized the scope, sequence and learning objectives for every topic. The standardization helps the community to implement the curricular in a similar pace so that they can have more meaningful sharing and reflection in the future. The teachers, then, decided the total period for each chapter and discussed students' common learning difficulties and misconceptions from their teaching experiences. One chapter often surrounds one focusing topic (i.e. living and non-living things, plants, animals, fungi and bacteria, materials) and the 5E model (Engage-Explore-Explain-Elaborate-Evaluate) was used by the community to design a learning cycle for one chapter. Some of the schools have already used 5E as lesson design framework, so adopting the 5E model was not difficult for them. The teachers all contributed ideas and the resources they have used for the activity, and appropriated those activities to fit in different stages. After going through all the five stages within 5E, the teachers volunteered to take up one or more lessons to detail down the lesson plan. Jane provided a template for teachers to elaborate the lessons, which comprised three columns: class activity, complementing home activities, and MLE (mobile learning environment) activities. In "class activity", the teachers described teacher's and students' activities respectively, while in "complementing home activities" teachers designed activities that students could do out of classroom with aids of the mobile devices before or after the lesson, and in "MLE activities" teachers specified the application they planned to use, the purpose and the objectives of using it. The last two focuses encouraged the teachers to integrate the characteristics of mobile learning, which is leveraging the mobile devices for students' learning and linking formal and informal learning. Besides, Jane also suggested teachers to consider about differentiated instruction while designing to cater to all students with different competency and needs.



Figure 2. Lesson Co-Design Process Experienced by the Teachers' Community.

In the co-design sessions, teachers not only discussed teaching, but also their understanding of the concepts. Some STs commented that through discussion in the lesson co-design, they gained clarification of certain science concepts, and hence improved their content knowledge. The diversity of the school context also provided teachers with more ideas to integrate the innovation and improve their lesson design. Anna mentioned that the community of the five school teachers was different from their own school teachers learning community, and the knowledge gained regarding designing learning journeys was valuable to her:

In our school, we know what we are doing, but we don't know about other schools. Let's say for teaching the same subject, (we don't know) how they (other schools) extend it, (and) where are the learning journeys that they bring their pupils to. For example, when we talk about animals, we mentioned we went for farm-hopping to the various farms, and I think some of other schools say 'oh, okay, we go to the mushroom farms', and there are some schools saying instead of farm hopping they went to zoo. So these are the things we learn from one another. And they even share, you know, when they teach certain topics what the major misconceptions are. What are the things that pupils are always not familiar with? Because every school's cohort is different, we get this type of knowledge, which is very valuable.

In the community, Jane was the only person that had years of experiences collaborating with researchers designing and implementing the mobilized curriculum, hence she is the best person to share the design experiences with the STs because of the "practitioners' sympathy", that is, as teachers, they share similar considerations and concerns. In the lesson observation, the STs observed her exemplar practices, but in the co-design session she also shared with the STs about her own challenges and failures and the observation of her mentees in the leading school. She always cheered the STs up and hoped them to hold a positive attitude toward technology integration, and at the same time advised the STs to be patient since change is difficult and takes time. The co-design benefited not only the STs, but also Jane. She reflected that when designed the lessons with the five school teachers, she applied a more "macro-view", which was different from the "micro-view" way in her own school. In her own school's lesson design, it is "activity per se" in which teachers discuss in detail about every activity, whereas with the STs, it is "learning objective per se" in that she only defined the main learning objectives for specific topics and sequences of topics to maintain the possibility for sharing in next year. She described the way that the STs and she co-designed the lessons for Material:

For example when we were talking about (the topic) material, then the teachers said 'why not we have activities like telling a story about Cinderella. So (it's) about the shoes of Cinderella... we can talk about material use to make that glass slippers and why Cinderella have that glass slippers.. Then you look at technology (about) how I can actually tell the story? Do I have to bring a physical book? If the teachers don't tell the story then someone said 'how about looking for one online virtual story or from Youtube and get the students to watch at home?' You know like a flipped classroom then that part will come in at the later stage... I think generally that the teachers are quite ok in suggesting this kind of activity. So they actually give even other suggestions like using other technology which they think can better help the students. Yeah, so

I thought for the seamless part of the activity, we are really looking more on how technology can support that instructional objectives... so generally it's quite automatic in them. I don't really need to tell them that "ay, you need to have this," and "ay, can it be done without technology?"

It can be seen that Jane took different approaches to preparing the teachers, which was leaving decision of activity details, resources and application to the STs so as to shift the ownership of curriculum design to them step by step. Jane also reflected that the dynamics between her and the STs were different from her and her school teachers. When communicating with the STs, she avoided to telling them what to do but suggesting them to try out something since the context was very different.

Epistemological change

STs described their perception of the core elements of the innovation. We count the frequencies of the key words shown in STs' responses, and the following words were used to describe the innovation with frequencies shown within the round brackets: inherent or intensive use of technology (8), student-centered and teacher as facilitator (4), self-directed (4), beyond classroom, or in and out of classroom (4), life-long learning (1), 21st century skills (1), and enhance students' interests in science learning (1). Thus when teachers were first introduced to the MSL, they seemed to be more impressed by the inherent use of technology in learning, which was absent in their own schools. They also saw the learning happening not only in the classroom, but also enabled by the mobile devices to happen beyond the classroom. Teachers saw the self-directedness in MSL because it encouraged students to ask questions and to initiate their own learning.

In the interview, teachers' ideas were further clarified. Most of the STs especially acknowledged the "seamless" element in the package and viewed it as a linkage between formal learning and informal learning. Wilson stated in the interview that the unique part of MSL was that the mobile devices served as a means to make learning a really 24/7 thing:

I think (the unique part of MSL) is that the students who are embarking on this programme have a means to an end. They have the means to do (inquiry), (and) they have been given a means to explore, research and to be able to do their research easily, how to say, validated, by their teachers, (and) by their peers. Using the mobile device, and like what the programme's name suggest, it is really seamless because they don't just do it in school. They do it at home, (and) they do it on the way home. You know, they can do it anywhere they wish to. So that's where I see the difference (between MSL and learning in my school), because right now here in school (the situation) is I (only) have 3 periods to teach. And after that they have other subjects to do and after they go home I also don't know what (they do at home). I mean I give them homework, but whether or not they revise and do, that's at home, (and) I am unable to access. But with MSL, because they have their mobile device, (so) whatever that they uploaded from home I also know. I mean I can tell that they are doing something at home.

Winston appreciated one uniqueness of the MSL in that the teacher could evaluate and monitor students' learning progress even they might be doing it at home. Other teachers also mentioned the value of MSL lies in students' easy access to vast information online. With the mobile devices, students can search information on the site. But that was not the case in other school. As Joanna mentioned, she once provided assignments asking students to search for information when at home, but some of her students were forbidden to use computer by their parents during week days.

STs also mentioned "self-directed learning" a lot, which is advocated by MOE as one of the desired student outcomes in 21st competencies. Anna gave an example of what she envisioned for her students, and elaborated her understanding of self-directed learning:

I mean you see it's like, we can give them a topic, and off we go, whether at home, along the road, even when they are in canteen with their friends, they may discover certain things, and then there they post. We can have the discussion forum. They may even notice something during holidays, even post and we have discussion. So that's what we mean by self-directed learning. It's no longer always teachers asking you must do this you must learn this, maybe the child can even post pictures of a creature that looks like an insect but doesn't have the full characteristics of the insect, but we can all discuss this. And teachers (perform) as facilitators. Of course trying to guide them to the right direction if they are too off-track, and maybe at the same time to facilitate the quality of discussion.

So STs see the potential of MSL as a means for students to become self-directed learners. They can spot problems, ask questions and initiate their research, which changes learning from passive receiving to constructing knowledge. Teachers recognize their role as facilitators, which might be quite a shift for them since most of them have been teaching in a teacher-centred way.

Despite of the affordances provided by the technology, teachers recognized that they key factor that leads to the success of the innovation is the teaching of the teacher, as expressed by Kabir:

It is how you use it to teach, I think that's the key factor. It's not just using technology for its own sake, it's that how we use it in a way that students are engaged and learn further, and learning is enhanced. So the way how teachers use it to enhance the learning is most important. Of course we have other things, but this is the most important one.

Discussion and conclusion

In this paper, we situated our study in a diffusion of MSL from one school to five more schools and proposed a model of teacher's epistemic learning to get their buy-in and prepare them for the future implementation. In the findings, we articulated how each activity in the model helped the STs to evolve their gradual understanding of MSL and what knowledge they have gained through interactions with teachers and school leaders from the leading school, and researchers. It was found that the learning experiences presented and hence convinced teachers of the advantages of the curricular innovation to teachers and students, as well as feasibility of implementation in their own school, which led to the STs high degree of buy-in of the innovation.

Research showed that teachers' perceptions of the five attributes of an innovation were critical for their adoption decision making and implementation. Through the embodied learning experiences, teachers were able to see many facets of the innovation and built their own understanding of the characteristics of the innovation. (Dearing, 2009) suggested an "exemplary demonstration" in a convincing manner to influence adoption decisions and thus increase the likelihood of diffusion, and in our study the infusion and lesson observation served the purpose of demonstration and enabled teachers to *observe the relative advantages* of MSL, specifically, students' significant improvement in answering open-ended questions and their engagement, enthusiasm, and scientist-like mind of thinking in the classroom. The curriculum design activity rendered them a sense of ownership of the innovation and let them see the *compatibility*; the understanding of the innovation highlight more on the pedagogy rather than the use of applications, as well as assurance of systemic supports got from Q & A sessions convinced them of the *simplicity* of the innovation. The curriculum package, which was a collective product of the community, make the innovation more *triable* at the first stages of implementation. The three types of knowledge described by Roger were also provided for the teachers through those learning experiences, such as how-to knowledge in the lesson observation in terms of how to ask scaffolding questions, how to manage a MSL classroom, and how to guide students to think and talk like scientists. It is also provided through the lesson co-design on how to design the MSL curriculum, and on how to integrate the package into individual school's existing package.

Compared with other teacher professional development programmes, our model not only emphasizes teacher's professional learning, but also provides infrastructure support through creating opportunities of communications between school leaders and teachers within and across schools. Teachers need to deal with multiple issues when implementing, but our learning model had endeavored to establish the systemic supports (from school leaders to the IT technicians and teaching assistants etc) for teachers to alleviate them from administrative matters and to enable them to focus on improving curriculum and instruction. Being different from other teacher PDs in the form of innovators/researchers-to-practitioners interaction, our model highlights the interactions between practitioners to practitioners-to-be. Teachers share similar considerations and concerns when adopting an innovation, so the advice and tips from peers would be more pragmatic and targeted. The learning within the community of practices benefited teachers in the preparation as we showed in the lesson co-design sessions, and will impact the future implementation and even dissemination within each individual school. The model we proposed here not only applies to the diffusion of educational innovation as in our case, but also to the diffusion of innovation in other domains. Thus, to prepare for adopting a potential innovation, the learning cycle should incorporate the following core elements: *effectiveness demonstration, embodied and epistemic learning journey, shift of ownership, learning community building, and systemic support provided.*

References

Dearing, J. W. (2009). Applying Diffusion of Innovation Theory to Intervention Development. *Research on Social Work Practice, 19*(5), 503-518.

- Fishman, B. (2005). Adapting innovations to particular contexts of use: A collaborative framework. In C. Dede, J. Honan & L. Peters (Eds.), *Scaling up success: Lessons learned from technology-based educational innovation* (pp. 48-66). New York: Jossey-Bass.
- Freeman, M. (2012). To adopt or not to adopt innovation: A case study of team-based learning. *International Journal of Management Education (Oxford Brookes University)*, 10(3), 155-168. doi: 10.1016/j.ijme.2012.06.002
- Fullan, M. (2001). *The new meaning of educational change*. London: RoutledgeFalmer.
- Jephcote, M., & Davies, B. (2004). Recontextualizing Discourse: exploring the meso-level. *Journal of Education Policy*, 19(5), 547-564.
- Jones, B. L., & Bjelland, D. (2004). International experiential learning in agriculture. *Proceedings of the 20th Annual Conference, Association for International Agricultural and Extension Education*, 963-964. <http://www.iaee.org/2004/Carousels/jones-carousel-NEW.pdf>
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review*, 85(5), 363.
- Kopcha, T. J. (2012). Teachers' perceptions of the barriers to technology integration and practices with technology under situated professional development. *Computers & Education*, 59(4), 1109-1121.
- Lavonen, J., Juuti, K., Aksela, M., & Meisalo, V. (2006). A professional development project for improving the use of information and communication technologies in science teaching. *Technology, Pedagogy and Education*, 15(2), 159-174.
- Looi, C.-K. (2011). *Sustained Innovation in Classroom Practices: The Role of Educational Researchers as Meso-Level Actors*. Paper presented at the Beijing Forum 2011, Beijing, China.
- Looi, C.-K., Zhang, B., Chen, W., Seow, P., Chia, G., Norrist, C., & Soloway, E. (2011). 1:1 mobile inquiry learning experience for primary science students: a study of learning effectiveness. *Journal of Computer Assisted Learning*, 27, 269-287.
- Makinster, J. G., Barab, S. A., Harwood, W., & Anderson, H. A. (2006). The Effect of Social Context on the Reflective Practice of Preservice Science Teachers: Incorporating a Web-Supported Community of Teachers. *Journal of Technology and Teacher Education*, 14(3), 543-579.
- Margolis, J., & Doring, A. (2012). The Fundamental Dilemma of Teacher Leader-Facilitated Professional Development: Do as I (Kind of) Say, Not as I (Sort of) Do. *Educational Administration Quarterly*, 48(5), 859-882. doi: 10.1177/0013161x12452563
- McNeill, K. L., & Krajcik, J. (2008). Scientific Explanations: Characterizing and Evaluating the Effects of Teachers' Instructional Practices on Student Learning. *Journal of Research in Science Teaching*, 45(1), 53-78.
- Ministry of Education, S. (2008). MOE Launches Third Masterplan for ICT in Education. <http://www.moe.gov.sg/media/press/2008/08/moe-launches-third-masterplan.php>
- Montecinos, C., Pino, M., Campos-Martinez, J., Domínguez, R., & Carreño, C. (2013). Master teachers as professional developers: Managing conflicting versions of professionalism. *Educational Management Administration & Leadership*. doi: 10.1177/1741143213502191
- Mustafa, J., & Al-Mothana, G. (2013). Using the Diffusion of Innovation Theory to Explain the Degree of English Teachers' Adoption of Interactive Whiteboards in the Modern Systems School in Jordan: A Case Study. *Contemporary Educational Technology*, 4(2), 138-149.
- Reiser, B., Tabak, I., Sandoval, W., Smith, B., Steinmuller, F., & Leone, A. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263-305). Mahwah, NJ: Lawrence Erlbaum.
- Rogers, E. M. (2003). *Diffusion of innovations* (4th ed.). New York: Free Press.
- Scribner, J. P., Cockrell, K. S., Cockrell, D. H., & Valentine, J. W. (1999). Creating professional communities in schools through organizational learning: An evaluation of a school improvement process. *Educational Administration Quarterly*, 35(1), 130-160.
- Stein, M. K., Smith, M. S., & Silver, E. A. (1999). The development of professional developers: learning to assist teachers in new setting in new ways. *Harvard Educational Review*, 69(3), 237-269.
- Sun, M., Penuel, W. R., Frank, K. A., Gallagher, H. A., & Youngs, P. (2013). Shaping Professional Development to Promote the Diffusion of Instructional Expertise Among Teachers. *Educational Evaluation and Policy Analysis*, 35(3), 344-369. doi: 10.3102/0162373713482763
- Urhahne, D., Schanze, S., Bell, T., Mansfield, A., & Holmes, J. (2010). Role of the Teacher in Computer-supported Collaborative Inquiry Learning. *International Journal of Science Education*, 32(2), 221-243. doi: 10.1080/09500690802516967

Disruptions to Practice: Understanding Suspensions of Youths' Interest-related Activities

Katie Van Horne, University of Colorado Boulder, katie.vanhorne@colorado.edu
Erica Van Steenis, University of Colorado Boulder, erica.vansteenienis@colorado.edu
William R. Penuel, University of Colorado Boulder, william.penuel@colorado.edu
Daniela DiGiacomo, University of Colorado Boulder, daniela.digiacomom@colorado.edu

Abstract. An emerging line of research in the learning sciences focuses on generating a better understanding of how youth develop and sustain in interest-related pursuits. In this paper, we focus not on what sustains engagement in interest-related activities but on what disrupts it. Disruptions have largely been a neglected object of theorizing, even though they are acknowledged in models of interest development, such as the four phase model (Hidi & Renninger, 2006). We build upon Azevedo's (2011, 2013) lines of practice theory by drawing on conceptions of social practice from German-Danish critical psychology to expand and complicate the notion of "conditions of practice" for the youth in our study. We examined youth in our interview study who reported experiencing disruptions to their interest-related pursuits. We found that some youth suspended their participation due to loss of access to necessary resources, and other youth suspended pursuits due to competing educational commitments and future desires.

Keywords: interests, learning across settings, social practice, conditions of practice, youth pathways, desired futures

Introduction

An emerging line of research in the learning sciences focuses on generating a more robust understanding of how youth develop and sustain interest-related pursuits (e.g., Azevedo, 2006; 2011; Barron, 2006; Greeno, 2006). An interest-related pursuit as we define it here is one that youth enjoy doing, believe they are getting better at or learning from over time, and seek out when they have the opportunity. A focus on such pursuits is important because engagement in interest-related pursuits can catalyze learning, open up future learning opportunities and possible careers (Ito et al., 2013; Barron, 2006).

In this paper, we focus not on what sustains engagement in interest-related activities, but on what disrupts it. Disruptions have largely been a neglected object of theorizing, even though they are acknowledged in models of interest development, such as the four phase model (Hidi & Renninger, 2006). Interest does not develop in either continuous or linear pathways (Azevedo, 2011; 2013). Sometimes pathways are interrupted or change course. Here, we attend closely to and theorize changes to conditions and opportunities that disrupt or lead youth to suspend their participation in interest-related activities.

We conjecture that a key source of disruption is a change to the conditions supporting participation related to material resources and time. Many of youths' interest-related pursuits are supported by settings and resources outside of school. These opportunities often require additional financial resources and supports. Young people who are dependent on in-school opportunities or resources from public institutions have differential supports and access to out-of-school learning opportunities for engagement in interest-related pursuits than youth who have ready access to these opportunities (Ito et al., 2013). For these youth, it can be difficult to turn interest-related pursuits into desired future careers or into desired resources. For youth for whom limited financial resources are less of a concern, interest-related pursuits and commitments out-of-school also compete with personal, local, and societal goals related to education and career pathways. Youth have to attend to the competing goals of school activities and future educational commitments and desires. For these youth, time is a scarce resource, and changing availability of time for pursuits can lead to youth suspending those pursuits. Because interest-related "lines of practice" (Azevedo, 2011) are often mangled by shifts in opportunity and competing engagements and because of the increasing recognition of the importance of interest for long-term learning outcomes, more attention is needed to understanding the disruptions in interest-related to pursuits.

Theoretical framework

We draw in part on Azevedo's (2011; 2013) "lines of practice" theory of individual preferences and conditions of practice that encourage persistence in interest-related participation over time. Azevedo (2013) argues that there

are four features of interest-related practices that encourage persistence in lines of practice over time by affording participants the ability to tailor their participation. The features of conditions include: “an extensive and varied material infrastructure,” opportunity to pursue the interest in multiple communities or sites, ways in which youth can participate in interest-related activities of various short and long durations, and space for collaboration and sharing of ideas (p. 1). We characterize young people’s lines of practice as influenced by both their preferences and the conditions of practice in which youth participate across contexts and over time. Within one hobby or interest-related activity an individual might participate in multiple lines of practice that change throughout the various contexts of their lives thus looking different than it did before.

However, not all youth have equitable opportunities for participation within the conditions of their interest-related activities and some youth experience disruptions and competition within their lines of practice. We build upon Azevedo’s lines of practice theory by drawing on conceptions of social practice from German-Danish critical psychology (Dreier, 1997; Mørck & Huniche, 2006; Nissen, 2005) to expand and complicate the notion of “conditions of practice” for the youth in our study. As we analyze the ways in which youth distribute their interest-related pursuits across settings, we argue for the need to pay attention to the contexts in of youths’ lives in relative to one another. In any given social context, youths’ “personal engagements and stakes in context depend on its status in relation to other contexts in their trajectories of participation” (p. 42). Youth have differential access to contexts and may lack access to positions within those contexts which can lead to inequitable scopes of possibility (Dreier, 2009).

We argue differential success pursuing interest-related pursuits is dependent on youths’ access to and roles in and across the many contexts of their lives. This is in contrast to dominant discourses that name interest development as a personal psychological attribute that operates in or out of formal learning environments (see Bathgate, Schunn, & Correnti, 2014). Using personal attributions (e.g., motivation, ability, grit) as justification of success or failure do not address the ways in which youth must organize their lives across the contexts in which they are participating. Holzkamp (2015) argues, “such personalizing attributions close off further questioning as to the possible disruptions and contradictions in the way students organize their lives that arise from shortages of specific resources but also from many other more or less unknown circumstances” (p. 66). We utilize this framework to investigate youths’ interest-related practices and the nuances of those more or less unknown circumstances. Our analysis seeks to understand how those conditions of practice affect their participation now and their ideas about desired futures.

Methods

In order to explore the conditions of practice that lead to youth suspension of interest-related pursuits, we asked two research questions:

1. Why do youth suspend interest-related pursuits?
2. What are the effects of suspension on youths’ future desires and career goals?

These questions informed how we methodologically carried out our study and investigated our data. Our larger project attended to youths’ engagement with interest-related pursuits and how those interests changed over time and across contexts. In an initial survey, youth answered several questions related to their experience of an interest-related pursuit, that is, a pursuit that youth said they enjoyed doing, believed they are getting better at or learning from over time, and sought out whenever they had the opportunity. For this study, we relied on an analysis of interviews with youth to learn about the nature of why youth suspended or experienced disruptions to their participation with those interest related pursuits. Using interviews allowed us to pay particular attention to how youth described the barriers and obstacles they faced when participating in interest driven activities. We elicited and attended to ways that youths’ engagement and persistence with activities is informed and influenced by their specific life circumstances (Azevedo, 2011, 2013; Mørck & Huniche, 2006). These in-depth discussions with youth via our interviews were particularly relevant to us developing an understanding of why and under what conditions youth suspended participation in activities.

Participants

Youth aged 13-17 years old representing 19 different youth programs across the United States and Canada participated in this study. In the initial waves of data collection, 266 youth completed surveys about their interest-related pursuits in both 2013 and 2014. Fifty-four youth from this group agreed to participate in an interview in wave three. Each youth participated in an interest-related activity such as video gaming, digital journalism, music production, writing, and illustrating.

We selected nine students to highlight in this study through a multistage process aimed at identifying representative cases of young people who reported experiencing disruptions in their interest related pursuits. First,

we created a representative data display for all 54 youth in the larger study; the data display included youth's demographic information as well as characteristics we identified as important to understanding youth's articulations of disruptions and activity suspension. These included the youth's articulations of their history of participation in their chosen activity, the type of disruption they experienced, their future goals, and the outcome of the experienced disruption. From this table we identified youth who we experienced a disruption that led to a suspension of their interest related activity. The data display allowed us to identify two broad categories of disruptions that the youth experienced. These included a loss of material or financial resources or the prioritization of formal schooling. From the list of youth who suspended their interest related pursuits because of these types of disruptions, we selected nine representative case studies. There are five females and four males, and collectively, they represent a number of ethnicities, including African American, Mexican American, Puerto Rican, Asian, and white. In this paper we present these case studies as representative of themes we found emerging from our larger corpus of data.

Initial survey results

We report on results from the initial surveys to provide context on youths' suspension of activities for this interview study. In Wave 1, we asked each young person to answer the question, "What is something you do that you enjoy and get better at, the more you do it?" Youth filled in an interest-related activity. In Wave 2, we re-presented that same activity to the youth and we asked them if they still engaged in the activity or not. A majority of youth with responses, 180 (71%), reported that they were still doing an activity while a minority, 74 (29%) reported changing their activity. Twelve youth did not respond to that question. All 74 youth provided some written explanation of why they stopped participating in their Wave 1 activity. By far the most common reason youth cited for stopping the activity was a conflict in their schedule or lack of time ($n = 26$). The next most frequent reason was a change in priorities or level of interest ($n=20$). We seek to provide depth to these explanations of suspension in this interview study.

Sources of data

We analyzed the data over a 7-month period which included constructing, administering, and analyzing interviews paying close attention to the ways youth talked about their interests and the influence of the greater world around them. Our interview protocol was informed by Dreier (2008), who argues that while people are usually studied within one context, researchers need to attend to how persons' participation unfolds in time and across contexts as they socially act in multiple settings. The protocol was constructed with the goal of collecting articulations of youth's experiences with interest related pursuits. Questions elicited youth's descriptions of their activities and purposes for participation, their current involvement in their activity, the networks (e.g. linkages and supports) they drew upon when participating in their activity, obstacles they experienced, and how they perceived the future as related to their participation. The interviews also elicited youths' perspectives on how their participation changed over time. This change over time, as youth described it, was influenced by disruptions and subsequent frustrations related to the various constraints youth faced when pursuing their chosen interest related activity.

Procedure

Our team interviewed the 54 youth who agreed to participate throughout summer 2015. The interviews were conducted over the phone and lasted between 45 minutes and one hour. Interviews were recorded, transcribed by an outside transcription service, and uploaded and analyzed in the software program Dedoose. During six code creation summits members of the research team participated in reading interviews and organizing codes into thematic patterns we saw across the data. Once these thematic patterns were established, we developed code constructs including parent and sub codes. Following the completion of the coding scheme, the coding of the corpus of data was divided amongst the research team. We established inter-rater reliability for all codes, with resulting Cohen's kappa ranging from 0.63 to 1.00. In this analysis, we used the codes designed to capture youths' experiences during their initial engagement with the activity and subsequent engagements with the activity. These "temporal participation" codes had inter-rater reliability measured with Cohen's kappa ranging from 0.73 to 0.84.

Approach to data analysis

Because we were interested in how young people experience disruptions in their interest related pursuits that lead them to suspending participation, we conducted analysis through a multistage process. First we coded the 54 youth interviews for the youth's relationship to their activity, their length of participation in the interest related pursuit, their duration of participation, how their participation changed and was affected over time, and how they imagined their possible future. Compiling these coded transcripts and coded data, we developed youth profile's that aimed

to trace the youth's relationship to their interest related pursuit over time. This level of analysis allowed us to see the ways in which their participation in the activity changed and was affected over time. We then created data displays for each student of initial activity, history of involvement in activity, youth articulations of disruptions, type of disruption, future goals, outcomes, and the program setting. These data displays allowed us to identify themes regarding (1) each youth's relationship to their interest related pursuit, (2) how this relationship changed over time, and (3) how their participation in the activity was disrupted by varying life circumstances and conditions. After discussing and finding agreement regarding youth's experience of disruptions and their subsequent suspension of activity, we then looked for types of disruptions occurring across the data. We then generated claims regarding these data and interrogated for other explanations.

Findings

Despite noting connections between their interest-related pursuits and their desired futures, some youth experienced disruptions in engagement or chose to suspend interest-related activities. Table 1 displays the types of disruptions youth in this study (n=54) reported experiencing. The top two disruptions youth mentioned were counted and are displayed in the table, for a total of 73 disruptions.

Table 1: Frequency of Disruptions

Type of Disruption	Frequency Count (Percent of Total)
School	24 (33%)
Material/Financial	13 (18%)
Geographical	1 (1%)
Social Support	8 (11%)
Career Feasibility	4 (6%)
Change of interest/new hobby	6 (8%)
Time	5 (7%)
No Disruption Mentioned	12 (16%)

Table 2 displays the case study youth (n = 9) who suspended their interest-related pursuits (column 2) along with frequency of participation at the time of the interview and the reported duration of engagement in the pursuit. The table also characterizes the types of conditions of practice that youth reported experiencing. These nine youth represent the top two types of disruptions experienced by the youth in our study: disruptions because of financial/materials conditions or disruptions because of educational commitments. Material or financial conditions denote instances in which youth could not participate because the material infrastructure was not available to them or financially, they did not have opportunity to participate. Educational conditions denote those in which youth reported having competing responsibilities related to formal schooling that led to limited or suspended engagement in the interest-related pursuit. Next, we examine each these conditions in depth.

Table 2: Youth and reasons for suspension

Youth Demographics	Interest-related Pursuit	Frequency of Participation; Duration	Reasons for Suspension		Desired Future
			Material or Financial	Educational	

<i>Caila, female, age 18, African American</i>	Film Editing, Film Making	Dabbling; 3 years	X		Film Maker
<i>Cedric, male, age 17, African American</i>	Music Production	Daily; 9 years	X		Rap Artist and Music Producer
<i>Gabe, male, age 18, Mexican American</i>	MC or Rapping	About two times per week; 5 years	X		MC or Rapper or TV Producer
<i>Ike, age 18, male</i>	Illustrating	Daily, 11 years	X		Graphic novel illustrator
<i>Emilia, age 19, female, African American</i>	Volunteering	Weekly; 1 year	X		Dentist
<i>Talia, age 17, female, American Indian, African American, Puerto Rican</i>	Writing scripts for media	A couple of times per year; 2 years		X	Creative Writing; still exploring options
<i>Claudia, age 17, female, Asian</i>	Playing <i>Starcraft</i>	Occasionally; 10 years		X	Physician or Medical field
<i>Angie, age 18, female, white</i>	Drawing	Not often, as a hobby; 7 Years		X	Tattoo Artist or Psychology-related career
<i>Danny, age 19, male</i>	Clubs at school	Variable; 4 years		X	Pediatrician or person working with special needs children

Loss of access to necessary material and financial resources

Many interest-related pursuits require material or financial resources to engage in them. For instance, youth may need access to a car to drive where public transportation does not go or access to pursuit-specific training or materials, such as video editing software. Youth in our study reported both types of barriers to participation in their interest-related pursuits. Youth who experienced financial and material barriers to participation named specific circumstances of their lives that led them to suspend their participation. Emilia, for example, while in high school enjoyed volunteering but was unable to pursue it more seriously because she lacked access to transportation. She emphasized this point when she explained that once in she was in college, she volunteered more frequently because her sister attended the same college and had a car that she offered to ease Emilia's issues with transportation to the homeless shelter.

Suspension of interest-related activities also occurred for youth in our study when life conditions stemming from family income required youth to take on extra responsibilities to support their parents or younger siblings. Gabe reported suspending his rapping activities whenever he needed to take care of his little brother and sister. He told us, "sometimes my parents work a lot so sometimes I need to help. I feel like chores too [get in the way]. I'm the oldest one, in a way I have to put in more work." Gabe suspended his participation in his interest-related activity because he prioritized providing support for his family at home.

Related, youth reported instances where they suspended their activities because they no longer had access to the material resources necessary for that specific pursuit. Often these pursuits -- related to new media arts or digital arts -- required computers and other technical equipment. Cedric, an aspiring rapper and music producer from Los Angeles, made beats and wrote and recorded raps to go with them. He started making beats in elementary school when his school allowed him to bring home a computer to work on. A desire to record rap music drove Cedric and he lacked funds to buy beats made by other producers so he made the beats himself. Sometimes he

sold or licensed his produced beats to other rap artists, but when we talked to him, he was working on his own album for the first time. He stored all his beats on his computer but then he lent his computer to his little brother who broke the machine. They were not able to recover the files. He lost all the beats for his album and suspended his practice. Eventually he was able to negotiate with his mother to borrow her computer to start anew on his album. Cedric's suspension of activity was relatively short because he was able to share material resources within his family to restart the work.

When we interviewed youth, we asked them what barriers they expected to encounter as they pursued their future goals. Some youth who wanted to pursue their interest-related activities in the future imagined they would face financial barriers to participation. For example, Caila, described her future goal of becoming a filmmaker:

The money aspect of it is a lot. Like you do have to buy a lot of your own supplies in order to set up that portfolio that do. You do have to buy a lot high technology and have large support from family, probably most likely your Mamma who's going to be buying the software for you. Then there's no guarantee that you might be a professional in the field or you may actually be good at what you do.

Calia was actively looking for universities where she could apply to a reputable film making program but she was not spending time film making at the time because she was no longer eligible to participate in her afterschool program where she primarily worked on films due to her age. Even though Caila did not stop pursuing a career in film production, available resources are a worry for her as she prepares for her future, signaling their salience as a key condition for pursuing an interest toward a career. Other youth described how expectations for future earnings led other people to expect them to suspend interest-related pursuits because "you need money to live". Ike, a graphic illustrator, reported that people in his life were not supportive of his time spent on illustrating because they expected him to make money, asking him "why work after something when you don't really get paid for it?" While this did not cause Ike to suspend his illustration work, he expected to encounter this tension in the future as he pursued his goal of publishing graphic novels.

Prioritizing formal school

In addition to the material and financial barriers and obstacles experienced by some youth, we found that other youth in our study suspended participation in an activity because of education-related responsibilities. School commitments like taking AP and honors classes, homework responsibilities, and college applications manifested in a lack of time for additional pursuits and often drove youth to suspend their interest-related activities. We understand this suspension of activity as being connected to the necessary attention youth felt they had to pay to school as a means to work toward their imagined future to attend college and become successful in context of their communities and greater world.

Talia, a youth who participated in a New Media Arts program that taught her to write scripts for new productions and supported her skill development, described homework as getting in the way of her ability to attend the program and write scripts. She told us: "I had to choose either homework or [going to the program]." When we pushed further, Talia explained that it was AP English specifically that took time away from writing scripts. For Claudia, a youth in our study who played the videogame *Starcraft* told us that exams got in the way of her game play, as well as "big projects that teachers give, and the homework that, most of the time, could get into the way." At one point *Starcraft* was a large part of Claudia's life but like Talia she suspended her participation because it took "time away that [she] could be more productive" and do her work that was integrally related to her desire to become a physician.

Angie, a youth interested in drawing, described similar sentiments to Talia and Will:

What's different is that I kind of dropped doing it because I'm more focused on school right now. I wish I had more time to be able to do it, so I can get better. I always thought I would get a career in art, but that dream has died off. I didn't really know it was worth it anymore, so I kind of pushed it back, and I guess that's the only difference.

In the case of Angie, we also see the way in which her focusing on school intersects with her commitment to drawing as a worthy future pursuit in terms of a career. When asked what sort of future careers illustrating prepared Angie for she described:

I honestly don't even know. I thought I was going to have a career with drawing; I just never went that way. Other interests and schooling got in the way. I never really branched out because I really don't know where I could go with it and what I could have from drawing other than a nice hobby to have.

Angie's lack of seeing a pathway to pursue illustrating as a future career in addition to her school commitments are significant. This example demonstrates the ways in which doing well in school or attending to schooling is driven by seeing a clear pathway into a future a career.

Discussion: Conditions of practice and desired possible futures

The conditions of practice that youth experienced – lack of access to material or financial resources or conditions characterized by commitments to schooling – varied among the youth in our study and affected their interest-related pursuit trajectories in different ways. For the first five youth in Table 1, interest-related pursuits were integral to their desired futures. For example, both Gabe and Cedric spent time working on rapping and music production and told us about the work they were doing to pursue music production as a career after college. The second set of youth shown on the bottom half of the table (n=4) described desired futures and careers different from or loosely related to their desired possible futures. Talia, for example, pursued script writing for digital media production in her afterschool program and she was still exploring options for her future. She was considering something related to creative writing -- a transformation of her previous line of practice. In general, youth who experienced suspension of interest-related pursuits due to loss of access to material or financial resources tended to pursue lines of practice toward possible futures directly related to their interest-related pursuit. In contrast, youth who experienced suspension of pursuits for educational careers saw their interest-related pursuits as more of a hobby than a possible future.

Contributions and implications

Interest-related pursuits are ever-changing and re-locating as youth navigate the circumstances of their lives. Azevedo (2013) describes this navigation as something people must organize in order to pursue an interest, “being interested in a practice requires weaving it with many other concerns, domains, values, goals, and practices in one’s life space, which makes the practice of interest meaningful in the short and long hauls” (p.44). We are concerned with the circumstances that are out of youths’ control and where this weaving or organization disrupts interest-related pursuits. We find in this study some circumstances that disrupt youths’ pursuits that are related to the pursuit of alternate, socially valued goals such as advancing one’s schooling. But in other circumstances, restricted access may be limiting possible future opportunities.

What does those mean in the long haul for youth who are working toward better socioeconomic outcomes and desirable resources? We need to elevate in a theory of interest-related pursuits the conditions of practice that enable and constrain those pursuits and that disrupt them. We argue German-Danish critical psychology lends this perspective to describe the key conditions of practice for youth persistence or suspension of interest-related pursuits and how these conditions lead to shifting engagement in youths’ pursuits of desired futures and resources. This lens allowed us to theorize the connection between youths’ interests and their future careers. People who design for and support youth in their interest-related pursuits need to attend to the key conditions that lead to disruptions in youths’ practices as well as recognize and mitigate the ways in which life circumstances affect participation across contexts of youths’ lives.

References

- Azevedo, F. S. (2006). Personal excursions: Investigating the dynamics of student engagement. *International Journal of Computers for Mathematical Learning*, 11, 57–98.
- Azevedo, F. S. (2011). Lines of practice: A practice-centered theory of interest relationships. *Cognition & Instruction*, 29, 147–184.
- Azevedo, F. S. (2013). The tailored practice of hobbies and its implication for the design of interest-driven learning environments. *Journal of the Learning Sciences*, 22(3), 462-510.
- Bathgate, M., Schunn, C. D., & Correnti, R. J. (2014). Children’s motivation towards science across contexts, manner-of-interaction, and topic. *Science Education*, 98(2), 189-215.
- Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development*, 49, 193–224.

- Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 289–325). Mahwah, NJ: Erlbaum.
- Dreier, O. (1997). Subjectivity and social practice. Aarhus, Denmark: Center for Health, Humanity, and Culture.
- Dreier, O. (2008). *Psychotherapy in everyday life*. New York: Cambridge University Press.
- Dreier, O. (2009). Persons in Structures of Social Practice. *Theory & Psychology*, 19(2), 193–212.
- Greeno, J. G. (2006). Learning in activity. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 79–96). New York, NY: Cambridge University Press.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 4(2), 111–127.
- Holzkamp, K. (2015). Conduct of everyday life as a basic concept of critical psychology. In E. Schrauble & C. Højholt (Eds.), *Psychology and the conduct of everyday life* (66-99). London: Routledge.
- Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., Schor, J., Sefton- Green, J., & Watkins, S.C. (2013). *Connected Learning: An Agenda for Research and Design*. Irvine, CA: Digital Media and Learning Research Hub.
- Mørck, L.L. & Huniche, L. (2006). Critical psychology in a Danish context. *Annual Review of Critical Psychology*, 5.
- Nissen, M. (2005). The subjectivity of participation: Sketch of a theory. *International Journal of Critical Psychology*, 15, 151-179.

Toward an Integrated Framework of Scientific Reasoning

Shiyu Liu, The Pennsylvania State University, liux0631@gmail.com

Abstract: This paper proposes an integrated framework, *PREP*, to capture the dynamic relationship among key influential factors of scientific reasoning, including prior knowledge (P), reading capacity (R), epistemic beliefs (E), and personal attributes (P). Empirical evidence from a qualitative study is presented as we elaborate and further develop this framework. In this study, twenty-six undergraduate students were interviewed to reason about competing arguments on global climate change. In-depth analysis of participants' interview responses demonstrated important characteristics of the cognitive processes and underpinning mechanisms of scientific reasoning. This work adds to ongoing discussion about the nature of scientific reasoning and provides a holistic approach to characterizing and evaluating scientific reasoning. Furthermore, by capturing important features of individuals' reasoning when faced with climate issues, this work has significant implications for classroom science teaching and learning.

Keywords: scientific reasoning, epistemic beliefs, prior knowledge, reading capacity, personal attributes

Introduction

Scientific reasoning is at the core of scientific inquiry, but is not just a vehicle that scientists are entitled to (Giere, Bickle, & Mauldin, 2006; Sinatra & Chinn, 2011). The ability to reason scientifically greatly affects how people understand and evaluate information from professional publications and public media (Sinatra, Kienhues, & Hofer, 2014). Hence, science educators have long been emphasizing the importance of teaching scientific reasoning in the classrooms so that students grow to be scientifically informed citizens (National Research Council, 2001, 2012; NGSS Lead States, 2013). However, obstacles abound in such efforts, one of which lies in the lack of theoretical consensus on what constitutes and affects scientific reasoning (Zeineddin & Abd-El-Khalick, 2010). To better inform teachers as they support students' development of reasoning skills, there is a call for more integrated frameworks to conceptualize the essence of scientific reasoning (Kind, 2013).

In this paper, we propose a framework to characterize the processes and mechanisms of scientific reasoning. Based on our work through two large projects that aimed to promote classroom instruction and assessment of scientific reasoning, we identified four key factors and investigate their dynamic interactions during reasoning processes. In the following, we first review relevant literature that informed our perspective. We then present our theoretical framework with findings from an empirical study and conclude with future directions of this work.

Theoretical backgrounds

Existing literature provides numerous perspectives toward what constitutes scientific reasoning and how to measure it. Here, we highlight previous research that has influenced our framework.

Reasoning is a process of drawing conclusions from principles and evidence so as to infer new conclusions based on what is already known (Wason & Johnson-Laird, 1972). As an important type of formal reasoning, scientific reasoning has drawn attention from scholars in different fields that hold distinctive perspectives (see Zimmerman (2005) for review), among which the main discrepancy falls in whether to focus on the cognitive processes during reasoning or the underpinning mechanisms that guide such processes. Chinn and Malhotra (2002) pointed out that more work is needed to unpack the complex mechanisms which underlie individuals' reasoning processes. They held that as the majority of previous studies tended to focus on exploring simplified aspects of scientific reasoning, such as the control-of-variable strategy, the epistemic nature of reasoning has been missing in not only scholarly research but also classroom teaching. Here, the epistemic nature mainly refers to how one evaluates evidence in authentic inquiry and coordinates theory and evidence when faced with complex science problems.

In the past decades, with growing concerns on teaching socioscientific issues—such as global climate change and genetic engineering—in science classrooms, the scope of research on scientific reasoning has been greatly extended (e.g., Mason & Scirica, 2006; Yang, Chang, & Hsu, 2008). For instance, Sadler and colleagues used the concept of *socioscientific reasoning* to explain how individuals process information about socioscientific issues to which there are no absolute solutions (Sadler, Barab, & Scott, 2007). In particular, they stressed that socioscientific reasoning, as a major type of informal reasoning, requires one to recognize the inherent complexity

of ill-defined issues, examine them from multiple perspectives, appreciate that they are subject to ongoing inquiry, and exhibit them from multiple perspectives. Yet, while studies of socioscientific reasoning have greatly enriched our understanding of how people reason, it remains unclear whether individuals are able to reason formally through scientific means about competing arguments on socioscientific issues.

More recently, the growing body of studies on epistemic beliefs has imposed significant influence on investigation of scientific reasoning (Sinatra & Lombardi, 2013). Current research on epistemic beliefs has moved from mainly focusing on individuals' beliefs toward knowledge and the process of knowing (e.g., Hofer & Pintrich, 1997) to investigating how such beliefs guide individuals' cognitive and metacognitive processes (e.g., Muis & Franco, 2009). As a result, many scholars start to incorporate discussion on epistemic beliefs into studies of reasoning processes (e.g., Barzilai & Eshet-Alkalai, 2015; Chinn, Rinehart, & Buckland, 2014). Such research movement corresponds to educational reforms which suggest further development of students' skills of engaging in authentic inquiry and participating in scientific argumentation (NGSS Lead States, 2013). In all, the converging research interests from various lines of work have brought in a refreshing look on what affects scientific reasoning and calls for a revisit of existing approaches to investigating scientific reasoning. Further work is needed to move toward an integrated framework of scientific reasoning.

The PREP Framework

Building upon related literature, we identify four interconnected factors that critically contribute to individuals' scientific reasoning processes: *prior knowledge*, *reading capacity*, *epistemic beliefs*, and *personal attributes* (henceforth, **PREP**) (Liu, 2014).

Prior knowledge

Prior knowledge is a double-edged sword in one's scientific reasoning process. On one hand, appropriate knowledge supports the selection of appropriate reasoning strategies (Schauble, 1996). The more content knowledge one has, the more likely they are to perform well in scientific reasoning about related issues (Mason & Scirica, 2006; Osborne, 2010). On the other hand, the prominent domain-specific approach to investigating scientific reasoning left it vague as to whether prior knowledge can place significant impacts on reasoning (e.g., Klahr & Dunbar, 1988). Thus, to maximize the positive effects of prior knowledge, more and more studies start to investigate the dynamic interactions between prior knowledge and other learner characteristics such as reading capacity and epistemic beliefs (e.g., Kendeou & van den Broek, 2007).

Reading capacity

Compared to prior knowledge, the effect of reading capacity on scientific reasoning is much less discussed. As research on scientific reasoning has been more and more embedded in contexts of argumentation, reasoning often involves comprehension and evaluation of written texts that include two-sided scientific arguments or multiple solutions in relation to controversial or ambiguous issues (Chan, Ho, & Ku, 2011). Hence, to reason about various arguments, it is critical that individuals have adequate reading capacity to coordinate the multiple sources of information presented. Bråten and colleagues held that individuals who could integrate information are more likely to process information in a way consistent with their epistemic beliefs (Bråten, Britt, Strømsø, & Rouet, 2011). At the same time, they also called for further research to specify the relationship between reading capacity and epistemic beliefs.

Epistemic beliefs

Epistemic beliefs entail "*individuals' beliefs about the nature of knowledge and the processes of knowing*" (Hofer & Pintrich, 1997, p. 117). People with limited reasoning abilities tend to find it difficult to process information from multiple perspectives, which in turn discourages them from endorsing sophisticated epistemic beliefs that acknowledge the tentative and complex nature of knowledge (Zeidler, Walker, Ackett, & Simmons, 2002). Conversely, naïve epistemic beliefs were found to be associated with withdrawal from in-depth reasoning (Sinatra, Southerland, McConaughy, & Demastes, 2003). Those who hold such beliefs tend to overlook the need to engage in scientific reasoning and are more likely to treat information that does not support their existing beliefs in a biased manner (Chan, Ho, & Ku, 2011; Kuhn, 2001; Weinstock & Cronin, 2003).

Discussion on the interaction of the above three factors adds an important lens to the study of scientific reasoning. It not only relates to constant debates on the domain-general and domain-specific nature of scientific reasoning, but also proposes alternative explanations of challenges one may encounter during reasoning.

The fourth factor: Personal attributes

With increasing emphasis on epistemic beliefs and other learner characteristics, research on scientific reasoning is demonstrating growing orientation toward capturing “warm cognition” (Dole & Sinatra, 1998). In particular, affective factors like motivations and beliefs are being paid increasing attention to in the investigation of reasoning processes. For instance, Lombardi, Seyranian and Sinatra (2014) have explored how emotions may relate to the level of scientific understanding students hold and their perception of plausibility of evidence regarding critical issues such as climate change. As early research on scientific reasoning mostly features the cognitive and metacognitive processes, the role affective factors play is a relatively under-studied area. Based on existing work on affective components of scientific learning, here we use “personal attributes” to refer to important aspects including one’s belief system, attitudinal orientation toward certain topics, as well as personal interests and motivation.

The study

While existing research has taken various paths to reveal the relationships among some of the four factors above, to our knowledge, there is little work that devotes to capture in depth how they interact with each other. Therefore, in the present study, we employed a qualitative approach, *grounded theory* (Glaser & Strauss, 1967; Strauss & Corbin, 1990) to answer a critical question: *How the four factors proposed in the PREP Framework interact with each other to guide people’s scientific reasoning processes?* Grounded theory research allows an exploratory development of theory grounded in data from the field (Glaser, 1978). Researchers conduct iterative data collection and analysis to allow an analytic, substantive theory to emerge from the phenomenon. Thus, following the principles of grounded theory, this work enables a more “naturalistic” insight into the nature of scientific reasoning.

Participants

Twenty-six undergraduate students (20 females and 6 males, $M_{\text{age}} = 19.65$ years, $SD_{\text{age}} = 1.06$) at a major university in the Midwestern U. S. were recruited through theoretical sampling (Glaser & Strauss, 1967). Theoretical sampling is a key feature of grounded theory: it requires the sampling process to be guided by the need for understanding the phenomenon of interest (Glaser, 1978). Following this principle, participants were purposefully recruited from a wide range of majors to increase the diversity of the sample in terms of their content knowledge and personal beliefs about climate change. Data collection was complete after 26 participants were interviewed when saturation of theoretical categories was reached and no new categories emerged.

Study design

An interview protocol was developed, including a 606-word reading document and 13 open-ended questions. The topic of interest is global climate change. Despite the ongoing educational efforts, students at all levels still experience difficulties in processing the multitude of perspectives and evidence regarding climate change (e.g., Braasch et al., 2013; Gil, Bråten, Vidal-Abarca, & Strømsø, 2010). Analyzing individuals’ reasoning processes on this topic can help reveal a more well-rounded view of the nature of reasoning and its influential factors. The reading document consisted of two opposing perspectives: climate change is human caused versus climate change is due to natural changes. This reading document involved three most commonly discussed topics regarding climate change: Earth’s temperature change, rising sea level, and extreme weather events. The three topics were presented on separate pages with arguments from both sides. The interview questions were designed to assess three aspects: 1) prior knowledge about climate change, 2) evaluation of evidence and arguments, and 3) perspectives toward climate science. As they read the documents, participants were asked to read aloud and think aloud to help us probe into their thinking processes. They then evaluated the evidence used in the reading, critiqued on the arguments, and discussed their viewpoints about climate science.

Data analysis

The interviews were audio-recorded and later transcribed verbatim. Transcripts were entered into NVivo 10 for further analysis. Constant comparative analysis (Glaser & Strauss, 1967) was employed for data analysis and it involved three stages of coding (*open coding*, *selective coding* and *axial coding*) to allow theoretical categories to emerge. Data analysis was conducted in two layers. The first layer featured identification of the four influential factors during reasoning. The second layer analyzed participants’ reasoning processes based on the complexity of their cognitive processes when thinking aloud and responding to the interview questions. The combination of the two layers thus generated fine-grained analysis of the dynamics among the proposed four factors.

To establish the trustworthiness of data analysis, techniques such as memoing and diagramming were employed throughout the coding processes. In addition, during data analysis, categories that were similar in their

definitions were further compared and contrasted to decide their final categorization. These efforts all helped to achieve the goal of establishing credibility, transferability, dependability and confirmability in the data analysis (Lincoln & Guba, 1985).

Results

Table 1 lists the outline of our coding system for this data set. We also identified sub-factors under reading capacity, epistemic beliefs, and personal attributes. In particular, reading capacity was analyzed based on participants' ability to use reading strategies such as monitoring their reading progress, epistemic beliefs captured participants' views about the nature of knowledge and expertise in climate science, whereas personal attributes involved their attitudes, as well as related belief systems, toward climate issues. Broadly, three patterns were identified to categorize reasoning processes—*withdrawal from reasoning*, *limited reasoning*, and *complex reasoning*—where the relationships of the proposed four factors were further investigated. Given the limited space, here we present three brief examples to demonstrate key findings at each reasoning pattern. Quotes from participants are specified with their assigned participant number (such as P1, P2, and so on).

Table 1: Outline of the core coding system

Factors	Sub-Factors	Definitions
Prior Knowledge		Existing understandings participants hold about climate change
Reading Capacity	Monitoring Comprehension	Timely reflection of reading progress and problems encountered
	Making Associations	Relating texts to existing beliefs or knowledge
	Making Inferences	Generating explanations or comments based on texts
	Others	Responses such as “I don’t like the wording here”
Epistemic Beliefs	Nature of Science	Perception of the nature of scientific knowledge in general, such as whether knowledge is fixed or fluid
	Relativity	Relative correctness of multiple perspectives
	Perception about Credibility	The trustworthiness of the source of evidence, the reliability of the source, the scientific foundations of measurement, etc.
	Scientist’s Expertise	Scientists’ certainty about the causes and consequences of climate change, consensus in the science community on the anthropogenic nature of climate change, etc.
	Flexibility of Beliefs	Reflection on one’s own and others’ stance and likelihood to change such stance
	Complexity and Uncertainty of Climate Science	Perception of the nature of climate science
Personal Attributes	Personal Beliefs	The role of political and religious beliefs
	Attitudinal Orientation	Preferences for and support of certain stance and arguments
	Personal Interests	Personal relevance of the topic of interest, curiosity toward the topic, etc.
	Motivation	Reasons to/not to actively seek relevant information on the topic of interest

Withdrawal from reasoning

This pattern consisted of responses from participants that were mainly repeating the information presented. Sometimes they may request specification or clarification after reading information that they were not familiar with or uncertain about. While they might recognize and briefly identify the consistency or conflict of information in the reading with what they already knew, they would not resort to their epistemic beliefs when evaluating evidence and making conclusions. For example, after reading about the composition of the atmosphere and the percentage of CO₂ in it, P14 responded “OK, I guess I didn’t know that CO₂ only constitutes less than 1% of the trace gases.” While she acknowledged that “scientists might realize ten years later that they’re missing some information”, P14 often simply conformed to or refuted a given argument without further reasoning about the evidence provided.

Limited reasoning

Reasoning processes identified into this pattern involved more cognitive efforts and included more details about how participants processed any given information. However, this reasoning pattern was limited as participants mostly emphasized the surface features of evidence as well as the arguments it supported. Their overwhelming focus on writing features, such as tone of writing, sometimes may override their attitudes and beliefs about the topic when they attempted to evaluate evidence and make conclusions. For instance, after reading the sentence “Changes in the frequency and intensity of extreme weather events are due to human-caused Earth’s temperature increase”, P22 commented that “even though I believe this, this sentence came off a little biased. Even though I do believe that it is human caused, it came off a little strong, so I don’t know (whether I agree with it).”

Complex reasoning

This reasoning pattern represented participants’ advanced skills in drawing upon their cognitive and affective resources. P25, for example, demonstrated her abilities to reason about statistical meanings of numerical values. When presented with numbers such as the percentage of CO₂ in the Earth’s atmosphere, P25 raised concerns that these numerical values alone might not be telling the whole story and requested clarification for the scientific meanings of these numbers. P25 explained her concerns about the need for more information regarding the statistical significance of evidence as follows:

Like if the increase of CO₂ went from like 0.5 to 1%, I mean I don’t know how significant that would be. Um, they’re also showing that the increase in the Earth’s average surface temperature, but it looks like a very small amount too. I don’t know how big an impact that would have.

Furthermore, P25 also actively related information to her religious beliefs and reflected on how her beliefs may have affected her thinking process. When evaluating the evidence that was used to support “Rising sea level is not human-caused”, P25 stated that “I would have tried to find sources that were like based on the Creation (God’s creation of the world) rather than the Theory of Evolution”. Similarly, after reading “In particular, the number of hurricanes and tropical storms during 1995 and 2004 doubled that during 1970 and 1994 in North Atlantic”, P25 reported the following as she was thinking aloud:

The Bible warns in the book of revelation when Christ returns his coming, that there will be more like catastrophic events like that, like as his return years. So, like, as a Christian, as a follower of Jesus, it’s kind of exciting, because you will wait for him to come back a second time, so I mean because God’s words warn about that, that makes me think of (this argument) just from that perspective.

The PREP Framework: A step further

The two-layered analysis in this study provided rich information on the complex dynamics among the four factors proposed. Based on the current findings, the nature of the relationships among these factors makes it difficult to come to a conclusive model. Nonetheless, some main themes emerged in this work presented opportunities to further develop the PREP Framework and inform future efforts in capturing scientific reasoning.

First, the potential role each factor may play is relatively consistent and predictable. Figure 1 presents a simplified process model based on the current findings to help illustrate the functional characteristics of these factors during reasoning. In particular, when individuals are exposed to multiple arguments in writing, information encoding takes place and their capacity in reading comprehension acts as the first filter for what they may choose to focus on as they continue to further interpret the information. During their reasoning processes, intentionally not unintentionally, individuals relate to prior knowledge to construct mental models of the information. As they align such models with their epistemic beliefs about climate change, individuals choose to engage in reasoning or withdraw from it. Individuals’ personal attributes play a critical role in how they reason. Those who show lack of interests in issues related to climate change are more likely to avoid reasoning about the different arguments. At the same time, when individuals integrate their personal values such as religious beliefs into their perspectives about climate change, they approach reasoning much more differently. It should be noted that the relationship between personal beliefs and the chosen paths of reasoning can be reciprocal: while the former affects how individuals evaluate different aspects of the arguments, their focus on arguments that support their own perspectives can further strengthen their beliefs and even biases. This process model can serve as a preliminary application of the PREP Framework to help us understand the proposed factors and their interactions.

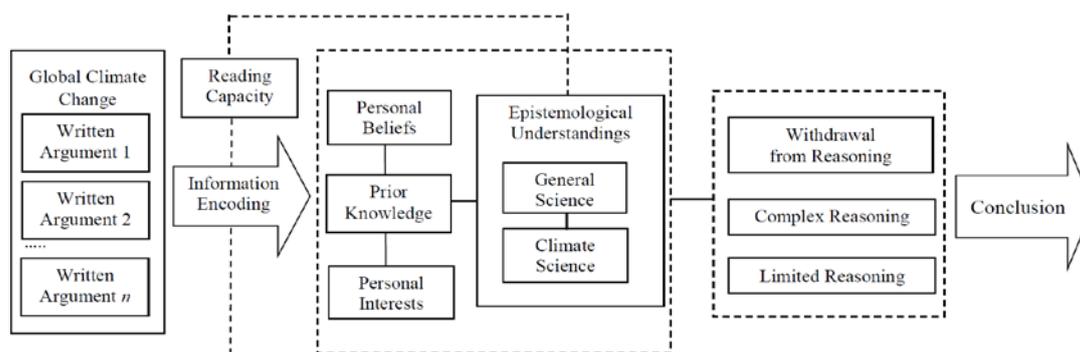


Figure 1. A process model of scientific reasoning on global climate change.

Second, interactions among the four factors fluctuate not only between individuals but also within individuals. For instance, based on the present results, whether more prior knowledge and advanced epistemic beliefs may result in more complex reasoning depends on the level of reading comprehension one achieves at the moment. Therefore, rather than pushing for a uniformed characterization of scientific reasoning, PREP leaves sufficient flexibility for future research to probe into the dynamics among these factors across contexts.

It should be noted that while findings from this qualitative work can serve as a foundation for further development of the proposed PREP Framework, cautions should be taken when employing this framework. First, this study adopted a qualitative research approach to investigate scientific reasoning processes. While integrating the grounded theory approach has yielded in-depth discussions on how undergraduate students reason about climate issues, like all qualitative studies, concerns may arise regarding the trustworthiness and credibility of this work. While great efforts were made throughout data collection and analysis to avoid biases and minimize preconceptions for grounded theory, given the nature of qualitative studies, it is open for further investigation whether the PREP framework can be generalized across subject domains for different populations. Moreover, the topic of interest here was global climate change, whereas scientific reasoning processes may differ as the topic varies. Follow-up studies will continue to explore how this theoretical framework may apply to reasoning with other socioscientific issues, such as genetic engineering and water pollution. To obtain a more comprehensive understanding about the contributing factors of scientific reasoning, later researchers may find it helpful to expand the scope of investigation and consider aspects that have not been discussed very much in scientific reasoning research such as social and religious factors.

Research on scientific reasoning has been of great interests to the learning sciences community. The PREP Framework proposed in this paper aimed to pursue a more integrated view of scientific reasoning and investigate its influential factors. Through capturing the dynamic interactions among four factors, including prior knowledge, reading capacity, epistemic beliefs, and personal attributes, this framework demonstrates a holistic view toward scientific reasoning and reveals the complexity of its underpinning mechanisms. PREP adds to the current efforts in facilitating the theoretical construct development for scientific reasoning. The uniqueness of the grounded theory methodology makes the exploration of individuals' reasoning processes more naturalistic and brought a new lens to research on scientific reasoning.

Educational implications

This work has important implications for classroom science teaching and learning. Despite the numerous studies in the field of scientific reasoning, no consensus is reached on what instructional support should be provided to facilitate student reasoning (Osborne, 2010). A central goal of science education is to enhance students' skills in effective communication of scientific issues. National science education standards in the U.S. have emphasized that students should be able to reason scientifically in order to engage in scientific argumentation and thus communicate about issues that impact their daily lives (NGSS Lead States, 2013). Osborne (2010) suggested that there are several aspects of scientific reasoning skills that science education might seek to develop such as identifying patterns in data and resolve uncertainty of scientific inquiry. However, there have not been many empirical studies that provide empirical support for this proposal. Rooted in empirical data, the PREP Framework is consistent with the proposal Osborne made, but extended its scope in the context of a socioscientific issue. By revealing participants' perspectives and thinking processes about climate issues, this study provided critical information for teachers to consider as they develop their curricula.

Furthermore, findings from this work add to the ongoing debates in climate change education about how to enhance climate literacy. One of the most critical educational objectives is for students to learn about how socioscientific issues are handled and evaluated within society so that they are able to act as responsible citizens in the future (e.g., Höttecke, Henke, & Riess, 2012). Educational and policy documents have suggested that students should develop reasoning skills that can help them evaluate the causes and effects of global climate change (NRC, 2001, 2012). Students should be more actively engaged in evidence-based reasoning about human impacts on the Earth's climate system to propose, test and modify possible solutions to current climate issues. Incorporating scientific reasoning into climate change education will help fulfill this goal. It is essential that the general public come to appreciate the relevance of scientific reasoning and its impact on climate literacy. However, although the importance of scientific reasoning in climate change education has been established, there have not been many detailed discussions on effective approaches to promote scientific reasoning and climate literacy. As this grounded theory study looked into the complex relationship between scientific reasoning and prior knowledge as well as identifying multiple factors that may have affected students reasoning about climate issues, it may serve to initiate conversations between scientists and educators for potential collaborations in their efforts of enhancing the public's climate literacy.

References

- Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36*, 86-103.
- Braasch, J. L., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology, 38*(3), 180-195.
- Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J. F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48-70.
- Chan, N., Ho, I., & Ku, K. (2011). Epistemic beliefs and critical thinking of Chinese students. *Learning and Individual Differences, 21*, 67-77.
- Chinn, C., & Malhotra, B. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86* (2), 175-218.
- Chinn, C., Rinehart, R., & Buckland, L. (2014). Epistemic cognition and evaluating information: Applying the AIR Model of epistemic cognition. In D.N. Rapp & J.L.G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences*. Cambridge, MA: MIT Press.
- Dole, J., & Sinatra, G. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist, 33*(2/3), 109-128.
- Giere, R., Bickle, J., Mauldin, R. (2006). *Understanding scientific reasoning*. Belmont, CA: Thomson Wadsworth.
- Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Summary versus argument tasks when working with multiple documents: Which is better for whom?. *Contemporary Educational Psychology, 35*(3), 157-173.
- Glaser, B.G., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Glaser, G. (1978). *Theoretical Sensitivity*. Mill Valley, CA: The Sociology Press.
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.
- Höttecke, D., Henke, A., & Riess, F. (2012). Implementing history and philosophy in science teaching: Strategies, methods, results and experiences from the European HIPST Project. *Science & Education, 21* (9), 1233-1261.
- Kendeou, P., & van den Broek, P. (2007). Interactions between prior knowledge and text structure during comprehension of scientific texts. *Memory & Cognition, 35*, 1567-1577.
- Kind, P. (2013). Establishing assessment scales using a novel disciplinary rationale for scientific reasoning. *Journal of Research in Science Teaching, 50* (5), 530-560.
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*, 1-48.
- Kuhn, D. (2001). How do people know? *Psychological Science, 12* (1), 1-8.
- Liu, S. (2014). How people reason: A grounded theory study of scientific reasoning about global climate change. *Unpublished doctoral dissertation*, University of Minnesota.
- Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: SAGE.
- Lombardi, D., Seyranian, V., & Sinatra, G. (2014). Source effects and plausibility judgments when reading about climate change. *Discourse Processes, 51*(1-2), 75-92.

- Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction, 16*, 492-509.
- Muis, K., & Franco, G. (2009). Epistemic beliefs: Setting the standards in self-regulated learning. *Contemporary Educational Psychology, 34*, 306-318.
- National Research Council. (2001). *Educating teachers of science, mathematics, and technology*. Washington, DC: National Academies Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Committee on a conceptual framework for new K-12 science education standards. Board on Science Education, Division of Behavioral and Social Sciences and Education, Washington, DC: The National Academies Press.
- NGSS Lead States (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.
- Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science, 328*(23), 463-466.
- Sadler, T., Barab, S., & Scott, B. (2007). What do students gain by engaging in socioscientific inquiry. *Research in Science Education, 37*, 371-391.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology, 32*(1), 102-119.
- Sinatra, G., & Chinn, C. (2011). Thinking and reasoning in science: Promoting epistemic conceptual change. In K. Harris, C. McCormick, G. Sinatra, & J. Sweller (Eds.), *Educational Psychology: Contributions to education* (Vol. 1, pp. 257-282). Washington, DC: American Psychological Association.
- Sinatra, G., Kienhues, D., & Hofer, B. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist, 49*(2), 123-138.
- Sinatra, G.M., Southerland, S.A., McConaughy, F., & Demastes, J.W. (2003). Intentions and beliefs in students' understanding and acceptance of biological evolution. *Journal of Research in Science Teaching, 40*(5), 510-528.
- Strauss, A.L., & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage.
- Wason, P., & Johnson-Laird, P. (1972). *Psychology of Reasoning: Structure and Content*. Harvard University Press, Cambridge, MA.
- Weinstock, M., & Cronin, M.A. (2003). The everyday production of knowledge: Individual differences in epistemological understanding and juror-reasoning skill. *Applied Cognitive Psychology, 17*, 161-181.
- Yang, F.Y., Chang, C.Y., & Hsu, Y.S. (2008). Teacher views about the constructivist instruction and personal epistemology-A national study in Taiwan. *Educational Studies, 34*, 527-542.
- Sadler, D., & Zeidler, D. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching, 42*(1), 112-138.
- Zeidler, D., Walker, M., Ackett, W., & Simmons, M. (2002). Tangled up in views: Beliefs in the nature of science and responses to socioscientific dilemmas. *Science Education, 86*(3), 343-367.
- Zeineddin, A., & Abd-El-Khalick, F. (2010). Scientific reasoning and epistemological commitments: Coordination of theory and evidence among college science standards. *Journal of Research in Science Teaching, 47*(9), 1064-1093.
- Zimmerman, C. (2005). The development of scientific reasoning skills: What psychologists contribute to an understanding of elementary science learning. Final draft of a report to the National Research Council Committee on Student Learning Kindergarten through Eighth Grade. Washington, DC: National Research Council.

Acknowledgments

The author thanks Dr. Frances Lawrenz for her helpful suggestions on earlier manuscripts of this work.

Perceptions of Productive Failure in Design Projects: High School Students' Challenges in Making Electronic Textiles

Breanne K. Litts, Yasmin B. Kafai, Kristin A. Searle, and Emily Dieckmeyer
breanne.litts@usu.edu, kafai@upenn.edu, kristin.searle@usu.edu, e.dieckmeyer@gmail.com
University of Pennsylvania

Abstract. The concept of productive failure has emerged as one of the key ideas for designing effective learning activities in well-defined problems. Here we report on moments of productive failure from an eight-week long workshop with 16 high school freshmen (13-15 years) who engaged in an open-ended design problem, making an electronic textile. Using students' self reports, we found that students encountered failure mostly in the design and crafting of circuits, and in these exploration phases students generated a multitude of successful and unsuccessful solutions. Our findings indicate that students' design artifacts function as a source of within-task feedback supporting their persistence through failures. Additionally, our findings highlight a need for further research examining what Kapur's consolidation phase looks like in an open-ended design environment, particularly focusing on the development of canonical problems from design activities as a way to have students experience failure and success in constrained design contexts.

Keywords: electronic textiles, learning through making, productive failure, learning designs

Introduction

One of the current debates in the learning and educational sciences is the role of instructional supports in learning and problem solving. While one group argues for the "productive success" model by providing direct instruction and scaffolding (Kirshner, Sweller & Clark, 2006), others point to "productive failure" with delayed instruction as an equally promising direction for scaffolding student learning (Kapur, 2008; Kapur & Bielaczyc, 2012). Much like the proponents of productive success who see scaffolds and collaborations as essential in propelling students' learning forward, the supporters of productive failure focus on better understanding the role of multiple representations and solutions, their role in activating prior knowledge, and the nature of peer support during the generation phase to identify which dimensions are most productive for which students and under which conditions (Kapur & Rummel, 2012). Most of these studies, however, have focused on getting students to solve well-defined canonical problems, not examining the role that failure plays in solving open-ended design problems.

In open-ended design activities, failure plays a constant and prominent role in the overall learning process. In projects that involve designing software (e.g., Kafai & Ching, 2004), building car ramps (e.g., Kolodner, Camp, Crismond, Fasse, Gray, Holbrook, Puntambekar & Ryan, 2003), or engineering bridges (e.g., Roth, 1998), students working in teams or alone constantly run into challenges as they imagine, then implement and iteratively cycle through the process on their way to product completion. For instance, in designing software, the process of debugging, or fixing the problems in the program code, is a constant part of programming and not just located in the beginning (Soloway & Spohrer, 1988). In fixing bridge structures, students learn not only about the qualities of materials but also about structural properties of different engineering designs (Roth, 1998). Identifying and addressing the small and big challenges in conceptualizing and implementing a design, whether physical or virtual, is what makes these design activities rich learning experiences. Could we consider the challenges and resolutions encountered in such design activities a form of productive failure? The answer to this question could provide the field of learning sciences with new insights and contexts in how to structure opportunities for productive failure for different kinds of learning activities.

We report on a study conducted with 16 high school freshmen ages 13-15 years who chose to participate in an eight-week long workshop in which they learned to design and program their own electronic textile using the LilyPad Arduino microcontroller (Buechley, 2006), sensors and actuators. Electronic textiles (hereafter: e-textiles) include microcontrollers, sensors for sound, touch and light and actuators such as LEDs and buzzers, that can be sewn into textiles to make interactive wearables and teach programming and engineering concepts (Buechley, Peppler, Eisenberg & Kafai, 2013). For our analysis, we focus on debriefing interviews conducted with all the students where we asked them specifically about periods of troubleshooting and debugging during their e-textile design processes. Our analysis was directed by two research questions: (1) What range of challenges do youth report encountering when crafting e-textiles? and (2) How do youth respond to and resolve these challenges? In the discussion we address what our findings from examining challenges and resolutions contribute to the growing work on designing for productive failure in educational settings.

Background

The idea that we can learn better through prior failure initially seems counterintuitive but has emerged as one of the new directions for designing effective learning activities. Productive failure defined by Kapur (2008) is first “engaging students in solving complex, ill-structured problems without the provision of support structures” (p. 379). Then, students are provided with instruction and tasked with solving canonical problems. Students who first solved ill-structured mathematics problems were significantly more successful at later solving well-structured mathematics problems compared to students who only solved well-structured mathematics problems (Kapur & Bielaczyc, 2012). The critical feature of productive failure—the use of ill-defined problems in an initial phase to achieve better problem solving performance of well-defined problems in a later phase (Kapur, 2008)—also provides promising connections to open-ended design activities. Integrating a “failure” phase into instruction during which students generate multiple representations and solution methods is instrumental in equipping those learners with more agile problem solving abilities and greater representation flexibility to later outperform students who only experienced direct instruction (Kapur 2012; Kapur & Bielaczyc, 2012). These research findings also highlight a central dimension of design learning where students turn an ill-defined into a well-defined problem by creating a solution in the form of a virtual or tangible artifact (Bielaczyc & Kapur, 2010; Kapur & Bielaczyc, 2012).

One of the key distinctions of learning in open-ended design activities, however, is that failure is not an initial one-time generative phase followed by successful completion but rather a frequent occurrence possible at multiple time points throughout the design process (e.g., Kafai & Ching, 2001; Kolodner et al., 2003; Roth, 1996; Litts & Ramirez, 2014). Learners addressing design problems not only experience but also resolve multiple failures. Hybrid designs projects where students need to identify and solve problems in multiple, overlapping domains provide a rich context for experiencing and addressing failure. For instance, in the context of robotic design, students not only have to build motor-driven models with active sensors but also have to code the programs that operate motors and sensors independently in the field, without any human direction (Sullivan, 2009). In the context of designing e-textile wearables, students are learning about crafting, circuitry and coding. We found how these overlapping domains, where materials behave in unexpected ways, are rich spaces for student learning (Kafai, Searle, Fields, Lee, Kaplan, & Lui, 2014). Students had to learn and coordinate multiple representations, i.e. blueprints for circuit design and program code for the microcontroller. We observed how students’ experiences with failure occurred not only within each respective domain (and thus would make them more comparable to the traditional productive failure studies), but also at the intersections between crafting and circuitry, coding and circuitry (Kafai, Fields, & Searle, 2014).

In this particular study, we were further interested in understanding how students themselves perceive moments of failure and resolutions as they are designing e-textiles. Such retrospective accounts are rare in the current research on productive failure, which has mainly focused on examining the impact of different conditions on students’ learning outcomes. Understanding what students consider moments of failure in the design process, how and where they deal with failure provides critical insights for designing supports in learning tools and interactions. It is this particular dimension of design learning that potentially could inform not only the larger literature on productive failure but also the work on design projects that has emphasized success, i.e., getting students to complete a design project by providing scaffolds and tools. By looking at failure as a productive rather than counterproductive dimension of learning, we can shift pedagogical directions as well students’ perceptions of such events that are inevitable in the design and learning process.

Context, participants, and data

Workshop participants and e-textile design activities

We conducted this study with 16 high school freshmen (7 boys, 9 girls, 13-15 years old) from a science magnet school in a metropolitan city in a US northeastern state. Students identified as follows: 56% Black, 19% Asian, 19% White, and 6% Multiracial. Participants selected our e-textiles workshop from a variety of workshop options as part of an immersion partnership between their school and a local science museum. The workshop spanned eight two-hour-long sessions at the museum and a ninth wrap-up session at the school. A researcher led the workshop and was supported by two graduate students who assisted with student projects and data collection. Students engaged in a series of three e-textile projects: 1) an introductory activity crafting a simple circuit with an LED, coin cell battery, and conductive thread; 2) a starter wristband project with two LED lights, coin cell battery, conductive thread, and snaps that closed the circuit; and 3) a more advanced ‘human sensor project’ where students programmed a LilyPad Arduino to activate a lighting pattern with 2-3 LED lights by touching two conductive patches that closed the circuit. While projects were scaffolded in that they built in complexity, students designed

their own circuitry and remixed existing code. We focus on the students' design processes in the human sensor project. We framed this project as a "Logo Remix" in which students remixed an existing logo (e.g., their favorite sports team or brand) by making it interactive with up to 3 LED lights, touch sensors, and a LilyPad Arduino. Each student created his or her own Logo Remix, but many worked collaboratively. As part of the design process, we instructed students to create a circuit design blueprint, which we approved before they began sewing. All students uploaded their design blueprints and received feedback via an e-textile website (ecrafting.org) from graduate students who were more advanced in e-textiles.

Data collection and analysis

We collected a range of data on students' design processes focusing on challenges students encountered and resolution strategies they employed. In addition to photo documenting the progression of students' artifacts and code, we kept extensive field notes throughout the workshop and interviewed students after they completed their projects. Interviews were semi-structured and aimed at the challenges students encountered in their crafting, coding, and circuitry, and how they resolved each challenge they shared. Ten students were interviewed in pairs due to time constraints. Our main analysis focused on students' post-interviews, which we triangulated with field notes and photos of students' design processes. Two researchers coded all of the data and iteratively developed a coding scheme around challenges and resolutions. Researchers discussed and resolved disagreements or inconsistencies. Initially, researchers completed open line-by-line and in-vivo coding (Saldaña, 2009) to get an overall sense of the data. Inspired by the productive failure literature, researchers then coded through a four-phase process: (1) Identified points of iteration/challenge (95 total); (2) Generated and applied a coding scheme of the types of challenges (e.g., knots/tangles, polarity of LEDs, etc.) remaining open to adding new codes; (3) Determined whether and how the challenges were resolved: alone, with a peer, with a teacher, or not at all; (4) Clustered codes around larger themes. For instance, challenges like knots/tangles and stray ends are connected by the underlying challenge of crafting with conductive thread. One researcher applied the coding schemes for "challenges" and "resolutions" to the field notes to triangulate with the findings from the self-report interview. No new codes were generated, but the fieldnotes shed more light on resolution strategies to remedy challenges.

Findings

We present the findings of this data analysis in two parts: 1) outline the two most prevalent themes of challenges: crafting with conductive thread and designing spatial circuitry; and 2) highlight the resolution strategies students employed to overcome these two most challenging elements of the design process.

Encountering challenges within e-textile designs

Across the interviews, students listed two overarching challenges—crafting with conductive thread, and spatial circuitry—that accounted for 68 out of 95 (roughly 72%) of all statements. Every student reported encountering both of these challenges in making his or her e-textile project. Other observed challenges related to project time constraints and working in the programming environment, Arduino.

Crafting with conductive thread

Learning to craft with e-textiles is not just about crafting skills, which are challenging in their own right (Lee & Fields, 2013), but also about how these skills intersect with circuitry knowledge in a domain specific way (Peplur & Glosson, 2012). For novices who have not fully grasped the properties of conductive thread, however, it often appears that crafting and circuitry can be separated out as discrete entities, resulting in failure to complete a functional circuit. In our analysis, crafting with conductive thread accounted for 28% of the total challenges faced by students. Asked to recount some of her sewing challenges, Kerry remarked:

Oh! Let's do all the problems I went through. Hey,...so when you wanna start [thread] the needle, it doesn't want to go in. And then when you finally get it in, you have to make a knot. And the knot doesn't wanna be formed. So then you finally make the knot and you're sewing, and sometimes...either the knot and the entire thread comes through and the needle just disappears and the the thread's left there alone, or, you know what happens? Knots come out of nowhere when you don't need them! It's like, hey, where were you when I was trying to knot you? (Int., 6/19/15, pp. 4-5).

In this reflection, Kerry highlights several of the ways in which she experienced failure in the process of crafting her e-textile project, including difficulty threading the needle, difficulty tying knots in appropriate places, and the

propensity of the thread to tangle easily. Rhett, a novice sewer, experienced a different but related problem when he tried to connect components with one giant stitch rather than many small ones. He explained, “[o]ften times I forgot to do a running stitch and I would just put it straight there and that was frustrating because I had to cut it and try again” (Int., 5/11/2015, p.2). These large stitches were easily pulled loose and failed to hold the components in place. They could also move around and touch other lines of conductive thread, creating unanticipated circuitry problems. As Kerry and Rhett illustrate, these basic sewing challenges were especially frustrating and became even more frustrating for students when they realized that, because of the uninsulated, conductive nature of the thread, sewing mistakes and circuitry mistakes were often intertwined.

All students had short circuiting issues in which their positive and negative threads tangled together or were connected by stray ends of thread and many students discussed having to repeatedly cut out and re sew lines of thread because of incorrect or crossing circuitry. Taraji was an experienced sewer who liked the thicker feeling of the conductive thread and often mended her own clothing outside of school but even she admitted that, “getting to understand, like, sewing without making the thread in the back like touch, or going over other [threads]” was difficult (Int., 5/08/15, p.3). Similarly, Sabrina observed that “another challenge was connecting it to the wrong LilyPad circuit. I messed that up, like, I know when I first did it I didn’t connect it to the negative and that messed up my whole project” (Int., 5/08/15, p.2). In these ways, students experienced multiple failures related to crafting with conductive thread in the context of their e-textiles projects. Though we did not measure their learning gains, anecdotally we can say that students developed a more robust knowledge of the properties of conductive thread and of key circuitry concepts (e.g short circuits) through their failures.

Designing spatial circuitry

In addition to experiencing failure related to an incomplete understanding of the properties of conductive thread as connected to circuit design, students also experienced failure related to *spatial circuitry*, the design and physical construction of the circuit. Spatial circuitry accounted for 43% of all challenges (41 out of 95). Spatial circuitry is not just about making sure that lines of conductive thread do not cross or that the whole project fits on a sheet of felt, but also takes into account making sure that the correct circuit components are connected to the correct ports on the LilyPad, since not all ports, for example, can read input from a sensor. Although we required students to first draw a paper and pencil representation of their design (including tracing the LilyPad and its ports and labeling the components) and to consult with one of the instructors before moving onto constructing their design, translating the design blueprint into a physical artifact proved challenging in terms of both functional and aesthetic considerations.

As Kerry recounted in her reflective interview, the placement of various design components often evolved through a degree of failure in the construction process (see Figure 1):

When I did [my project], I had to switch [the patches] like four different times and I also had to switch my LEDs because the way that my design worked is...there wasn’t a lot of space over here or through here [*pointing to top of her project*] so it was like, you had to try to cut through things and...hide some things but not a lot. And I was like trying to make everything be there but not touch and it was really confusing...LEDs had to switch and ideas had to switch” (Int., 5/08/15, p.11).

In this reflection, Kerry highlights several common aspects of spatial circuitry that challenged students, including the size, shape, and placement of their conductive patches and also an aesthetic desire to hide some of the circuitry while still having a functional project. Similarly, Taraji explained, “I didn’t really estimate all the amount of space I would have on the felt [as opposed to the piece of paper]. So when I actually started sewing on the felt I had to like think, like before I got to put my conductive patches on I had to think like how big they could be, or where exactly they could fit” (Int., 5/08/15, p.4). Across projects, students experienced failure related to realizing their design goals within the constraints of their circuitry and coding knowledge.

Overcoming challenges within e-textile designs

Working with conductive thread and designing spatial circuitry were the two most common challenges students discussed-in their reflections of design processes. In post-interviews, 6 out of 16 students reported that they did not successfully complete their project (e.g., they had not finished sewing or did not upload their code to their LilyPad). Interestingly, though, all of the students employed various resolution strategies to overcome these two challenges. In this section, we outline the range of those resolution strategies according to who students reported

they relied on to resolve the challenges: alone (55%, 52/95), teacher support (26%, 25/95), peer support (4%, 4/95), and unresolved (15%, 14/95).



Figure 1. Kerry's circuitry blueprint and partially completed project, front (l) and back (r).

Crafting with conductive thread

A large part of resolving challenges of crafting with conductive thread is having the foresight to see how crafting causes short-circuiting. Students must understand that because the thread is conductive (and not insulated) it cannot cross and their stitches must be tight to ensure strong connections. Students largely worked by themselves (55.6%) to troubleshoot challenges of crafting with conductive thread, but some required teacher (29.6%) support, and a few solicited peer support (7.4%) or were not able to resolve their challenges (7.4%).

Working on his own, Mack realized the importance of solid connections, but struggled to sew accordingly, so he generated a creative resolution. He expounded, "The hardest part was connecting the LEDs to the thread, because if I would stick it in, I only did it single [thread] so it would pop [break] and then I had to restart" (Int., 5/08/15, p.2) In response, he sewed plastic beads behind the LEDs to guide his sewing on the back of his project and secure his components (see Figure 1). Other students resolved this particular challenge by taking more traditional approaches like trimming their stray ends to prevent short-circuiting with thread crossing in the back.



Figure 2. Mack's project partially complete, front with LEDs (left) and back with beads (right).

Likewise, Kerry described how she resolved the issue of short-circuiting caused by stray ends, "I did that on my own after I was given tape. Because I saw and I was like, oh, you know, this part is so long and then [the teachers] just placed tape down on the table and so I was like, ok and decided to tape things" (Int., 5/08/15, p.7). While Kerry believed that she resolved this challenge on her own, she also illustrates how minimal, implicit teacher support (placing the tape down on the table) structured her ability to do so. In contrast, Jonathan, needed direct teacher support to tackle a short circuiting issue he was having with stray ends. He recounted, "I had to trim down a bunch of stuff in the back so that they didn't cross over...I had to be told that's what I had to do" (Int., 5/08/15, p. 3).

In other cases, students solicited help from their peers. For instance, Kendra described how her friends helped her make a proper running stitch, "Well previously my threads were really far apart, and then my

friends,...noticed that it could like really be easy to get loose...So they told me to make my threads tighter, so it'll be like difficult to [get] loose and they wouldn't be tangling. That was a helpful suggestion, so I just sewed tinier" (Int., 5/11/15, p. 6) Unlike traditional sewing where typically only one side of the project is important, both sides of e-textiles project are equally important for aesthetics and circuitry.

While we instructed students to secure their components by sewing the connection at least three times, several students still faced challenges with this technique. Kendra explained, "This [LED] is pretty loose. I didn't sew it tight enough, and it's pretty dull, too. So I don't know, those could be contributing factors [of] why...it doesn't really light up that much" (Int., 5/11/15, p. 2) Kendra struggled to understand *why* securing her components was important until she made the mistake herself, and though she left this problem unresolved, she articulated the most likely cause of her dim LED light in her reflective interview.

Designing spatial circuitry

In spite of the challenges students faced in translating their paper and pencil design blueprints into a physical reality, many referenced their original design blueprints or created multiple versions of their design blueprints in order to resolve challenges related to spatial circuitry. More than half of the students (58.5%) relied mostly on themselves to solve challenges related to spatial circuitry, but nearly one third (31.7%) requested teacher assistance to troubleshoot their circuitry. Only one student reported soliciting peer support (2.5%) and a few were unable to resolve their spatial circuitry issues (7.3%).

Students who worked alone to resolve their spatial circuitry problems often reported generating multiple design blueprints along the way. Taraji, for example, reported a careful planning process, "Before I would actually sew, I'd draw the pattern out and make sure none of them like crossed each other, and once I had that I'd just copy it off the paper and to the actual [felt]" (Int., 5/08/15, p. 3). Each time Taraji made changes to her design plan she modified her blueprint to match. Taraji's proficiency in using her design blueprint streamlined her overall design process, as she described, "I just followed each step that was listed on here and it was pretty simple" (Int., 5/08/15, p. 4). In contrast, Kerry attempted to work through spatial circuitry challenges in her design (see Figure 1) by making in-the-moment changes. In her reflective interview, she recalled:

The touch sensors, they were supposed to be...I have this right here, I have the basic idea (*showing design blueprint*). My design never changed, it's just the way I did it, I switched it. Like I had these were gonna be the conductive touch sensors (*showing cheek on design blueprint*), because these were gonna be touch sensors and they were supposed to go right here on the edge (*pointing under the eyes*), but you see how close they are to the eyes and everything it made everything a lot harder to navigate so things got weird" (Int., 5/08/15, p. 11).

Since Kerry did not reflect the changes to her spatial circuitry on her design blueprint, she failed to realize that her design iterations would result in crossing negative and positive threads, which caused a lot of frustration, because she "had to do redo everything" (Int., 5/08/15, p. 12). Both Taraji and Kerry resolved their spatial circuitry challenges on their own, however they used different strategies and this resulted in different outcomes.

Students also requested teacher assistance to troubleshoot their spatial circuitry challenges. Eagan, for example, described her biggest mistake, "Well the circuiting obviously. Cuz I just kept messing on up, but yeah I had great teachers help me out" (Int., 5/08/15, p. 8). She further explained that with teacher support she tried rotating her Lilypad and that helped resolve a lot of of her issues connecting her LEDs and touch sensors to the Lilypad without crossing positive and negative lines. Another student, Jess, reported assisting her friend with connecting her LEDs to the Lilypad, "I was helping Lauren with stuff, because she didn't know where to connect it" (Int., 5/11/15, p. 5). Even though Jess struggled herself with the concept, she understood the challenge of thinking through, "how [the] design will fit into the circuitry... [and]... not to get the thread all tangled up in each other" (Int., 5/11/15, p. 6). A few students, like Mel, didn't reach out at all: "Since we don't know anything about this part, this Lilypad, I don't know where to put which...I don't know, where to put things and stuff. So, or how it works, cuz I don't know how this thing works... it's just really confusing" (Int., 5/11/15, p. 5). Even at the end of the project Mel still struggled to understand the importance of her design blueprint and the function of the Lilypad, but did not report asking for help to resolve her confusion with these concepts. While we might read students' spatial circuitry challenges and resolutions as just a lot of frustration, we argue that students learned valuable lessons about design processes and how to use resources to solve ill-defined problems.

Discussion

We examined students' perspectives on challenges and resolutions with the goal of better understanding how the idea of productive failure can apply within an open-ended design task. All students reported moments of failure with sometimes successful, and at other times unsuccessful responses. The majority of students reported working on their own to troubleshoot and debug their projects without direct instruction. While it is likely that students received more external help than they reported, what is important is that they persisted to overcome challenges even if they did not produce fully functioning and finished products. In the following sections, we discuss whether or not our findings of productive failure resonate with those observed in previous research and outline potential instructional applications.

Understanding the productivity of failure in open-ended design tasks

The productive failure literature introduces the "solution generation effect" meaning the more solutions students generate, the more they learn (Kapur, 2015) and illustrates that this occurs in two phases: exploration and consolidation. However, the bulk of this research has been conducted using math tasks where there exists a clear canonical solution to which students can compare and contrast their self-generated solutions. In this study, we identified two ill-structured exploration phases that exist in e-textile design: (1) crafting with conductive thread and (2) designing spatial circuitry. In these exploration phases, students generated multiple solutions, many of which were non-traditional. For instance, Mack's creative solution of using beads to guide his LEDs and secure his components illustrates the productive nature of repetitively making the same mistake. But in open-ended designs, there is not a parallel canonical solution to which students can compare their self-generated solutions. Instead, there are a range of design styles and techniques students can leverage to suit their project. Hence, it is difficult to model well-structured tasks like the ones used in productive failure studies. Though open-ended designs have multiple successful trajectories and a vague definition of failure, e-textiles designs require integration of knowledge from multiple domains (crafting, circuitry, and coding). In the next phase of examining productive failure in open-ended designs, we must more intentionally explore what a well-structured consolidation phase looks like.

In this study we collected students' perspectives rather than examining their learning outcomes. Students' reflections shed light on the role of feedback that physical artifacts can provide when persisting through iterative failure. A productive failure learning design builds from unguided (exploratory) to guided (consolidation) problem-solving conditions and wraps up with direct instruction (Kapur, 2015). In these studies the role of the teacher is to provide the scaffolding that guides learning. Our study suggests that the physical artifacts students make can guide their problem solving without teacher intervention by providing enough within-task feedback to support them to persist beyond their failures. With tangible, three-dimensional artifacts as a learning support, students were able to rapidly cycle through many failures before finding a solution that worked for their project and process. Unlike a math problem, e-textile designs can provide live feedback to a student through their function. For instance, when debugging a short circuit a student can check knots, stray ends, etc. and debug the project until the LED turns on. The feedback from physical artifacts perhaps introduces an additional form of guidance to consider in future productive failure learning designs.

Engineering productive failure in open-ended tasks

The findings from our research illustrate that students encounter many moments of failure among the process of completion, two of which we presented and discussed in more detail. Our argument is that failure is embedded in the design process rather than a separate pre-planned activity that sets the stage for later successful problem-solving. Observing and hearing about the multiple difficulties that students experience along the way suggest that we should consider the design of supports. In previous design studies such supports that have taken the form of reflections or design diaries (e.g., Kafai, 1995; Roth, 1998) to help novice designers with completion of their artifacts, alone or together. Perhaps the presentation of supports could also come in form of challenges or problems that mirror features of productive failure. For instance, we might present students with artifacts that have intentional problems built-in that they will be asked to fix. In the context of e-textiles this could actually mean that teachers or researchers design and make an e-textile that has faulty crafting, circuits, or code by design (See Figure 2). Based on what we know from observing students' challenges, we could incorporate intentional problems in an e-textile that we ask students to fix or debug (Fields, Searle, Kafai, & Min 2012).

Furthermore, we could think carefully about the time point in the design process when we would ask students to fix the problem. The later beginning phase when students have already some experience with crafting and circuit design seems like a reasonable time point. Indeed, this idea of providing students with faulty or buggy programs is not a new one and has been used in teaching novice programmers. For instance, the "Debug-me Studio" in Scratch presents students with a set of programs that have strategically planted bugs to help them better

understand programming in playful way (Griffin, Kaplan & Burke, 2012). We see no reason why we cannot adopt the same approach for tangible artifacts such as e-textiles that often feature a larger and interconnected array of challenges cutting across the disciplinary boundaries of crafting, circuitry, and code.

We conclude our paper with a quote from a student, Eagan, who perhaps best highlights the benefits of working through multiple challenges, “Even though I never got one to light up (laughing)... I learned from my mistakes. I did just, ‘oh I made a mistake, ugh, give up,’ but I learned from my mistakes” (Int., 5/08/15, p. 8). Whether or not failure occurs and promotes learning in well-defined or open-ended problem settings, it is clear that mistakes which are often perceived as roadblocks can become stepping stones to success.

References

- Buechley, L. (2006). A construction kit for electronic textiles. In *Proceedings of IEEE International Symposium on Wearable Computers (ISWC)* (pp. 83-92). Montreux, Switzerland.
- Buechley, L., Peppler, K., Eisenberg, M., & Kafai, Y. (2013). *Textile Messages: Dispatches from the World for e-Textiles and Education*. New York, NY: Peter Lang.
- Fields, D. A., Searle, K., Kafai, Y. B. & Min, H. (2012). Debuggemes to assess student learning in e-textiles. In *SIGCSE '12 Proceedings of the 43rd ACM technical symposium on Computer Science Education* (pp. 699-699). New York, NY: ACM
- Griffin, J., Kaplan, E., & Burke, W. Q. (2012). Debug'ems or other Deconstruction Kit for STEM Learning. In *Proceedings of Integrated STEM Education Conference (ISEC)* (pp. 1-4). IEEE.
- Kafai, Y. B. (1995). *Minds in Play*. Mahaw, NJ: Lawrence Erlbaum Publishers.
- Kafai, Y. B., and C. C. Ching. Affordances of Collaborative Software Design Planning for Elementary Students' Science Talk. *Journal of the Learning Sciences* 10, no. 3 (2001): 323-63.
- Kafai, Y. B., Searle, K. A., Fields, D. A., Lee, E., Kaplan, E. & Lui, D. (2014). A Crafts-Oriented Approach to Computing in High School: Introducing Computational Concepts, Practices and Perspectives with E-Textiles. *Transactions on Computing Education*. 14(1), 1-20.
- Kapur, M. (2008). Productive Failure. *Cognition and Instruction*, 26(3), 379-424.
- Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, 40 (4), 651-672.
- Kapur, M. (2015). Learning from productive failure. *Learning: Research and Practice*, 1(1), 51-65.
- Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *The Journal of the Learning Sciences*, 21 (1), 45-83.
- Kapur, M., & Rummel, N. (2012). Productive failure in learning and problem solving. *Instructional Science*, 40 (4), 645-650.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work. *Educational Psychologist*, 41(2), 75-86.
- Kolodner, J., Camp, P., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting Learning by Design™ into practice. *Journal of the Learning Sciences*, 12, 495-457.
- Lee, V. R. & Fields, D. A. (2013). A clinical interview for assessing student learning in a university-level craft technology course. In the *Proceedings of FabLearn 2013*, Palo Alto, CA.
- Litts, B., & Ramirez, D. (2014). Making people fail: Failing to learning in games and making. In *Proceedings of the tenth annual Games+Learning+Society Conference*. ETC Press: Pittsburgh, PA.
- Roth, W. M. (1998). *Designing communities*. Dordrecht: Kluwer Academic Press.
- Saldana, J. (2009). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.
- Soloway, E. & Spohrer, J.C. (Eds.). (1989). *Studying the novice programmer*. Hillsdale, NJ: Lawrence Erlbaum.
- Sullivan, F. R. (2009). Robotics and science literacy: Thinking skills, science process skills and systems understanding. *Journal of Research in Science Teaching*, 45(3), 373-394

Acknowledgments

This material is based upon work supported by the National Science Foundation (ITEST-1509245) to the second author. The views expressed are those of the authors and do not necessarily represent the views of the National Science Foundation, Utah State University, or the University of Pennsylvania.

The Effects of Coaching on the Teaching and Learning of English in Indian Government Schools

Harini Raval, Center for Learning Resources, hariniraval@gmail.com
Chittaranjan Kaul, Center for Learning Resources, chittaranjan.kaul@gmail.com
Susan McKenney, University of Twente, susan.mckenney@utwente.nl

Abstract: Although English is mandatorily introduced as a second language early in majority of government primary schools in India, its quality remains dismal due to lack of appropriate curricula, poor ESL teacher competencies and weak professional development opportunities. In an effort to promote the adoption of an innovative ESL program in urban government schools, coaching was introduced to support teacher and implementation of the new program. Through classroom observations, structured interviews and learner pre/post tests, the impact of coaching on teacher practices, teacher beliefs and student outcomes were studied and compared with those of teachers who were not coached. The powerful results found in the coached-teacher classrooms build a strong case for reviewing the current professional development policies within Indian schools, both for English and in general. Implications for research, policy and practice are discussed.

Keywords: ESL, teachers, coaching, student learning

Introduction

This paper presents a study investigating the role of coaching support for primary school teachers while adopting an innovative bilingual approach to teaching English as a second language (ESL). An interactive ESL program was implemented in 250 regional medium government schools in an urban Indian city across grades three and four. In grade three, the program aimed at building English reading skills (Let's Read and Write – LRW), and in grade four, it focused on developing basic speaking and comprehension skills (We Learn English – WLE). This widely tested program, created by a non-government resource agency, was designed to aid teachers with little to no ESL competencies and habituated to rote-based teaching methods. Extensive real-time bilingual audio-input was used in the program to introduce new English content and guide teachers in the use of new interactive teaching strategies. Under a directive by the school administration, all the teachers participated in an experiential workshop which oriented them towards the principles of the pedagogical approach in the program, while offering an opportunity to observe and practice the teaching strategies through demonstrations and micro-teaching sessions, respectively. The school administration planned to launch this initiative merely on the basis of these workshops, in a manner characteristic of the broader Indian education system where professional development is synonymous with one shot trainings without any ongoing guidance and supervision (Ramchandran, Pal, Jain, Shekar, & Sharma, 2005; Tyagi, 2010). The resource agency, in order to build a case for creating on-going support critical for adoption of reforms, suggested that coaching support be introduced to 30 schools as a pilot. A study was commissioned to examine the impact of coaching support on the 30 schools. This paper describes the nature of this impact in terms of practices and beliefs of teachers who were coached and differences in the student outcomes related to the coached (C) and uncoached (uC) teachers.

Theoretical underpinnings

Coaching, as a form of support for professional growth, has been widely endorsed for its potential to impact teachers practice positively, by helping them integrate newly learned skills into their ongoing practice and in improving their efficacy (Cantrell & Hughes, 2008; Joyce & Showers, 1982; Poglinco, et al., 2003). According to Joyce and Showers (1982) a coach's role is to provide support to teachers in their new endeavor through provision of companionship, technical feedback, analysis of application, adaptation to the students and personal facilitation. Coaching support is provided on an individual basis or in a group. In this program, the coaching support was individual in nature, in which a skilled peer would visit a teacher's ESL class and provide support. This support comprised of activities like, co-teaching, modelling, debriefing the challenges of implementation, observation and feedback on enactment (Raval, Mckenney, & Pieters, 2010; Gibbons & Cobb). It was aimed at helping teachers adapt the program realistically to their dynamic classroom challenges (Putnam & Borko, 2000), involving students who had no background to English or to the new learning approach that involved audio-technology and extensive interaction.

Clarke & Hollingsworth (2002) emphasize the importance of understanding the process by which teachers grow in order to facilitate their professional development effectively. Since the purpose of this study was to draw attention of administrators and decision makers towards the significance of ongoing professional support for teachers especially during the early stages of a reform, it sought to examine both the changes in different kinds of professional outcomes of teachers, as well as how these changes took place. The model of professional development advocated by Clarke and Hollingsworth, aided the study's intentions well, as it (a) highlights the principle domains of the teacher's world in which change occurs as a result of professional development, and (b) unpacks the processes that mediate this change. Their model features four domains: (a) the external domain represents an external source of information and in this study refers to the coaching intervention; (b) the personal domain represents the teachers' knowledge, beliefs and attitudes, which in this context would include ideas including but not limited to general pedagogical knowledge, pedagogical content knowledge, and motivation, (c) the domain of practice represents professional experimentation which specifically refers to the implementation of the ESL program and, (d) the domain of consequences represents salient outcomes that pertain to ESL competencies of students in this study.

The interaction between these domains happens as a result of two processes that teachers undertake: enactment and reflection. Enactment refers to the conscious putting into action of a(n) (new) idea or practice, and plays a critical role in transferring change from one domain to another domain. For instance, when a teacher develops an interest or conviction in a new approach (the personal domain), it would lead to enactment of the approach and hence influence the domain of practice. Reflection takes place not just on student outcomes but also on other perceived consequences of their actions during enactment. Similar two-way reflective links connect teachers' practice, the interpretation of outcomes (consequences of the practice) and revision of knowledge or beliefs.

Methods

The aim of this study was to assess if and how coaching support influenced the teaching of ESL through the bilingual interactive program in the primary schools. Based on the Clarke and Hollingsworth (2002) domains, the research was guided by the following three questions:

1. What changes in the practice and personal domains were exhibited by coached teachers?
2. According to them, what engendered the changes?
3. What were the consequences for WLE and LWR learners of coached and uncoached teachers?

Data collection procedures

Domain of practice: Professional experimentation

Changes in the domain of professional experimentation, that is, the implementation of WLE and LRW by teachers, were determined by analyzing the structured classroom observations conducted by the coaches for grades three and four. The coaches collected this data as a part of their ongoing responsibility for a period of six months. 256 classroom observations covering 64 teachers were conducted for LRW at the grade three level. Therefore each teacher received about four visits from a coach in the six months. For grade four, 140 classroom observations of WLE from 35 teachers were analyzed. Again, each teacher received about 4 coaching visits in the six-month period. Additionally, 10 supervisors of teachers were interviewed to understand what changes they had perceived in the teachers' ESL teaching practices.

Personal domain: Knowledge, beliefs and attitudes

The changes in the personal domain were examined through semi-structured personal interviews with 20 teachers from C-schools. Teachers were asked to reflect on what specific changes they experienced as a result of the coaching support. Each interview lasted 20 to 30 minutes. The teachers were selected through convenience sampling, based on time available for participating in the interview.

Perceptions regarding causes of change

The perceptions about engendered changes were examined through interviews with the same 20 teachers mentioned above. The teachers were asked to reflect and elaborate on if and how changes that they experienced were influenced by the coaching support.

Domain of consequence: Salient outcomes

The domain of consequence was examined through a pre/post test assessing ESL outcomes of students of grade three and four. Approximately 20 percent of both coached (C) and un-coached (uC) schools were randomly selected (10 out of the 40 C-schools, and 35 of the remaining uC-schools). Within each selected school, 10 students of grade three and of grade four were randomly selected for the pre-test. The students from grade three (n-C=100, n-uC=344) took the LRW post-test, whereas those from grade four (n-C=166, N=uC=343) took the WLE post-test.

Instruments

Classroom observation

Classroom observations were conducted with the help of a curriculum profile. Curriculum profiles are sets of statements about activities and intended behaviors of teachers during the observed lessons (Ottevanger, 2001; van den Akker & Voogt, 1994). The extent to which teachers realize these intentions is established by crediting scores based on classroom observations, which results in the actual practice profile of the teacher. Three curriculum profiles were used to develop a profile of the quality of implementation of the LRW program in grade three, and the WLE program in grade four. For each item the observer could provide a score of A, B, C or D wherein 'A' referred to 'Throughout the lesson/always; 'B' referred to 'Often, 'C' referred to sometimes, and 'D' as Never/Not attempted. Both tools had items which represented strategies needed to implement the English language materials.

The grade three observation tool had three main dimensions, (1) general pedagogical skills required to use the materials well, (2) ESL pedagogical content skills and (3) learner responses. Each of these parameters had other sub-items. Together, the tool comprised 17 items divided across these three parameters. The grade 4 tool contained 16 items as the WLE program was far more structured and simple to execute as compared to the LRW program. These items were divided across three parameters: general pedagogical skills, pedagogical content skills and learner responses. The items for both curriculum profiles are shown in Tables 1 and 2, respectively. Observers practiced completing observation profiles in non-sampled classrooms until acceptable levels of inter-rater reliability were met. (Cohen's kappa of .7 or higher).

Teacher and supervisor interviews

Supervisor interviews were semi-structured, focusing on: their perceptions of the benefits of the coach, changes in teaching strategies perceived in the C- and uC schools, changes in the responses of learners between C- and uC schools. Similarly teacher interviews were also semi-structured and focused on: nature of support received from the coach, benefits or dis-benefits to their teaching and learner behavior as a result of the coaching intervention and perceptions about how the change took place.

ESL pre/post-test

The grade three test assessed the impact of LRW by measuring reading and writing skills. The test administered in grade four aimed at assessing impact of WLE by measuring speaking and comprehension skills. The specific skills tested are reflected in Tables 3 and 4, respectively.

Data analysis

Data on classroom observation was analysed by examining the teachers' transition from lower to higher levels (from D to A) within the six months. The positions recorded for September and February were compared. The percentage of teachers who had progressed by one level, two levels and three levels across various parameters in the curriculum profile between September and February was calculated. Differences between the pre and post test scores for the C-schools and the uC-schools were calculated. Effect sizes were determined to indicate the strength of the learning gain for both types of schools. Interview data was analysed based on the interconnected model of teacher growth. Specifically, to answer the question on changes initiated by the coaching support, the reflections by teachers indicating a change were categorized into domains of *practice, consequence or personal* using Atlas ti. Once domains of change were identified, causal relationships and interactions across domains were sought in the teacher-given descriptions.

Findings

Domain of practice

Data on professional experimentation obtained by classroom observation from class 3 (Table 1) revealed that there was a general improvement amongst teachers over the six months. The average percentage of teachers who moved up by one level, two and three levels in the general pedagogical skills was 32.3 percent, 10.6 and 4.3 percent respectively. In all, 47.2 percent of teachers reflected progress in the general pedagogical skills. Similarly, for the pedagogical content skills, 38.7 percent teachers moved up by one level, 9.8 percent teachers moved up by two levels and 5.8 percent teachers moved up by three levels. This totaled 54.9 percent teachers who had progressed at least by one level. Finally in the learner responses, the average percent of teachers who moved up by one level was 34.3 percent, those who moved up by two levels was 10.9 percent, and those who moved up by three levels was 1.6 percent. In all 46.6 percent teachers reflected a progress in the quality of learner response.

Table 1: Teachers in grade 3 whose classroom practice moved up by 1, 2, 3 or 4 categories

Items	No of categories				Total
	1	2	3	4	
<i>General Pedagogical Skills</i>					
Teacher is prepared (has read instructions, ready with teaching aids)	20.7%	10.3%	6.9%	0.0%	37.9%
Teacher gives clear instructions when introducing and conducting activities	34.5%	10.3%	3.4%	0.0%	48.2%
Teacher uses teaching aids	30.8%	11.5%	7.7%	0.0%	50.0%
Teacher asks questions to students.	35.7%	10.7%	0.0%	0.0%	46.4%
Teacher gives students time to answer	40.0%	10.0%	3.3%	0.0%	53.3%
Average	32.3%	10.6%	4.3%	0.0%	47.2%
<i>Pedagogical content skills</i>					
Teacher teaches phonemic sounds appropriately	46.4%	7.1%	3.6%	0.0%	57.1%
Teacher teaches oral vocabulary appropriately	25.0%	8.3%	8.3%	0.0%	41.6%
Teacher teaches sight vocabulary appropriately	40.0%	13.3%	13.3%	0.0%	66.6%
Teacher helps students read/write by giving them contextual clues to answer/understand	42.3%	3.8%	3.8%	0.0%	49.9%
Teacher helps children decode when reading/writing	28.6%	10.7%	0.0%	3.6%	42.9%
Teacher clarifies meanings of unfamiliar or new words and sentences	59.1%	18.2%	4.5%	0.0%	81.8%
Teacher uses a bilingual approach	29.6%	7.4%	7.4%	0.0%	44.4%
Average	38.7%	9.8%	5.8%	0.5%	54.9%
<i>Learner Response</i>					
Students are able to make letter sound association	31.0%	6.9%	3.4%	0.0%	41.3%
Students are able to decode	33.3%	18.5%	0.0%	0.0%	51.8%
Students are able to read sight vocabulary	38.1%	14.3%	4.8%	0.0%	57.2%
Students read with understanding	40.0%	8.0%	0.0%	0.0%	48.0%
Students ask questions.	27.6%	6.9%	0.0%	0.0%	34.5%
Average	34.0%	10.9%	1.6%	0.0%	46.6%

Data on professional experimentation obtained by classroom observation from grade 4 (Table 2) also revealed a general improvement amongst teachers over the six months. The average percentage of teachers who moved up by one level, and two levels in the general pedagogical skills was 41.7 percent and 4.1 percent respectively. In all, 45.7 percent of teachers reflected a progress in the general pedagogical skills. Similarly, for the pedagogical content skills, 36.2 percent of teachers moved up by one level and 4.7 percent of teachers moved up by two levels. This totaled to 40.9 teachers who had progressed at least by one level. Finally in the learner responses the average percent of teachers who moved up by one level was 43.0 percent and those who moved up by two levels was 6.9 percent. In all 49.9 percent of teachers reflected a progress in the quality of learner response.

Table 2: Teachers in grade 4 whose classroom practice moved up by 1, 2, 3 or 4 categories

Items	No of categories				Total
	1	2	3	4	
<i>General Pedagogical Skills</i>					
Teacher gives prior instructions and introduces the lesson	33.3%	0.0%	0.0%	0.0%	33.3%
Teacher is listening attentively	43.5%	0.0%	0.0%	0.0%	43.5%
Students are listening attentively	43.5%	8.7%	0.0%	0.0%	52.2%
Teacher provides interaction opportunities as directed	35.3%	5.9%	0.0%	0.0%	41.2%

Teacher synchronizes time with the audio	57.9%	5.3%	0.0%	0.0%	63.2%
Teacher engages all students	36.4%	4.5%	0.0%	0.0%	40.9%
Average	41.7%	4.1%	0.0%	0.0%	45.7%
<i>Pedagogical content skills</i>					
Teacher uses English as classroom language	28.6%	4.8%	0.0%	0.0%	33.4%
Teachers pronounces clearly	36.4%	13.6%	0.0%	0.0%	50.0%
Teacher creates time for practicing listening and speaking	42.1%	5.3%	0.0%	0.0%	47.4%
Teacher clarifies meaning of new words	38.9%	0.0%	0.0%	0.0%	38.9%
Teacher ensures that they respond with understanding	35.0%	0.0%	0.0%	0.0%	35.0%
Average	36.2%	4.7%	0.0%	0.0%	40.9%
<i>Learner Response</i>					
Children answer with confidence	42.9%	4.8%	0.0%	0.0%	47.7%
Children answer with understanding	57.1%	4.8%	0.0%	0.0%	61.9%
Children answer accurately	33.3%	9.5%	0.0%	0.0%	42.8%
Children ask questions/give instructions confidently	31.6%	10.5%	0.0%	0.0%	42.1%
When a child is answering, other children listen attentively	50.0%	5.0%	0.0%	0.0%	55.0%
Average	43.0%	6.9%	0.0%	0.0%	49.9%

From external domain to personal domain to domain of practice

During interviews, fifty percent of the participants shared that regular monitoring of their work, the sense of someone watching over, was a motivator towards regular implementation of the program, in absence of which they succumbed to the pressure of administrative tasks. Teachers shared sentiments like, *“When she comes, then I remembered that I have to teach, because otherwise I would be busy with various (administrative) routines.”* and, *“Because of the her presence I would remember that someone is coming and I have to undertake this program; I would feel someone is watching my progress and that kept up my motivation to teach regularly.”*

From external domain to domain of professional practice

Forty percent of the teachers shared examples of direct changes in their *professional practice* as a result of the support provided by the external stimulus - the coach. One teacher shared, *“As a result of the coach’s intervention, we started using English words in the class while teaching.”* While another quoted, *“There was a question in the radio lesson about the occupation of the parents. I did not know the name of several occupations in English. So the coach suggested that sometimes I may not know the English word, in which case I can retain the Marathi word for the specific occupation, but I should encourage the use of the English sentence structure. I took this suggestion and asked the children to use the Marathi word for their parent’s occupation.”* Speaking about the LRW experience, yet another respondent expressed, *“She explained to me how to teach letter sound association using the picture cards. She first demonstrated how to use the materials, and after that I began to use the material.”*

Personal domain

Changes in the personal domain and perceptions about how they were engendered were understood through interviews with the teachers. They shared that they had experienced personal changes in terms of greater regularity in teaching, a greater sense of motivation and confidence, sense of learning and support, new knowledge and beliefs about ESL as well as general pedagogical strategies.

From external domain to personal domain

Eighty percent of the teachers expressed the availability of the coach gave them access to new ideas, and new knowledge which was unavailable earlier. For instance, one teacher shared that she viewed the coach as someone who helped her become more aware of the gaps in her teaching, and offered alternatives. Another teacher expressed how the coach exposed her to new ESL concepts: *“When the coach used to come they would give us new ideas, tell us how to teach the lesson, they would talk about what children should learn in English.”*

Several teachers gave examples of how, with the help of demonstrations, the coach helped them clarify how to use the material effectively. One of them described that, *“She would also demonstrate how to teach. When children had to listen to the radio lesson, if children did not understand something, we would pause the radio, and she would guide me: explain why a particular question is answered in a particular way. For e.g., if the question was what is your name in the first lesson. If children did not understand the question, she would pause the question, then she would explain what this question means and instruct the child that you must give your name to this question. She would also clarify that you are not to simply give your name, but you must answer: My name is....”*

From external domain to domain of professional practice to (reflective link) personal domain

Fifty percent of the teachers talked about a change in their personal domain, in terms of an increase in their knowledge and confidence about teaching ESL after trying it out with the coach's help. Many teachers expressed similar sentiments as this one who expressed that, *"Because of the coach's presence we learnt to teach in the way we were trained. My confidence has increased as a result of teaching the children."*

From external domain to domain of professional practice to domain of consequence to personal domain

Fifty percent of the teachers who were interviewed, referred to personal changes on account of the positive effects of their professional experimentation on children. The following excerpts from two different teachers reflects a confidence developing in the new approach of teaching ESL, which is grounded in actual changes in student ESL outcomes after trying the new approach. Both teachers ascribe changes to the coach's role in helping them try out the material in the first place. One of them explained, *"I used to teach according to the old method, where I used to talk and children used to listen; but as a result of the coach's help I have been able to use the new material well, and therefore I see a definite change in children's learning."* While the other shared, *"The coach helped me to encourage students to speak various sentence forms (tense, singular-plural). For e.g., I am playing, I am cooking etc. She also encouraged me to do likewise with is, are, am etc. Even today students are able to recognize tense, and speak in the correct tense. I have not seen this in 17 years of teaching, but I have seen such difference in the last two years."* Another teacher reveals a growing belief in the merits of the new approach which she trialed with the coach's help, which improved responsiveness in her students in this reflection, *"The coach helped me to use the material correctly. With the right use material, there were many changes in my teaching style. Earlier I used to be the only one speaking, and I would not get a response from the children. I was myself confused whether the children understand me or not. I would make them write the same word repeatedly. Sometimes I used to force a response out of children, and they responded with fear. But as I used this material correctly, I got an enthusiastic response from across the classroom. Those children who did not participate at all, now take an interest in the subject. Their fear of English has disappeared. I have also learnt to try out various ways of eliciting student's participation."*

From external domain to domain of consequence to personal domain

Ten percent of the teachers indicated that the indicated an impact on their personal domain as a result of the coach's interaction with the students. *"They would come and ask questions, and when children were able to respond, I would feel satisfied that someone from outside has come, and children are able to respond. I would feel very encouraged when they would appreciate how my children are progressing."*

During interviews supervisors of teachers also commented on the differences they perceived between coached and un-coached schools. They unanimously stated that the learning outcomes coached schools were better than in non-coached schools. For instance, several of them found that the children in the coached school spoke English with a clearer pronunciation of words. Some of them also discussed the relative difference of interest and engagement level amongst students towards the subject. For instance, a supervisor shared, *"Children in coached schools love English, and those in non-coached schools have lower levels of interest in English."* Another added that, *"the confidence of children in coached schools has risen, and their fear of using English has reduced."* A third supervisor shared his observation stating that, *"Children in non-coached schools do not respond to questions as quickly as from coached schools."*

Domain of consequence

Pre and post-test scores of students' ESL outcomes were collected. Effect sizes reflected the comparative learning gains between the students from coached and un-coached schools (Table 3). Class 3 scores for LRW outcomes, reflect substantial difference between the learning gains of the coached (C) versus un-coached (uC) school students. The effect size scores for the un-coached schools are small, which reflect a poor learning gain between the pre and the post-test scores. Whereas the effect sizes of the coached-schools are moderate in the first two skills of understanding capital and small letters, and large in the remaining skills, which indicates an impressive learning gain.

Table 3: Learning gains of students for the LRW program conducted in Grade 3

ESL items	Baseline		Post-C		Post-uC		Gain-C	Gain-uC
	Mean	SD	Mean	SD	Mean	SD	Cohen's d	Cohen's d
Identifying Capital Letters	3.30	1.19	3.94	0.24	3.67	0.79	0.66	0.39
Identifying Small Letters	3.08	1.25	3.79	0.50	3.04	1.35	0.67	0.03

Identifying First Letters	1.80	1.47	3.12	0.91	2.55	1.43	1.01	0.13
Identifying Last Letters	1.73	1.44	3.23	1.11	2.41	1.57	1.12	0.32
Identifying Middle Letters	1.00	1.13	2.42	1.19	1.75	1.46	1.24	0.36
Identifying Correct Word	1.40	1.14	2.92	0.99	2.28	1.25	1.39	0.46
Matching Picture with Word	1.92	1.23	3.19	1.24	2.29	1.50	1.03	0.35
Matching Picture with Sentence	1.31	1.12	2.70	1.22	2.03	1.46	1.20	0.30

Similarly, the outcomes for the WLE program (Table 4) indicate that the learning gains of the coached (C) and un-coached (uC) schools are negligible for the question on listening skills. This is likely due to a ceiling effect, as the pre-test scores for this skill were high. However, in the more complex skills for production, which include answering the question in a partially correct, and fully grammatically correct form, the learning gains of the coached skills are visibly better as reflected in the large effect size.

Table 4: Learning gains of students for the WLE program conducted in Grade 4

Skills	Baseline		Post-C		Post-uC		Gain-C	Gain-uC
	Mean	SD	Mean	SD	Mean	SD	Cohen's d	Cohen's d
Listening Skill	7.51	2.39	7.95	2.49	7.41	2.50	0.02	-0.04
Speaking: Full sentences	1.27	1.53	2.70	2.14	1.46	1.78	0.77	0.11
Speaking: Broken sentences	1.30	1.76	2.51	1.39	1.93	1.61	0.76	0.37
Speaking: Single words	3.62	2.74	2.53	1.81	3.66	2.50	-0.47	0.02

From external domain to domain of practice to domain of consequence

Forty percent of the teachers expressed that the coach had inspired them to adopt the new method and that had led to various improvements in student learning. Teachers voiced feelings such as, *"I had a student who learnt how to spell words phonetically, she could read the long passages from the LRW book. This was because the coach guided me about how to teach those lessons, and even corrected the limitations in the way I was teaching the lesson."* A commonly echoed sentiment was, *"With their help our teaching skill has improved and children's English (has improved)."*

From external domain to domain of consequence

Forty percent of the teachers expressed that their students experienced direct benefits as a result of the coach's visit. Teachers shared vividly how they saw this happen by explaining that, *"They ask children questions, and children feel enthused by the fact that someone from outside is coming and asking questions."* And, *"When children took lead in participating during the coach's visit, it encouraged other children, for e.g. some children go up and read from the book, then those children who were slow in reading also got encouraged to do so and they felt that I should also know how to read and get appreciation from a visitor. The coach once appreciated four students and that got 8 other kids motivated."* Some more teachers expressed similar experience expressed in this reflection, *"As a result of the coach's visits, because she used to conduct good demonstrations, children really got interested in the language."*

Conclusion and implications

This paper presents powerful evidence of how coaching support elevated the quality of ESL practices and outcomes within a short period of six months. A substantial number of teachers in the coached schools adopted desirable pedagogical practices with the help of the coaching support and also exhibited more engaged learner responses over a period of time. Learner outcomes in the coached schools, especially in the more complex skills were better than those in the un-coached school. Teachers' perceptions not only validate this direct evidence but also help nuance it further by explicating how coaching as the external stimulus triggered a variety of interactions across the various domains, and resulted in an experience of growth and learning. They candidly admitted that the coach motivated them to direct their focus from administrative tasks to teaching ESL regularly. It is important to note that in a context such as Indian schools where teacher absenteeism and inactive teaching time are pervasive (Kremer, Murlidharan, Nazmul, Hammer, & Rogers, 2005), an intervention like coaching may have a substantial contribution to make towards regularizing teaching. Further, where teachers themselves do not speak English, the

challenges of implementing an innovative ESL program are multiplied. While on one hand they themselves do not speak the language, they are faced with the challenge of fostering interaction in the classroom sensitively, so that students can develop the language skills necessary for confident expression (Kapur, 2013). The teachers in this study illustrate how coaches helped them meet these challenges in real time, which eventually led to improved ESL outcomes of students. Such a situated instructional focus is considered an important attribute of good quality professional development (Raval, Mckenney, & Pieters, 2010; Putnam & Borko, 2000). Finally the study also illustrates that besides offering tangible knowledge and skills to teachers, coaching also has the potential to support the development of their efficacy while they are implementing the new program (Cantrell & Hughes, 2008). Even more noteworthy are teachers' reflections about how, by experimenting with new teaching methods and encountering varied student learning, teachers start believing in new ways of teaching ESL. Consistent with the description of Clarke & Hollingsworth (2002), teachers shared vivid examples of exploring a new teaching strategy, reflecting on the consequences of that exploration and identifying notable outcomes, in turn beginning to believe in the value of that strategy.

Evidence on the effects of coaching is relevant for Indian schools (and other under-resourced countries) as they confront the paradox of declining educational outcomes, low teacher attendance and efficacy in spite of large investments in infrastructure and teacher training. While resolving this stalemate will involve paradigm shifts in terms of pedagogy and governance (Murlidharan, 2013), this paper highlights the contribution of professional support solutions like coaching which offer situated learning opportunities for teachers, and thus, not only have motivational value, but also lead to enhanced teaching and learning. Much of the existing research base on Indian primary education focuses on the current gaps in policies and practices. This paper emphasizes the need for more widespread research on promising professional development interventions, their benefits, the manner in which they are implemented and the challenges that they encounter.

References

- Cantrell, S. H., & Hughes, H. K. (2008). Teacher Efficacy and Content Literacy Implementation: An Exploration of the Effects of Extended Professional Development with Coaching. *Journal of Literacy Research, 40*, 95–127.
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 9*, 947–967.
- Gibbons, L. K., & Cobb, P. (n.d.). "Content-Focused Coaching Practices Implicated in Designing Potentially Productive Coaching Activities to Support Teachers' Learning."
- Joyce, B., & Showers, B. (1982). The Coaching of Teaching. *Educational Leadership, 40*, 4–16.
- Kapur, K. (2013). The Role of Mentoring in Developing Reflective Practices for CPD. In R. Bolitho, & P. Amol (Eds.), *Continuing Professional Development: Lessons from India* (pp. 96-107). New Delhi, India: British Council.
- Kremer, M., Murlidharan, K., Nazmul, C., Hammer, J., & Rogers, H. F. (2005). Teacher Absence in India: A snapshot. *Journal of the European Economic Association, 3*(2–3), 658–667.
- Muralidharan, K. (2013, April). Priorities for primary education policy in India's 12th five-year plan. In *India Policy Forum* (Vol. 9, pp. 1-46).
- Ottevanger, W. J. W. (2001). *Teacher support materials as a catalyst for science curriculum*
- Poglinco, S., Bach, A., H. K., Rosenblum, S., Saunders, M., & Supovitz, J. (2003). *The heart of the matter: The coaching model in America's Choice schools*. Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education.
- Putnam, R., & Borko, H. (2000). What do new ways of research and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4-15.
- Ramchandran, V., Pal, M., Jain, S., Shekar, S., & Sharma, J. (2005). *Teacher Motivation in India*. London: DFID.
- Raval, H., Mckenney, S., & Pieters, J. (2010). A conceptual model for supporting para-teacher learning in an Indian nongovernmental organization. *Studies in Continuing Education, 32*(3), 217-234.
- Tyagi, R. S. (2010). School based instructional supervision and the effective professional development of teachers. Compare: *A Journal of Comparative and International Education, 40*(1), 111-125.
- van den Akker, J., & Voogt, J. (1994). The use of innovation and practice profiles in the evaluation of curriculum implementation. *Studies in Educational Evaluation, 20*(4), 503-512.

Acknowledgments

We thank participating teachers and students. This work was funded by Michael and Susan Dell Foundation. Grant #999-888-1111.

Listening Versus Looking: Learning About Dynamic Complex Systems in Science Among Blind and Sighted Students

Ran Peleg, University of Haifa, rpeleg@gmail.com
Sharona T. Levy, University of Haifa, stlevy@edu.haifa.ac.il
Orly Lahav, Tel-Aviv University, lahavo@post.tau.ac.il
Noha Chagab, Tel-Aviv University, K_noha@walla.com
Vadim Talis, Tel-Aviv University, talisv@yahoo.com

Abstract: Listening to Complexity (L2C) is a novel sonified representation of complex systems that addresses the lack of exploratory materials in science for students who are blind. It complements agent-based models of systems by representing a micro-level individual's properties, events and location, together with macro-level variables of the system. The study examines whether this perceptual compensation creates a comparable learning environment. Students who are blind worked with auditory models; sighted students used visual models. Conceptual understanding and systems reasoning were assessed with pre- and post-test questionnaires. Learning processes were captured by analyzing their workbook writings. The L2C environment not only supported blind students' learning similarly, but also furthered their learning with one of the more challenging concepts, diffusion. Some concepts relating to the micro-level were learned earlier among the blind students. It seems that auditory representation increases sensitivity to the micro-level interactions in a way less accessible in visual representations.

Keywords: model-based learning, complex systems, science education, special education, sonification

Problem statement

How can we support students who are blind in gaining access to exploratory learning materials in science? Blind students have been integrated into public schools for decades. However, since most learning materials in science classes are based on visual information, they are often excluded from full participation (Beck-Winchatz & Riccobono, 2008). Several manuals have been written on how to teach science to students who are blind and visually impaired (Kumar et al., 2001; Willoughby & Duffy, 1989). Few learning environments based on assistive technologies have been created to support science learning and research regarding such technologies is sparse (Farrell et al., 2001, 2008; Wies et al., 2000; Zaborowski, 2006).

Addressing the problem: Listening to complexity (L2C)

Listening to Complexity (L2C) addresses a central need among people who are blind: providing equal access to the science classroom, allowing them to interact with exploratory materials, independently collect data, adapt and control their learning process. Its design is based on the assumption that the supply of appropriate information through compensatory sensory channels may contribute to science learning among blind students (Passini & Proulx, 1988). L2C is a representational form designed to overcome the lack of visual information, by providing dynamic auditory feedback to models of complex systems. Sonification, the use of non-speech audio to convey information (Kramer, 1994), is used to represent micro-level individuals, their properties and related events in space (via stereo), and macro-level descriptors of the system. Over the years, several auditory technologies have been developed for blind people (e.g. vOICe, Meijer, 1992; Lahav et al., 2008). The L2C models enables users to interact with dynamic objects that are computed in real time, providing a heightened sense of reality while learning about complex scientific phenomena. It is unique with respect to state-of-the-art learning technologies for blind people, in its sonified representation of *dynamic complex* systems, providing access to quickly changing information of both the micro- and macro-levels in a system. Important to future dissemination, it is a low-cost technology based on the robust and continually developing free NetLogo platform (Wilensky, 1999).

In this project, we addressed the learning of Kinetic Molecular Theory and the Gas Laws in middle-school chemistry, topics fundamental to the understanding of many advanced concepts in the physical sciences (NGSS, 2013). The learning environment is based on two earlier successful curricula (Levy & Wilensky, 2009ab; Samon & Levy, 2013). It consists of eight activities and includes guided exploration of agent-based NetLogo computer models, laboratory demonstrations and class discussions. The models represent a gas in a container in the form of particles (points) located within a rectangle (representing a vessel). The first models introduce students to the concept of a scientific model and to the use of NetLogo. With each additional model students explore new

concepts, relationships and phenomena, which allows them to discover a different characteristic of the nature of gases. A workbook for the use of the students accompanies the unit. For most of the unit, students can progress at their own rate with the workbook that provides guiding information and questions.

The models were adapted for blind students by sonifying events and properties of a single particle such as collisions and speed, as well as global variables such as pressure. It is important to note that the information provided in the visual model is not identical that delivered in the auditory model. Individual events and properties are highlighted to a greater extent in the auditory mode. Global information such as spatial patterns of multiple individuals is provided in the visual but not the auditory model. Finally, in the auditory mode, information is provided in real time, but cannot be accessed later, as in a visual graph. The curriculum was available to the participants as text-to-speech files and in Braille. Figures were presented as tactile images.

In previous studies (Levy & Lahav, 2011) we had investigated the viability of such an approach with one model and one activity with a small sample of blind adults and found learning gains for some but not all science and systems concepts. Later studies (Lahav et al., 2014) have investigated various properties of sound that would optimally convey the most information in a way convivial for our target audience. In the present study, we examine whether such perceptual compensation creates a comparable learning environment for blind and sighted students. We compare learning outcomes and processes for a set of curricular materials in an auditory mode among blind students and in a visual model among sighted students. The research expands knowledge about how the auditory channel may compensate and complement for the visual channel among blind individuals for learning complex systems.

Theoretical perspectives

One of the challenges of gaining a well-developed understanding is that chemical systems can be described in at least three different modes: an invisible molecular submicroscopic level, an experienced macroscopic level and symbolic representations (Johnstone, 1993). Research has shown that students often lack deep understanding of all three modes and often fail to coordinate between them (Novick & Nussbaum, 1981; Gilbert & Boulter, 2000).

Complex systems approaches have come into the limelight in several different domains of science and in education and are based on the following idea: a system can be modeled as many entities that operate according to a small set of simple rules, their concurrent actions and interactions emerging into global patterns (Bar-Yam, 1997). This research aligns with the recently published US framework for science education (NGSS, 2013) that underscores system models as one of the central crosscutting concepts. A complex systems approach to teaching chemistry has been shown to help students overcome these obstacles (Levy & Wilensky, 2009ab; Holbert & Wilensky, 2014). One way of introducing students to dynamic complex systems is by means of agent-based modeling (ABM) in which a computer model simulates the many autonomous, interacting entities (agents) of the system.

We assess both conceptual learning in science and reasoning about complex systems among sighted and blind students with guided exploration of an ABM learning environment in which micro- and macro-level variables and events are either visual or sonified. Our research question is: How do conceptual learning, systems reasoning and learning processes with sonified feedback models for blind students compare with such learning through a visual feedback model for sighted pupils?

Methodological approach

To investigate the comparative learning of blind and sighted students a two-group pre-test-intervention-post-test quasi-experimental design was used. The process of learning is examined through analyses of the students' writings during four periods interspersed throughout the activity span.

Participants

The experimental group consisted of ten students, who are totally blind with no additional impairments (cognitive, auditory), their ages ranging 17-33, high school and university students, as well as two teachers. They all had experience in using computers and had studied science in junior- and high-school. The comparison group consisted of 31 seventh grade sighted students who attend a school serving a high socioeconomic population. Students were distributed approximately evenly between genders. Seven students' data was excluded (completed less than 18/24 questionnaire items, workbook was largely empty). The differences between the two groups results from the scarcity of individuals who are blind and recruitment issues.

Data sources

Data collected included students' responses to pre-post questionnaires and workbook items.

The questionnaires included three open items and 22 closed items, and were developed to align with the curriculum. To do so, the items in the workbook were coded for scientific content, systems thinking and reasoning types (declarative, procedural, schematic or strategic, Shavelson et al., 2003). Questionnaire items were designed in similar proportions. Most items were validated and used in previous research (Levy & Wilensky, 2009b; Samon & Levy, 2013) and printed in four differently-ordered versions.

The workbooks included 229 items, out of which four groups of items were selected as windows for analyzing students' ideas as they change throughout the learning process.

Procedure

Blind students worked and were observed individually. Each session lasted 60 minutes, and the research consisted of ten sessions that were distributed over 5-8 weeks.

Sighted students worked in their school's computer lab, 1-2 students to a computer, during four double-periods over two weeks. Each student had a workbook. The teacher and researchers conducted relatively few conversations, mainly to support students' understanding of the workbook instructions.

Both groups completed identical pre- and post-test questionnaires. Data collection was conducted differently for the two groups. While blind students worked individually with a researcher, sighted students worked individually or in pairs within a classroom setting. The choice results from the study's main goal to compare the blind students' learning with learning in normal school settings, and the scarcity of blind individuals.

Analysis schemes

Overall learning gains: An overall score for each pre- and post-questionnaire was calculated by awarding 1 point for each correct answer. From this data overall scores for each student were calculated. The aggregated scores for sighted and blind students were compared.

Learning of specific scientific concepts and systems reasoning: Questionnaire items were grouped according to the scientific concepts and central systems ideas. Scores were computed for individual students and were then aggregated and compared. In terms of scientific concepts, students' answers were coded for 26 concepts and principles that relate to the particulate nature of gas, various particle interactions and behaviors, macroscopic properties such as pressure and temperature, relations between these properties as seen in the Gas Laws and processes, such as diffusion. With respect to systems reasoning, the answers were coded for ideas regarding levels (micro, macro), emergence, interactions among micro-level entities, control, randomness and dynamic equilibrium.

Learning process: In-depth analysis of four groups of items or 'windows' in the workbook provided information about the learning process. Answers were analyzed according to a coding scheme for both scientific content as well as systems reasoning as described above for the science concepts and systems reasoning in the questionnaires.

Validity

The construct and criterion validity was determined by having five experienced science teachers and two science education experts the content knowledge questionnaires. All confirmed that the test items were appropriate for examining the issues studied in the two learning environments. Coding was checked by the researchers and by an external expert in science education. Agreement was high and disagreements were resolved uniformly by discussion.

Findings

Overall learning gains

Group scores rose significantly from pre- to post-tests (Table 3). Blind students started out with a similar score to that of the sighted students, but their post-test score is significantly higher (Mann-Whitney $U=36.0$, $p<.01$). Learning gains are comparable, though insignificantly higher for the blind students.

Science conceptual learning and systems reasoning

With respect to conceptual learning in science, the sighted students mainly learned the Gas Laws, KMT and density to a certain degree and regressed with respect to understanding diffusion. Blind students' learning gains were higher than those of sighted students. Quite distinct from the sighted students, their largest learning gain relates to diffusion. Distinct from the sighted students the Gas Laws showed least improvement, this, however, may be a ceiling effect, as the score in the post-test is 85%. Regarding systems reasoning, the groups are similar

in the pre-test and in the post-test.

Learning processes

Processes of learning along the four windows for both groups are presented in Appendix A. Comparing these temporal curves shows that for most concepts, the shapes of the curves are similar for the two groups. This is not true for including interactions, uncertainty and decentralized control, where the blind students provided more such explanations earlier in the learning process. Another result from the comparison is the higher proportion of students expressing each of the concepts (systems, science) among the blind group with respect to the sighted group.

Table 3: Pre-test and Post-test Scores (%) and Learning Gain Comparisons for the Blind and Sighted Students, Overall, by Concept

Learning Concepts (# of items)	Blind Students		Sighted Students		Learning Comparison	Gain
	<i>Pre-test M (SD)</i>	<i>Post-test M (SD)</i>	<i>Pre-test M (SD)</i>	<i>Post-test M (SD)</i>	<i>Mann-Whitney U</i>	
Overall (22)	58 (12)	83 (10)	53 (17)	64 (18)	90.0	
Science concepts						
KMT (10)	48 (19)	78 (14)	50 (21)	60 (21)	177*	
Diffusion (4)	45 (20)	80 (16)	53 (32)	47 (39)	145**	
Density (3)	47 (28)	77 (16)	56 (35)	72 (25)	138	
Gas laws (9)	68 (13)	85 (7)	55 (19)	72 (20)	112	
Systems thinking						
Micro-level (7)	49 (22)	87 (16)	43 (22)	58 (24)	158	
Macro-level (7)	70 (8)	93 (14)	57 (21)	64 (21)	160	
Emergence (8)	56 (17)	70 (9)	57 (21)	70 (20)	107	

Conclusions and implications

The study compared blind students' use of auditory computer models to sighted students use of visual models. With respect to sighted students, the auditory representation not only supported blind students' learning similarly, but was even related to greater learning of diffusion, a challenging concept. In analyzing the progressions we can see that the individual science and systems concepts were learned at similar times, corresponding to the learning materials. However, regarding two central systems concepts: interactions between individuals and uncertainty, the blind students learned and applied these concepts earlier.

Earlier, we have compared the visual and auditory models, noting the greater salience of micro-level individuals. The L2C's auditory representation seems to have increased sensitivity to micro-level interactions (e.g. colliding) in a way that is less accessible in visual representations. We had chosen to sonify particular events for one particle and time progressions of global variables such as pressure. In fact, much information in the visual array is missing in the auditory array: many other particles, each moving about, colliding and bouncing off walls. It would seem that this filtering of information helps students focus on these very interactions that are subtle in an array of many particles, and notice their random changes over time, evidence from which uncertainty can be derived.

Limitations to this study include the different character and treatments of the two groups: the experimental group included young adults, who interacted one-on-one with a researcher, while the comparison group included middle-school students who worked in pairs within a classroom. Such limitations were inevitable due to a limited sized blind population.

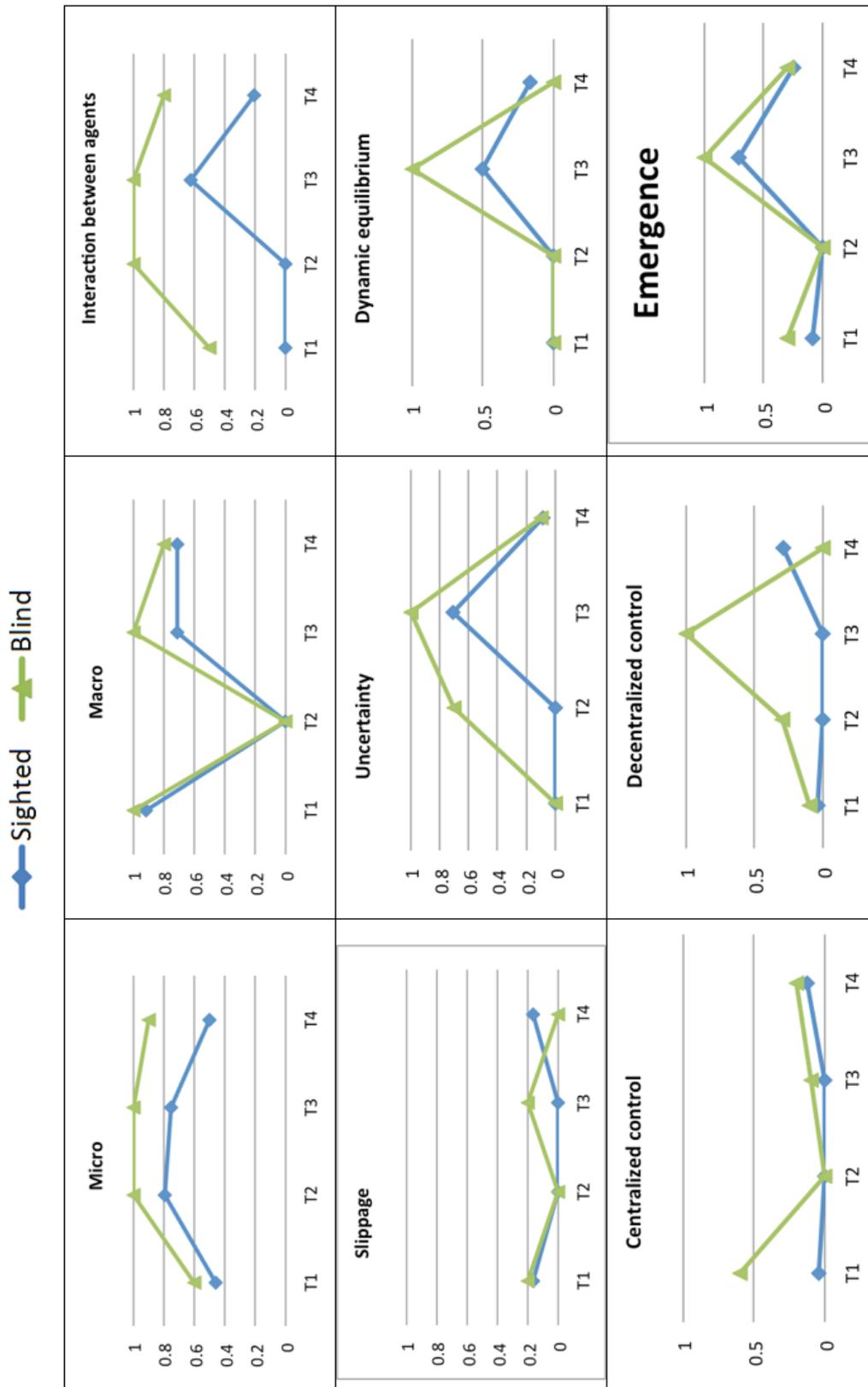
Given the success of this low-cost learning environment, extending this design to learning of other STEM systems by using sonified models opens the way to equitable participation of blind people. It also raises questions for further research into the use of auditory representations to facilitate learning among sighted learners.

References

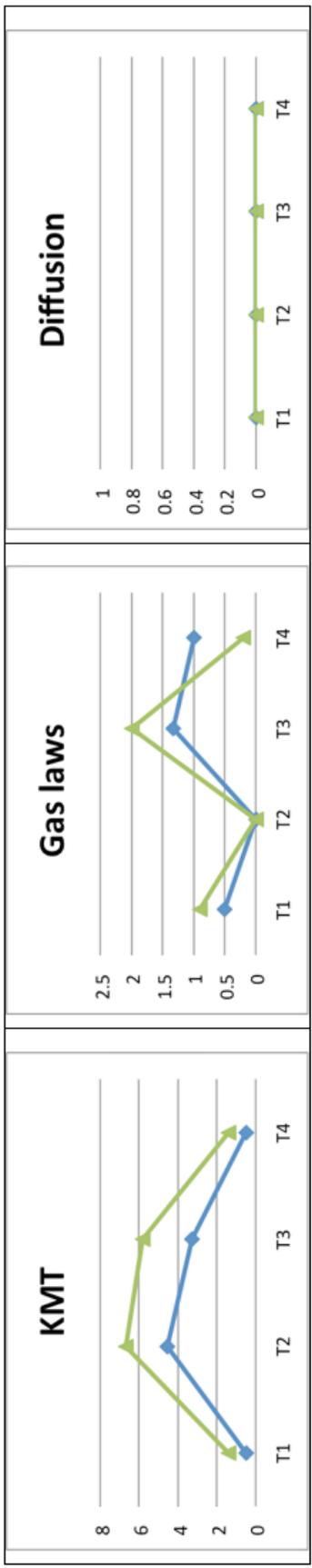
- Bar-Yam, Y. (1997). *Dynamics of Complex Systems*. Reading, MA: The Advanced Book Program.
- Beck-Winchatz, B., & Riccobono, M. A. (2008). Advancing participation of blind students in Science, Technology, Engineering, and Math. *Advances in Space Research*, 42, 1855–1858.
- Connell, B. R., Jones, M., Mace, R., Mueller, J., Mullick, A., Ostroff, E., et al. (1997). *The Principles of Universal Design*, Version 2.0. NC State University, Raleigh, NC: The Center for Universal Design.
- Farrell, B., Baldyga, D. D., & Erlandson, R. (2001). Force feedback mouse used for a physics experiment. Mansfield, CT. In J. D. Enderle & B. Hallowell (Eds.), *NSF 2001 Report: Engineering Senior Design Projects to Aid Persons with Disabilities*. Wayne State University, Creative Learning Press, Inc.
- Gilbert, J.K., and Boulter, C.J. (1998). Learning science through models and modeling. In Fraser, B.J. & Tobin, K.G. (Eds.), *International Handbook of Science Education* (pp. 53-66). Great Britain: Kluwer Academic Publishers.
- Holbert, N., & Wilensky, U. (2014). Constructible authentic representations: Designing video games that enable players to utilize knowledge developed in-game to reason About Science. *Technology, Knowledge and Learning*, 19(1-2), 53–79. doi:10.1007/s10758-014-9214-8
- Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education*, 70(9), 701–705.
- Kramer, G., (1994). An Introduction to auditory display. In: Kramer, G. (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings vol. XVIII. Reading, MA: Addison-Wesley.
- Kumar, D., Ramassamy, R., and Stefanich, G. (2001). Science for students with visual impairments: Teaching suggestions and policy implications for secondary learners. *Electronic Journal of Science Education*, 5(3), 1-9. Retrieved July 27, 2010, from <http://wolfweb.unr.edu/homepage/crowther/ejse/kumar2etal.html>
- Lahav, O. & Levy, S.T. (2011). Listening to Complexity: Blind people’s learning about gas particles through a sonified model. *International Journal on Disability and Human Development*, 10(1), 87-91.
- Lahav, O., Kittany, J., Levy, S.T., & Furst, M. (2014). Perception of Multi-varied Sound Patterns of Sonified Representations of Complex Systems by People who are Blind. Paper presented at 10th International Conference Series on Disability, Virtual Reality and Associated Technologies, Gothenburg, Sweden.
- Lahav, O., Schloerb, D., Kummur, S. & Srinivasan, M.A. (2008) BlindAid: a learning environment for enabling people who are blind to explore and navigate through unknown real spaces. Paper presented at the *Virtual Rehabilitation 2008* International Conference, Vancouver.
- Levy, S.T., & Lahav, O. (2011). Enabling blind people to experience science inquiry learning through sound-based mediation. *Journal of Computer Assisted Learning*, 28(6), 499-513.
- Levy, S.T., & Wilensky, U. (2009a). Crossing levels and representations: The Connected Chemistry (CC1) curriculum. *Journal of Science Education and Technology*, 18(3), 224-242.
- Levy, S.T., & Wilensky, U. (2009b). Students’ Learning with the Connected Chemistry (CC1) Curriculum:Navigating the Complexities of the Particulate World. *Journal of Science Education and Technology*, 18(3), 243-254.
- McGuire, J. M., Scott, S. S., & Shaw, S. F. (2006). Universal design and its applications in educational environments. *Remedial and Special Education*, 3, 166-175.
- Meijer, P.B. (1992). An experimental system for auditory image representations. *IEEE Trans Biomed Eng*, 39, 112-121
- NGSS. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC.
- Novick, S., & Nussbaum, J. (1981). Pupils’ understanding of the particulate nature of matter: A cross-age study. *Science Education*, 65(2), 187–196.
- Nussbaum, J. (1985). The Particulate Nature of Matter in the Gaseous Phase. In R. Driver, E. Guesne, & A. Tiberghien (Eds.), *Children’s Ideas in Science*. Milton Keynes: Open University Press.
- Passini, R., and Proulx, G. (1988). Wayfinding without vision: An experiment with congenitally blind people. *Environment and Behavior*, 20, 227-252.
- Samon, S., & Levy, S.T. (2013). What Does a Complex Systems Perspective Offer Science Learning? Junior High-school Students’ Learning of Chemical Systems with an Agent-based Viewpoint versus a Disciplinary Viewpoint? The Annual meeting of the *American Educational Research Association*, April 27-May 1, 2013, San Francisco, USA.
- Shavelson, R., Ruiz-Primo, M., Li, M., & Ayala, C. (2003). Evaluating New Approaches to Assessing Learning. *CSE Report.*, 1522(310).

- Wies, E. F., Gardner, J. A., O'Modhrain, M. S., Hasser, C. J., & Bulatov, V. L. (2001). Web-based touch display for accessible science education. In S. Brewster & R. Murray-Smith (Eds.), *Proceedings of Haptic HCI 2000* (pp. 52–60).
- Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. Retrieved July 27, 2010, from <http://ccl.northwestern.edu/netlogo>.
- Willoughby, D., and Duffy, S. (1989). Handbook for itinerant and resource teachers of blind and visually impaired students. Baltimore, MD: National Federation of the Blind.
- Zaborowski, B. (2006). Research issues from the consumer perspective [Message 1]. Retrieved July 27, 2010, from http://ecampus.pco.edu/webapps/portal/frameset.jsp?tab=course&url=%2Fbin%2Fvommon%2Fcourse.pl%3Fcourse_id%3D_355_1

Appendix A: Learning Progressions for Systems Reasoning and Conceptual Learning in Workbook Activity Windows



◆ Sighted ▲ Blind



Detecting Collaborative Dynamics Using Mobile Eye-Trackers

Bertrand Schneider, Stanford University, schneibe@stanford.edu
Kshitij Sharma, EPFL, kshitij.sharma@epfl.ch
Sébastien Cuendet, EPFL, sebastien.cuendet@epfl.ch
Guillaume Zufferey, EPFL, guillaume.zufferey@epfl.ch
Pierre Dillenbourg, EPFL, pierre.dillenbourg@epfl.ch
Roy Pea, Stanford University, roypea@stanford.edu

Abstract: Prior work has successfully described how low and high-performing dyads of students differ in terms of their visual synchronization (e.g., Barron, 2000; Jermann, Mullins, Nuessli & Dillenbourg, 2011). But there is far less work analyzing the diversity of ways that successful groups of students use to achieve visual coordination. The goal of this paper is to illustrate how well-coordinated groups establish and sustain joint visual attention by unpacking their different strategies and behaviors. Our data was collected in a dual eye-tracking setup where dyads of students (N=54) had to interact with a Tangible User Interface (TUI). We selected two groups of students displaying high levels of joint visual attention and compared them using cross-recurrence graphs displaying moments of joint attention from the eye-tracking data, speech data, and by qualitatively analyzing videos generated for that purpose. We found that greater insights can be found by augmenting cross-recurrence graphs with spatial and verbal data, and that high levels of joint visual attention can hide a free-rider effect (Salomon & Globerson, 1989). We conclude by discussing implications for automatically analyzing students' interactions using dual eye-trackers.

Keywords: joint visual attention, collaborative learning, dual eye-trackers

Introduction

Joint Visual Attention (JVA) is a central construct for any learning scientist interested in collaborative learning. It is defined as "the tendency for social partners to focus on a common reference and to monitor one another's attention to an outside entity, such as an object, person, or event" (Tomasello, 1995, pp. 86-87). Without joint attention, groups are unlikely to establish a common ground, take the perspective of their peers, build on their ideas, express some empathy or solve a problem together. As an example, autistic children are known to have difficulties in coordinating their attention with their caregivers (Mundy, Sigman, & Kasari, 1990), which in turn is associated with social disabilities. Joint visual attention is a prerequisite for virtually all of social interactions.

Previous work has highlighted the importance of joint attention in small groups of students: Barron (2000), for instance, carefully analyzed how groups of students who failed to achieve joint attention were more likely to ignore correct proposals and not perform as well as similar groups. Richardson & Dale (2005) found that the degree of gaze recurrence between individual speaker–listener dyads (i.e., the proportion of times that their gazes are aligned) was correlated with the listeners' accuracy on comprehension questions. Jermann, Mullins, Nuessli and Dillenbourg (2011) conducted a similar analysis with a dual eye-tracking setup. They used cross-recurrence graphs (Fig. 1) to separate "good" and "bad" collaborative learning groups. They found that productive groups exhibited a specific pattern (shown in Figure 1's middle graphs), and that less productive groups produced a fundamentally different collaborative footprint (right graph of Figure 1). In a different study, Mason, Pluchino and Tornatora (2015) showed that 7th graders who could see a teacher's gaze when reading an illustrated text learn significantly more than students who could not. Schneider & Pea (2013), using two synchronized eye-trackers in remote collaborations, showed that dyads who could see the gaze of their partner in real time on a computer screen outperformed their peers in terms of their learning gains and quality of collaboration. This intervention was highly beneficial to students since they could monitor the visual activity of their partner, anticipate their contribution and more easily disambiguate vague utterances. There are many more examples of studies finding joint visual attention to be associated with interactions of higher quality. Due to space constraints we will not conduct an exhaustive review of this line of research. The take away message is that JVA plays a fundamental role in organizing social interactions, and measuring its frequency using a dual eye-tracking setup can serve as a proxy for evaluating a group's collaboration quality.

This line of research has received renewed attention with the advent of mature eye-tracking systems. The main focus of this paper is 1) to use dual eye-tracking data to understand how well-functioning dyads of students

interact; 2) to use cross-recurrence graphs to guide qualitative analyses; and 3) a foray into studying the variety of ways that students use to maintain high levels of joint visual attention.

Dual eye-tracking settings and cross-recurrence graphs

Previous work done by Jermann and his colleagues used *cross-recurrence* graphs to make sense of dual eye-tracking datasets. A cross recurrence graph (shown on Fig. 1) is a visual representation of a dyad’s visual coordination: the x-axis represents time for the first participant, the y-axis represents time for the second participant, and black pixels show moments of joint attention (for a given time slice and distance between two gazes). Thus, a black diagonal signifies that the dyad was continuously looking at the same area of interest (e.g., the middle graph in Fig. 1); the absence of a dark diagonal means an absence of joint visual attention (e.g., the right graph on Fig. 1). In other words, good quality interaction exhibits higher recurrence rates of JVA compared to low quality interaction, and this difference can be visually detected using cross-recurrence graphs.



Figure 1. Reproduced from Jermann, Mullins, Nuessli & Dillenbourg (2011): on the left, a schematic cross-recurrence graph; in the middle picture, a cross-recurrence graph from a productive group; on the right side, a graph from a less productive group.

We use a similar methodology in this paper, with three new contributions: first, the data comes from a study that looked at co-located interactions. We captured students’ gazes using two mobile eye-trackers, whereas prior work has solely looked at remote interactions where students were looking at two different computer screens. This methodological development is a significant improvement in ecological validity, because so much of collaborative learning is among co-located individuals. Second, we augmented cross-recurrence graphs with speech data and spatial information to uncover students’ strategies when working on a problem-solving task. This provided us with compelling visualizations that guided our qualitative analysis. Third, we contrasted two groups that each exhibited high levels of joint visual attention (in contrast with comparing a productive versus a non-productive group, as Jermann and his colleague did). Our goal is to illustrate the multitude of ways that students use to successfully establish joint visual attention. We found that dual eye-tracking datasets can uncover particular types of collaborations.

In the next sections, we summarize the methods we used to collect our data. We then describe our two dyads of interest, and compare their cross-recurrence graphs. Finally, we provide quotes to illustrate crucial differences in the ways that those two dyads collaborated. We conclude with thoughts on the implications of this work, and future directions for studying joint visual attention in small collaborative learning groups.

Methods

Experimental design, subjects and material

The goal of our previous study was to conduct an empirical evaluation of a Tangible User Interface designed for students following vocational training in logistics. Our goal was to compare the affordances of 2D, abstract-looking interfaces (Fig. 2, left side) with 3D, realistic-looking interfaces (Fig. 2, right side) The system used in this study, the TinkerLamp, is shown on Figure 2 (bottom row): it features small-scale shelves that students can manipulate to design a warehouse layout. A camera detects their location using fiducial markers and a projector enhances the simulation with an augmented reality layer. Fifty-four apprentices participated in the study (28 in the “3D” condition, mean age = 19.07, SD = 2.76; 26 in the “2D” condition, mean age = 17.96, SD = 1.56). The activity lasted an hour and the goal for students was to uncover good design principles for organizing warehouses. The core of this reflection involves understanding the trade-off between the amount of merchandise that can be

stored in a warehouse and how quickly one can bring an item to a loading dock (i.e., a lot of shelves make it difficult to efficiently load/unload items, while few shelves limit the size of available stock). To help students understand this relationship, we followed the Preparing for Future Learning framework (Bransford & Schwartz, 1999). We designed a set of contrasting cases (Fig. 2, first row) and asked students to analyze three layouts based on the following criteria: in which warehouse they would prefer to work on a daily basis (prompt 1); which warehouse maximized space (prompt 2); and finally, which warehouse minimized the distance from each shelf to the docks (prompt 3). At the end, the experimenter revealed the numbers for those two metrics (referred to as the “reveal” on Fig. 3). The contrasting cases were set up to highlight good design principles (e.g., orienting the shelves toward the docks makes them more accessible, placing them back to back so that only one side is accessible frees up space, and so on). Students were supposed to use those principles in the next activity (Fig. 2, second row), where we asked them to build two warehouses: one, where they had to put as many shelves as possible in a given layout, and a second one where they had to minimize the average distance to the docks given a certain number of shelves. Before and after those two activities, we gave students a pre- and post-test on paper where we asked them to slightly change the layout of several warehouses in order to optimize them. Thus we have two main dependent variables: how well they designed the two warehouses, and how well they answered the learning test. During those two activities, students wore two mobile eye-trackers (SMI Eye-Tracking Glasses with binocular pupil tracking at 30Hz) and we used the scene cameras (1280x960 pixels) to record videos for post-processing. In a previous publication, we explain in detail how we used the fiducial trackers detected by the scene camera to remap the dyad’s gazes onto a ground truth and determine whether two students were jointly looking at the same location at the same time. We will not delve deeper in the general results of this study here; the interested reader can learn more about it in Schneider et al. (2015). In the next section, we conduct a more in-depth analysis of two dyads of students who achieved high levels of visual coordination, as measured by the mobile eye-trackers.

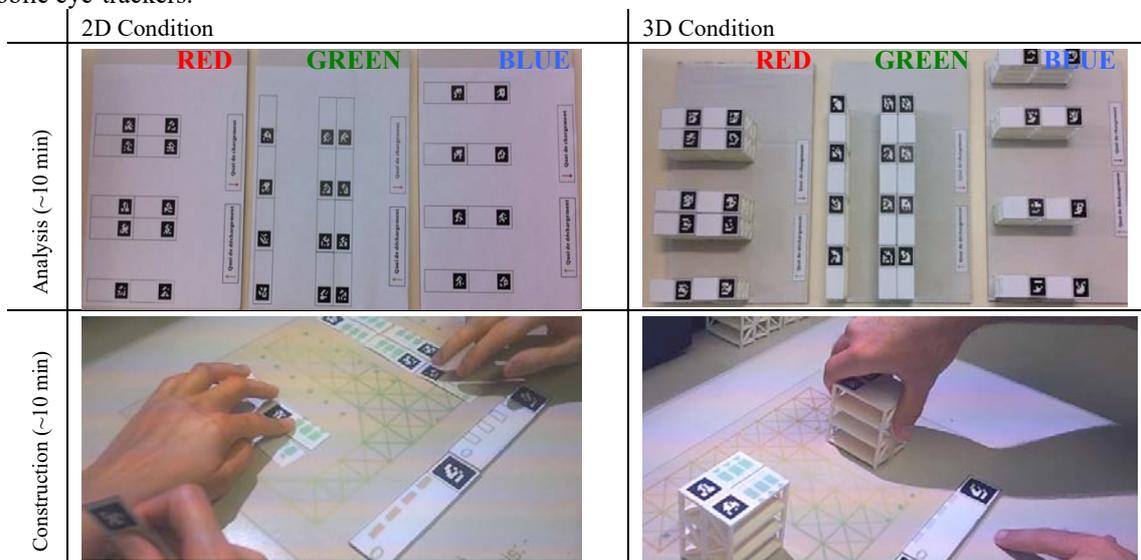


Figure 2. The two experimental tasks of interest: 1st row the contrasting cases students had to analyze; 2nd row shows the Tangible Interface for the construction task. The “red”, “green”, “blue” labels are used to color-code additional analyses below (e.g., in the cross-recurrence graphs in Fig. 3)

Analyses

We chose to focus on groups 13 and 20 because they had the highest levels of joint visual attention. They also performed above the average on the problem-solving tasks (but not on the learning test). Table 2 summarizes key information about the two dyads:

Table 2: Data of group 13 and 20.

	Task 1	Task 2	Learning gains	JVA	Speech
Group 13	13 shelves	7.47 meters	7.5 points	24 %	409 seconds
Group 20	18 shelves	7.39 meters	-0.5 points	29%	540 seconds

Total	M=13.5 SD=2.9	M=7.6 SD=1.9	M=1.3 SD=3.2	M=15.7 SD=8.0	M=405.7 SD=245.7
--------------	------------------	-----------------	--------------	------------------	------------------

As a reminder, the goal of Task 1 was to place as many accessible shelves as possible on a given layout. The goal of Task 2 was to minimize the average distance from each shelf to the docks, given a certain number of shelves. We computed learning gains by subtracting pre-test scores from the post-test scores, and then averaged them at the group level. After running the study, we found that the post-test was harder than the pre-test. This explains why some learning gains are negative (e.g., group 20 in Table 1). JVA scores were computed by counting how many times two gazes were in a radius of 70 pixels (i.e., the width of a shelf) within +/- 2 seconds of lag (Richardson & Dale, 2005); we then divided this count by the total number of data points in the eye-tracking datasets. This gave us a number that reflects the percentage of time that two students were jointly looking at the same location. Finally, we extracted speech information from the audio data to collect the number of seconds that each student spoke. JVA and speech data are displayed in Fig. 3 for groups 13 and 20: on the top, we first show the traditional black and white cross-recurrence graphs used in prior work (e.g., Jermann, Mullins, Nuessli & Dillenbourg, 2011); on the bottom, we show our cross-recurrence graphs augmented with spatial information (i.e., which warehouse the two students were jointly looking at) and speech data (who spoke when during the analysis task).

Augmenting cross-recurrence graphs

The graphs of Fig. 3 provide us with several interesting bits of information about groups 13 and 20. First, we can notice that the traditional cross-recurrence graphs are ideal for identifying moments of joint visual attention (black squares along the diagonal). Based on those graphs, one would predict that group 20 has a better visual coordination than group 13: the diagonal is darker, with more well-defined squares, meaning that this group jointly looked at the same area on the table more often and had longer moments of joint attention. Prior work suggests that an unfilled cross recurrence graph is likely to indicate a poor collaboration between group members. However, we would like to demonstrate that a “good” cross-recurrence graph (i.e., with a dark diagonal) is not necessarily indicative of a good learning group. We will illustrate this difference by exposing some of the diversity that exists between groups of highly visually-coordinated students.

Our augmented cross-recurrence graphs (middle section of Fig. 3) color-code each pixel to show which warehouse groups 13 and 20 were looking at (red pixels represent JVA on the 1st warehouse, green on the second one, and blue on the last one). We also added dotted squares to show when the experimenter gave various prompts to the groups (i.e., prompt 1 = “in which warehouse would you prefer to work”, prompt 2 = “which warehouse maximizes space, and why”, prompt 3 = “which warehouse minimizes the average distance from each shelf to the docks and why”, and reveal = “now I will show you numbers that answer those two questions to verify if your intuition was correct”). Two very different strategies appear: group 20 analyzed the three warehouses in a serial manner. For instance, after receiving prompt 2, they jointly looked at the first warehouse, then the second one, and finally at the last layout. Group 13, in comparison, continually compared the three warehouses. We do not have empirical evidence that one strategy is superior to the other, but common sense and the PFL framework would suggest that multiple comparisons between contrasting cases would increase students’ chances to notice important features between those layouts. One indication that this strategy might be more beneficial to learning is that group 13 achieved higher learning gains on the test (indeed, they achieve the highest learning gains in our experiment). We plan to quantify this behavior in future work to see whether it positively correlates with learning across all groups who participated in this study.

Coming back to groups 13 and 20, there is one last piece of information worth mentioning. We know from Table 1 that group 20 talked slightly more than group 13 (540 vs 409 sec.). On the bottom of Fig. 3, we show participants’ speech data. Observe that group 20 has more imbalanced participation (participant 39 talked for 117 sec. vs 423 sec. for participant 40) compared to group 13 (in which participant 25 talked for 170 sec. vs 239 sec. for participant 26); thus, one group 20 student talked 78% of the time whereas one group 13 student talked 58% of the time. Might these differences be symptomatic of collaborative learning issues?

To illuminate those results comparing groups 13 and 20, we qualitatively examined the videos of the experiment. It should be noted that those differences (strategy used, learning gains and speech distribution) are already a striking contrast to the supposedly superior (black and white) cross-recurrence graph of group 20.

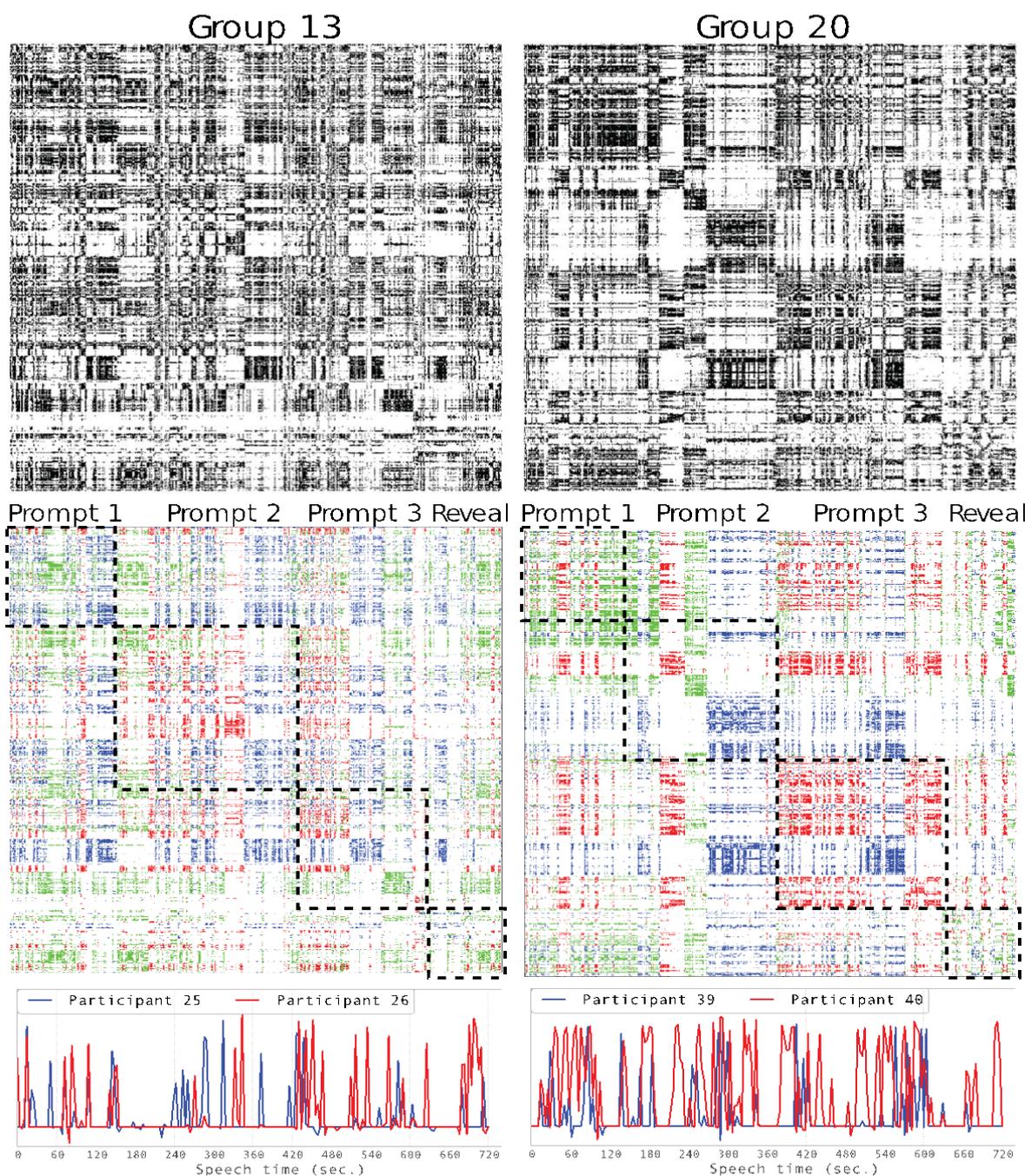


Figure 3: Cross recurrence graphs of two high-performing groups (13 and 20). The top figures show standard cross-recurrence graphs. The middle figures show colored graphs (red pixel = JVA on the 1st warehouse; green = JVA on the 2nd one; blue = JVA on the 3rd one). The bottom figures show speech data for each dyad.

Qualitative analysis

We further compared our two groups by analyzing videos of their interactions. We generated videos by combining the scene cameras of the students (top left section of Fig 4) and remapping their gazes onto a ground truth during task 2 (bottom left section of Fig. 4). We also added an animation of the cross-recurrence graph on the right side of Fig. 4, showing the graph up to that particular video frame. This video enabled a multi-modal analysis of the students' interactions, in particular by highlighting how they combined gestures and speech to achieve and sustain joint attention. Groups #13 and #20 were extremely similar in that regard, constantly using pointing gestures to ground their verbal contributions. It is reasonable to believe that those deictic gestures played a large role in increasing their levels of joint visual attention.

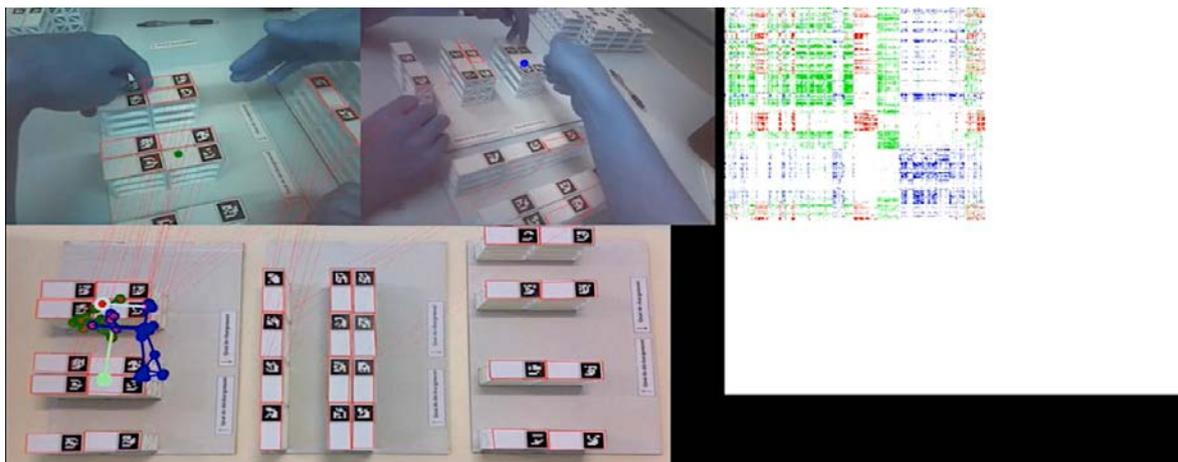


Figure 4: A video frame generated for the qualitative analysis. On the top left, we show the perspectives of the students with their gazes in blue and green (red dots show moments of JVA). On the bottom left, we remapped their gazes onto a ground truth. On the right side, we show the cross-recurrence graph up to that frame.

In the table below, we focus on the major differences that we observed for the two groups. Specifically, we focused on one behavior found to be an important predictor of a group's success: how peers react to a proposal (Barron, 2000). A proposal can be accepted, rejected, challenged, or ignored. One key difference between groups 13 and 20 was that one dyad was more likely to uncritically accept proposals, while the other was much more likely to challenge them (keywords highlighting this difference are marked in bold in Table 2):

Table 2: Excerpts of group 13 and 20's dialogues. (e.g., E=experimenter, 25 = participant 25)

Group 13	Group 20
<p><u>08:24 --> 09:38</u> E: The first question is, in which warehouse would you rather work, and why? 25: This one! 26: Yeah this one! 25: if you look at this one, you have less palettes than that one 26 Because of the width here... 25: Not really! 26: yes, it's wider 25: but wait, here that's 4 shelves 26: Here you can't get it out 25: but you can't get from behind 26: yes, that's what I'm saying [...] 25: 1,2,3,4 you're crazy. You have more space here.</p>	<p><u>00:08 --> 01:30</u> E: The first question is, in which warehouse would you prefer to work, and why? 40: I think this one is good (2nd warehouse), because you can use half of the warehouse for loading and the other half for unloading 39: yes but you can go faster in this one (1st warehouse) 40: yes because in this one you only have space for two palettes in front of the docks. It's a little tight. I think that I would prefer to work in this one (2nd warehouse). Maybe I would like this one too 39: there's enough space between the shelves 40: yes, and in this one [pointing at the 1st warehouse] the shelves are too tight</p>
<p><u>12:40 --> 13:52</u> E: if you could change something in each warehouse, how would you improve them? 26: Hmmm so... 25: I would add two shelves, like that, there. Otherwise... 26: What if we put those like that... to add more shelves 25: No; you can add two here, that's 18 additional palettes 26: Otherwise... one here 25: No, because then if you take a palette from here, you have to back off like that, even if someone's coming from that direction</p>	<p><u>08:12 --> 09:11</u> E: if you could modify those warehouses to minimize the average distance to each shelf, what would you do? 40: What would I do? For instance, in this one (3rd warehouse), I would move those shelves to the corners 39: yes right here 40: This one and that one, I put them here, and those two (in the middle), I would put them there 39: yeah 40: No no, this doesn't make sense. It doesn't change anything. I was thinking, those two you put them there 39: yes 40: but then you're too far away from this shelf 39: well, it's not too bad.</p>

We found that this difference was a recurring pattern in the videos and transcripts of groups 13 and 20. For group 20, participant 40 was extremely verbose and dominated the interaction by using many pointing gestures to illustrate his thought process. This was crucial for the group, because it allowed his peer to maintain joint attention on the warehouse layouts. Participant 39, on other hand, was very passive and very rarely contradicted his partner. His behavior is best described by the “free rider effect” identified by Salomon and Globerson (1989). In summary, the ideas generated by group 20 did not significantly change over the course of this activity because proposals were rarely challenged or contradicted.

Group 13 provides a strong contrast to this dynamic. In this group, we found that participant 26 tended to act like participant 40 (i.e., by driving the interaction and assuming a leadership role). This was mostly manifested by the amount of speech shown on the bottom of Figure 3. But participant 25 was not the kind of free rider that participant 39 was; whenever he did not agree with a proposal, he challenged it until the group reached a consensus. Those challenges were often initiated by using observations made on other layouts, which explains why the group was more likely to have some back and forth on different warehouse layouts (as shown on the colored cross-recurrence graph in Fig. 3). Thus, a continuous refinement of their proposals seemed to be beneficial for their learning, while group 20’s reflection stayed on a more superficial level.

Detecting imbalances of participation in the eye-tracking data

The next question is whether this free-riding behavior pattern is a general one in other groups, and whether it can be detected from the eye-tracking data. We propose one approach here; other measures might provide the same result. We started by identifying, for each moment of joint attention (i.e., each red dot on Fig. 4), which student initiated that episode. In our case, we define the initiator as the person whose gaze was most present in this area during the previous second. While we recognize that this definition is arbitrary, it is arguably a reasonable first step in capturing leadership behaviors. We then computed the proportion of the JVA moments that each student initiated, and then applied this proportion to the percentage of JVA moments achieved by the group. We took the absolute value of the difference between the score of each participant in a group to represent the (im)balance of a group’s “visual leadership”. As an illustration, a group may achieve joint attention during 25% of their time collaborating together; let us say that one student initiated 5% of those moments of JVA, while the other student initiated 20% of those moments. Following the computation above, this group would receive a score of $\text{abs}(0.05 - 0.20) = 0.15$. This measure is shown on the x-axis of Figure 5.

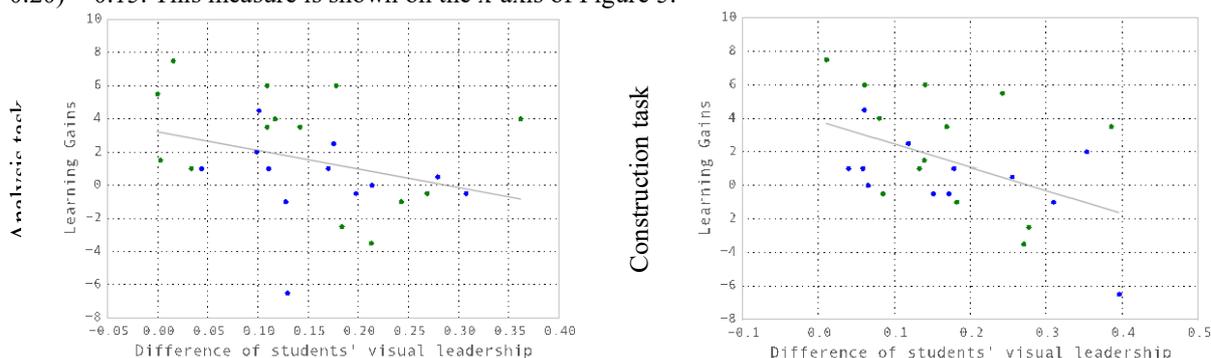


Figure 5: Negative correlation between students’ learning gains and their visual leadership (i.e., the difference between the percentage of moments of JVA initiated by each participant). Left side shows the scatter plot for the analysis task ($r=-0.33$, $p=0.10$) right side shows the scatter plot for the construction task ($r=-0.47$, $p=0.02$).

Green dots are dyads in the 3D condition, blue dots dyads in the 2D condition.

Points on the right side of the graphs (with higher values) represent groups where one person was more likely to initiate more moments of JVA; points on the left side of the graph (with values closer to zero) represent groups where both students were equally likely to either respond or initiate a moment of JVA. The y-axis shows learning gains computed at the group level. We found a negative correlation between dyads’ learning gains and the absolute difference of students’ visual leadership during the analysis task: $r(23)=-0.33$, $p=0.10$ and the construction task: $r(23)=-0.47$, $p=0.02$. It suggests that groups who did not equally share the responsibilities of initiating and responding to offers of JVA were less likely to learn. *This result shows that we could potentially recognize a form of the “free-rider” effect by looking at the eye-tracking data in a collaborative setting.*

Discussion

This paper makes three contributions to the study of collaborative learning groups: first, it analyzes data coming from two mobile eye-trackers capturing co-located interactions. Previous work (e.g., Jermann, Mullins, Nuessli, & Dillenbourg, 2011; Richardson & Dale, 2005) studied remote collaborations where participants were in different physical spaces and collaborated remotely. This makes our contribution more ecological and opens new doors for analyzing face-to-face and side-by-side interactions. Second, we showed that augmenting cross-recurrence graphs with spatial and verbal information provides researchers with new insights regarding students' strategies and interactions: color-coding them suggests whether a dyad is working in a serial or parallel manner when they are analyzing contrasting cases. We hypothesize that the latter strategy is more beneficial to learning, as measured by our pre and post-tests. Furthermore, adding speech data to the graphs was crucial in our analyses. It allowed us to visually detect imbalances in the group's interactions and dig deeper into their discussion. This observation generated the paper's third contribution: we found that highly coordinated dyads (as measured by dual eye-trackers) were not necessarily the best learning groups. Augmented cross-recurrence graphs revealed imbalances in students' verbal contributions, which are characteristic of the free-rider effect where one student does all the work while his/her partner stays passive. This difference was also observed by qualitatively analyzing the videos: we found that the second student was less likely to challenge his peer's proposals, which prevented the dyads from refining their understanding of warehouse management. Finally, we extended those results to the entire sample and found a negative relationship between students' learning gains and their tendency to share the responsibility of initiating and responding to offers of joint visual attention.

Conclusions and implications

The implications of this work are that cross-recurrence graphs are highly valuable for distinguishing between productive and unproductive groups. But they should ideally be complemented with spatial and verbal information to provide a more refined representation of a group's interactions. Past a certain threshold, high levels of joint visual attention make higher learning gains possible but they do not guarantee them. There are multiple ways in which student dyads can establish and maintain joint visual attention, which is a necessary but not sufficient condition for socio-cognitive conflicts. This paper makes a first step in detecting the absence of those conflicts in visually coordinated students by identifying imbalances in students' tendency to initiate or respond to offers of joint visual attention.

References

- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences, 9*(4), 403–436.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of research in education, 24*, 61-100.
- Jermann, P., Mullins, D., Nüssli, M.-A., & Dillenbourg, P. (2011). Collaborative gaze footprints: correlates of interaction quality. *CSCL2011 Conference Proceedings., Volume I - Long Papers*, 184–191.
- Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of text and picture integration during reading: Effects on processing and learning. *Contemporary Educational Psych., 14*, 172-187.
- Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders, 20*(1), 115–128.
- Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cog. Science, 29*(6), 1045–1060.
- Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Educational Research, 13*(1), 89–99.
- Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *Int. Journal of Computer-Supported Collaborative Learning, 8*(4), 375–397.
- Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2015). 3D Tangibles Facilitate Joint Visual Attention in Dyads. *CSCL2015 Conference Proceedings., Volume I*, 158-165.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum.

Acknowledgments

We gratefully acknowledge support from the National Science Foundation for this work from the LIFE Center (NSF #0835854) as well as the Leading House Technologies for Vocation Education, funded by the Swiss State Secretariat for Education, Research and Innovation.

Communication Patterns and Their Role for Conceptual Knowledge Acquisition From Productive Failure

Christian Hartmann, Nikol Rummel, Katharina Loibl
christian.hartmann@rub.de, nikol.rummel@rub.de, katharina.loibl@rub.de
Ruhr-Universität Bochum

Abstract: In *Productive Failure* (PF), students struggle with a problem in small groups before receiving instruction on the correct solution, which has been shown to enhance conceptual knowledge acquisition. Explanations for the effectiveness of PF point at the importance of contrasting the student's solution ideas with critical aspects of the correct solution during instruction. However, other mechanisms of PF have not yet been investigated in detail. Here, we focus on the role of communication patterns during the initial problem-solving phase. Using existing process data from a PF study, we investigated communication patterns based on the ICAP-model and correlated them with prior knowledge, quantity and quality of student solutions, and conceptual knowledge at post-test. Our findings suggest that interactive communication during the initial problem-solving phase only fosters conceptual knowledge if students discuss those ideas that particularly lend themselves to being contrasted with critical aspects of the correct solution during the subsequent instruction phase.

Introduction

Imagine a learning situation where a small group of students has to deal with a mathematical problem to an unfamiliar concept (e.g., variance). Presenting students the canonical (i.e., correct) solution of a specific mathematical problem first and letting them practice afterwards is a common approach to teaching students mathematical concepts. From a constructivist point of view, this instructional design may neglect the potential of students using their prior knowledge and intuitive ideas for fostering a more elaborate understanding of the learning material (cf. Moshman, 1982).

An alternative instructional design for teaching mathematical concepts that goes along with constructivist principles is *Productive Failure* (PF). The central design principle of PF is to ask students to solve a domain-specific problem prior to receiving instruction on that concept (Kapur, 2014). Research on PF suggests that successful conceptual knowledge acquisition is strongly intertwined with the possibility of struggling with a problem before learning how to solve the problem correctly. Struggling with an – at the time of problem solving – unfamiliar problem helps students to activate their prior knowledge and to explore the problem-solving space by generating and contrasting different solutions (Kapur, 2014).

During the problem-solving phase of PF, the students typically generate a certain number of solutions, which are more or less useful for solving the given problem. Because the given problem is based on new learning objectives, the students generally fail to find the correct solution. However, when students try to solve a problem by generating their own ideas, they activate prior knowledge and intuition to create solutions (Kapur, 2014), which can be contrasted to the canonical solution during subsequent instruction. *“One explanation of the better performance of PF students comes by way of prior knowledge activation and differentiation during the problem-solving phase, which may help them better notice and attend to critical features of the concept”* (Kapur, 2014, p. 1018). In other words, struggling with a complex problem before getting the correct solution can 1) activate prior knowledge and intuitive ideas and 2) help students to understand the canonical solution afterwards because they already noticed the limitations of their ideas during problem solving. This awareness can be useful to get a deeper understanding of the topic because students are likely to focus on the relevant aspects of the problem during the instructional phase.

This awareness can also prepare the students for the conceptual knowledge acquisition that happens during the instruction phase even if they were not able to solve the problem correctly (Kapur & Bielaczyc 2012; Loibl & Rummel, 2014). Research has shown consistently that PF promotes conceptual understanding more than presenting students the correct solution of a problem first (e.g., direct instruction) followed by solving practice problems afterwards (Kapur, 2014, Kapur & Bielaczyc 2012; Kapur & Kinzer, 2009; Loibl & Rummel, 2014). So far, research of PF has identified necessary conditions for the instruction phase to be effective. Loibl and Rummel (2014) showed that student ideas that contrast with the canonical solution presented during instruction might be relevant for the effectiveness of PF. Yet it is still unclear what processes during the initial problem-solving phase prepare students best for the following instruction. The fact that the problem-solving phase is usually implemented

in small group collaboration may play an important role here. Different patterns of communication (i.e., discussions) may be more or less suited to foster the possibilities of contrasting different solutions.

Nearly all studies of PF implemented the initial problem-solving phase in small groups (Kapur, 2014; 2014; Loibl & Rummel, 2014). During small group collaboration, students perform activities such as discussing, explaining, or listening to explanations. Ample research (including several meta-analyses) has shown that collaboration has positive effects on conceptual knowledge acquisition (Johnson & Johnson, 2002, 2009; van Boxtel, van der Linden, & Kanselaar, 2000). With regard to PF, Sears (2006) found that students showed more knowledge-sharing behavior when having to solve unfamiliar problems in comparison to students who first received instruction and solved familiar problems afterwards. Thus, struggling with an unfamiliar problem seems to go along with (or require) specific communicative patterns. A quasi-experimental study by Mazziotti, Loibl and Rummel (2015) investigated the impact of collaboration on the effects of PF. However, they did not find differences between an individual and a collaborative PF condition. A reason for their finding could be that the students did not collaborate effectively (Mazziotti et al., 2015). More generally, it is likely that communication can enhance the effectiveness of PF only under certain conditions. Therefore, it is interesting to look closely at communication patterns taking place in a PF setting and to investigate how they affect the effectiveness of PF. Moreover, the importance of the student constructed ideas generated during the problem-solving process have not been investigated. There is a need to investigate two aspects of learning from PF: 1) How relevant are specific communication patterns during problem solving for knowledge acquisition? 2) About what intuitive ideas students have to communicate during problem solving?

In this paper, we are presenting process analyses of data from a typical PF setting. By analyzing different communication patterns during a problem-solving phase, our findings provide the first evidence for answering the questions above. We analyzed the process data of a collaborative problem-solving phase from a PF study (Loibl & Rummel, 2014) by building upon the ICAP-Model by Chi and Wylie (2014) to identify different communication patterns (e.g., passive observation or interactive communication) and evaluated their relation to students' conceptual knowledge acquisition in a PF setting. With this analysis, we want to provide a basis for experimentally testing hypotheses concerning the relationship between collaboration (specifically, communication patterns) and learning in PF. In the following paper, we deal with these issues by building upon the ICAP-Model as a central framework for our analysis.

The ICAP-Model by Chi and Wylie (2014) is a framework that characterizes central learning activities that can emerge during learning situations. The acronym 'ICAP' (**I**nteractive-**C**onstructive-**A**ctive-**P**assive) describes these learning activities. The activities are listed in descending order with reference to their attributed potential to foster learning outcomes. Chi and Wylie (2014) argue that interactive learning activities are most effective and passive activities are the least. It is important to stress that – with the exception of interactive activities (defined as strongly intertwined contributions of at least two individuals) – the described learning activities do not solely focus on social learning situations. Nevertheless, the mentioned learning activities are useful to characterize communication patterns during problem solving in small groups (cf. Deiglmayr, Rummel, & Loibl, 2015). For instance, Chi and Wylie (2014) describe passive activities as the 'passive' perception or observation of information (e.g., watching a video or reading a text). Such an activity can also be performed during communicative learning situations (e.g., listening to an explanation). Active learning activities are defined as 'actively' manipulating or dealing with the learning environment (e.g., coordinating by highlighting an important object) and constructive activities as generating new information by elaborating on prior knowledge (e.g., explaining, arguing for new ideas). In a collaboration setting, students can engage in monologues that fit these definitions. Thus, we argue that ICAP is a manageable way to categorize communication patterns during collaborative problem solving.

Interactive learning activities describe learning situations where at least two individuals co-construct ideas (e.g., discussions). If a large extent of the communication processes within small groups are interactive, this implies that individuals are dealing with each others' ideas during the communication process (cf. Stahl, 2013; Teasley, 1997). In other words, communicative actions are interactive when ideas or thoughts of more than one individual are strongly intertwined. One important criterion of strong (i.e., interactive) collaboration is synchronous communication (Dillenbourg, 1999). Along with most theoretical approaches on collaborative learning (Dillenbourg, 1999; Stahl, 2013), the ICAP-Model points out that interactive activities outperform non-interactive activities concerning conceptual knowledge acquisition (Chi & Wylie, 2014). Therefore, in particular, interactive communication may have the potential to prepare students for conceptual knowledge acquisition during instruction.

During the problem-solving phase of PF, learners struggle with each others' ideas to deal with an unknown problem. Mugny and Doise (1978) argue that struggling with different points of view during strongly-interactive learning situations can lead to socio-cognitive conflict, especially when students have to bring together

diverging aspects or ideas. With reference to Piaget (1977), a socio-cognitive conflict describes a situation where learners try to assimilate their ideas to the ideas of other learners (see Piaget, 1977; Mugny & Doise, 1978). In the case that assimilation fails, students may realize that there is a need for accommodation in their existing concepts because their ideas do not fit to other ideas or are not useful for solving the given problem. Therefore, interactive communication during problem solving may foster knowledge acquisition from PF because socio-cognitive conflicts are a good preparation for the following instruction.

The theory of socio-cognitive conflicts could explain why PF is potentially more productive when students have the chance to interact with each other. When students are communicating interactively, they may 1) bring in intuitive ideas, 2) contrast them during conflicts, and 3) discuss possible differences, which may lead to the co-construction of new ideas to solve the given problem. From that point of view, interactive communication can increase the effectiveness of PF. With reference to Chi and Wylie (2014), we assume that interactive sequences during problem solving outperform passive, active, and constructive communication processes concerning conceptual knowledge acquisition.

Building upon this work, we analyzed process data from the study by Loibl and Rummel (2014), which will be explained in more detail below. More specifically, the focus of our analyses was to carve out which communication patterns during problem solving prepare students the best for the instruction afterwards. For instance, it could be better for knowledge acquisition if students were communicating actively (e.g., by giving explanations) or interactively (e.g., by discussing with each other) rather than just observing communication processes (e.g., by listening to their partner's explanations). To analyze the video material, we used the following communication pattern categories: 1) passive observation, 2) active-constructive, and 3) interactive communication. In our data, active and constructive activities could not be meaningfully separated. We therefore combined these activities into one category. Because communication patterns are probably not the only important preconditions for effective knowledge acquisition from PF, we also took into account the quantity and quality of students' solutions, and their prior knowledge. By analyzing communication patterns together with these measures, we are able to investigate under which conditions the problem-solving phase prepares students most effectively for the following instruction.

Study by Loibl and Rummel (2014)

The process analyses presented here are based on data collected by Loibl and Rummel (2014). Therefore, we will first briefly describe the methods from this study before we explain in more detail our analyses. Loibl and Rummel (2014) compared the *timing of instruction* (i.e., before or after problem solving) and the *form of instruction* (i.e., with or without contrasting student solutions to the canonical solution) in a quasi-experimental 2x2 design with pre- and posttest measures. Participants were 10th graders from secondary schools in Germany. For our analysis, we used one condition of the study. In this condition, students first worked on a mathematical problem in small groups. After struggling with that problem, the students were introduced to the canonical solution through direct instruction. During the instruction phase, student solutions were contrasted with the canonical solution. The length of the two phases was inspired by a typical teaching unit in Germany, and therefore took 45 minutes for each phase. We analyzed 14 groups, consisting of three students each for a total of N = 42 students. The average age was M = 15.85 (SD = .66). The mathematical problem students received was about the concepts of variance with the two formulas: *mean absolute deviation* ($MAD = \sum |x_i - \text{mean}| / N$) and *standard deviation* [$SD = \sqrt{\sum (x_i - \text{mean})^2 / N}$]. The problem was adopted from Kapur (2012). Students had not received instruction on variance in school prior to the study. In the following sections, we describe variables used by Loibl and Rummel (2014), which we took into account for our analysis: 1) prior knowledge, 2) quantity and quality of student solutions, and 3) conceptual knowledge at post-test. We then describe the new process analysis (communication patterns) of our analysis.

Pretest

A pretest measured students' knowledge about descriptive statistics and graphical representation (as a prerequisite for learning the concept of variance) and asked students to report their final grades in mathematics of the last two school periods. In Germany, '1' represents the best grade while '6' represents the worst grade. To facilitate interpretation, grades were reverse-coded so that '1' represents the worst grade and '6' the best grade.

Quantity and quality of student solutions

The quantity and quality of student solutions refer to the knowledge co-construction during the problem-solving phase implemented by Loibl and Rummel (2014). As described above, the problem students received was related to the MAD and SD formulas. Both formulas have the following components: 1) sum up deviations for all data

points to obtain a precise result, 2) take absolute or squared deviations (i.e., positive values) to prevent positive and negative deviations from canceling each other out, 3) take deviations from a fixed reference point (the mean) to avoid sequence effects, and 4) divide by the number of data points to account for sample size. For the process analyses, the number of components included in each student solution was counted. Component 4 (division by the number of data points) was excluded from this paper, because for the analyzed condition, there was nearly no communication concerning that component, and the given problem did not require the division by the number of data points (see component 4 above) to solve the problem correctly. For all of the solutions, we assigned them to a category (C0-C3). Solutions that do not include any of the above components were classified as category 0. Solutions with one correct component were categorized as category 1 and so on. Therefore, solutions of a lower category are 'more incorrect', because they disregard more components. Notably, the categories focus on the number of correct components within one solution and do not indicate which of the components are included. In addition, an overall indicator of quality for the student solutions, which we refer to as *total quality*, was formed by determining the maximum number of components included across all solutions. The quantity of student solutions was measured by counting different solutions, notwithstanding how useful they were to solve the given problem.

Posttest

Conceptual knowledge was tested by four posttest items. Items included an explanation for a graphical solution, the identification and explanation of presented mistakes, and the correct assignment of graphical representations. The maximum score on the test for all four items was seven points.

Analysis of communicative patterns during problem solving

In order to investigate the relevance of small group collaboration for the effectiveness of PF, we performed a process analysis of the PF-condition in which students first engaged in problem solving and afterwards received instruction building on student solutions, implemented by Loibl and Rummel (2014). For the analysis, we used video data of the initial problem-solving phase. As described above, each video of the group problem-solving process had a length of 45 minutes.

Building upon the ICAP-Modell by Chi and Wiley (2014) we differentiated between three communication patterns: 1) passive observation, 2) active-constructive communication, and 3) interactive communication. As mentioned above, these communication patterns are describing different participative, task-relevant behaviors during the problem-solving phase. We cut the videos (45 minutes each) of the problem-solving phase in 270 10-second sequences. Each sequence was analyzed separately for each student in the groups (see *Table 1* for an example). During each sequence, a student was communicating about task-relevant information concerning the different solutions, which was assigned to a specific category (C0-3). For example, as you can see in *Table 1*, a student's communication status was rated as *interactive* during a given 10-second segment when there was at least one more contribution of another student within the same segment. When there was not, the contribution was classified as *active-constructive*. If a student did not communicate during a 10-second sequence while at least one other groupmate did, we rated the non-communicative behavior of the student as *passive observation*. We only considered task-relevant communication (e.g., explanations, confirmations, discussions, accentuations) and did not analyze the quality of a contribution (e.g., high elaboration). This resulted in three dependent variables: how often a student communicated during problem solving, in what form or pattern communication occurred (passive, active-constructive, interactive), and to what kind of solution-category (C0-3) it was related. These variables were useful to measure, how students communicated during problem solving, as you can see in *Table 1*.

Table 1: The coding matrix to measure the three communication patterns

	Sequence 1	Sequence 2	...	Sequence 269	Sequence 270
Time in sec.	00:00:00	00:00:10	...	00:44:40	00:44:50
Student A	C 1 Interactive	C 2 Active-constructive	...	-	C 3 Interactive
Student B	C 1 Interactive	- Passive	...	-	- Passive
Student C	C 1 Interactive	- Passive	...	-	C 0 Interactive

To account for differences in verbosity, we computed relative frequencies of each communication pattern for each group member by dividing his or her number of communication patterns by the total number of all communication patterns for the group as a whole. For example, one group had a total of ‘100’ interactive 10-second sequences during problem solving. In this case, student A participated on ‘80’, student B on ‘30’, and student C on ‘100’ 10-second sequences. We divided the values of each student by the total frequency of interactive communication (in this case ‘100’). If one member of the group communicated in every interactive sequence during problem solving (e.g., student C), the student gets assigned a value of ‘1’, which means, that the student was involved at 100% of the interactive 10-second sequences during the 45 minutes. Using this method, we computed the values for each member for each of the three communication patterns.

Results

The findings we present focus on 1) prior knowledge, 2) quantity and quality of student solutions, and 3) the three communication patterns. We analyzed which of these variables has an impact on conceptual knowledge acquisition. We analyzed the data with *Pearson* correlation coefficient (r). If variables were not normally distributed, we used the non-parametric coefficient *Spearman’s Rho* (r_s).

Pre-test scores did not correlate significantly with conceptual knowledge acquisition ($r_s = .278$, ns), but average mathematics grade did correlate with conceptual knowledge acquisition (described in detail below; $r = .365$, $p < .05$). In our process analyses, we therefore used the average grades in mathematics as an indicator for relevant prior knowledge. This grade can be seen as a measure of conceptual knowledge of mathematical contents and also the ability to solve mathematical problems. Prior knowledge also correlated with task-relevant communication ($r = .358$, $p < .01$). Students with better grades in mathematics showed less passive observation ($r = -.404$, $p < .01$) and participated with more active-constructive communication than their other group members ($r = .419$, $p < .01$). However, the correlation between the grade in mathematics and the frequency of interactive communication was only marginally significant ($r_s = .285$, $p = .065$).

Concerning the quantity and total quality of the generated solutions within the groups, there were no significant correlations with conceptual knowledge acquisition. Neither the quantity of generated solutions ($r_s = -.086$, ns) nor the quality ($r_s = -.014$, ns) within the groups correlated significantly with conceptual knowledge acquisition. Hence, the quantity and also the quality of generated solutions seemed to be of no importance for conceptual knowledge acquisition.

Next, we analyzed the relation between the communication patterns and posttest results. Concerning the three communication patterns, there were no significant findings with respect to knowledge acquisition. Neither for passive observation ($r = -.172$, ns), active-constructive ($r = .149$, ns), nor for interactive communication ($r_s = .073$, ns) was a positive correlation with knowledge acquisition found.

The following analysis of the communication patterns takes into account the quality of individual student solutions, each coded for the number of components included (C0-C3). With respect to the four qualitative categories of student solutions (C0-C3), we analyzed the communication patterns separately. It is notable that only for solutions of category 1 (i.e., solutions including one component) was there a significant correlation between the frequency of active-constructive communication and conceptual knowledge acquisition ($r_s = .381$, $p < .05$). This same pattern was found for interactive communication as well ($r_s = .424$, $p < .01$). It seems that active-constructive and interactive activities for students along with generated solutions of category 1 are associated with better conceptual knowledge acquisition. Further analyses of variance have shown that this is true particularly for interactive activities, thus we only describe the analysis of variance for interactivity. For the further analyses, we

took into account the students' score of active-constructive communication (high or low) and if solutions of category 1 were available during group discourse or not. We did the same for students' interactivity scores. As already mentioned, analyses of variance only showed effects when we took into account interactivity, as we describe it in more detail below. Therefore, higher frequencies of interactive behavior during problem solving only seem to be more or even only effective when groups struggled with solutions of category 1. Remember that solutions, which were assigned to this category, only included one solution component.

In order to support the assumption of an interrelation between more frequent interactive behavior and the availability of solutions assigned to category 1 as described above, we compared the factors *Interactivity* (high / low) and *Solutions of Category 1* (available / not available) within an ANCOVA, using mathematical grade as a covariate and scores on the post-test as the dependent variable. *Interactivity* (high / low) was based on a median-split of frequency of interactive communication patterns ($Md = .77$). In Table 2 you can see the means and standard deviations of the four subsequent conditions that we analyzed.

Table 2: Descriptive statistics which was analyzed within an ANOVA (Interactivity vs. Solutions of Category 1)

Solutions of Category 1	Interactivity	<i>N</i>	Conceptual Knowledge Acquisition <i>M (SD)</i>
available	high	13	3.50 (1.52)
available	low	13	2.34 (1.46)
not available	high	8	1.56 (.86)
not available	low	8	2.31 (1.09)

As you can see in *Table 2*, students who struggled with solutions of type category 1 and participated with high interactivity during problem solving, outperformed the other conditions with respect to conceptual knowledge acquisition. If students participated with low interactivity during problem solving, it seems to be irrelevant if they struggled with category 1 solutions or not (see *Table 2*). The worst results for conceptual knowledge acquisition were achieved by students who were interactive but did not struggle with category 1 solutions. The main effect of the factor *Solutions of Category 1* (available / not available) was significant, $F(1, 37) = 4.480, p = .041, \eta^2 = .108$, but the main effect of the factor *Interactivity* (high / low) was not, $F(1, 37) = .064, p = .802, \eta^2 = .002$. Further, there was a significant interaction effect between the two factors, $F(1, 37) = 5.320, p = .027, \eta^2 = .126$. This interaction particularly suggests that students who participated often in forms of interactive communication achieved better posttest results when they struggled with solutions of category 1. If there was no communication about solutions of category 1, interactive students achieved the lowest conceptual knowledge scores despite being interactive.

Conclusions and implications

Our findings suggest three initial conclusions: 1) favorable prior knowledge goes along with conceptual knowledge acquisition from PF; 2) solutions of category 1 seem to be of high importance for learning, particularly when 3) solutions of that category are linked to student's interactive communication.

First, we want to discuss the positive correlation between prior knowledge and conceptual knowledge acquisition. Our findings are in line with Kapur's (2014) assumption that the activation of prior knowledge during problem solving has an important role for the learning process. At this point it is important to stress that this requires the learner to have enough prior knowledge to work on the unknown problem and to not have too much knowledge to find an adequate solution for the given problem effortlessly. "*So the productive failure method depends strongly on choosing a problem that is inside this window with respect to the students' prior knowledge*" (Collins, 2012, p. 733). This could explain why the math grade, as an indicator for more comprehensive prior knowledge, correlated significantly with conceptual knowledge acquisition, and the specific prior knowledge did not. Maybe it is more important that students are highly skilled in solving mathematical problems than knowing much about the actual problem. Nevertheless, better prior knowledge seems to be an important condition for knowledge acquisition from PF.

Further, our findings suggest that, in particular, working on solutions of category 1 seem to foster conceptual knowledge but only when engaging in interactive communication. To interpret this finding, it is necessary to understand why solutions of category 1 are so fruitful for knowledge acquisition. This can be explained against the background of the final phase of PF, where student solutions are contrasted with each other and the canonical solutions. In contrast to the other categories (C0, C2, and C3), solutions of category 1 can be described as an intermediate step to the canonical solutions. In other words, these solutions are neither too far (as C0) nor too close from the correct solution (as C3 and possibly C2). As we mentioned above, the categories represent the number of components of the canonical solution that were correctly included in the solution attempt. Taking that into account, one can argue that solutions that do not include important solution components cannot be contrasted to the canonical solution during the instruction afterwards because there are no aspects, or too many aspects, to contrast. Further, if a solution includes all (C3) or close to all (C2) solution components, there is less space for contrasting because the differences between the canonical solution and the student solutions are too low. Solutions of category 1 may offer the highest potential for contrasts and differentiation. If this is true, why did solutions of this category only enhance conceptual knowledge acquisition when linked to interactive communication?

It is important to stress that the knowledge acquisition from PF does not take place during the problem-solving phase. Rather the conceptual knowledge acquisition happens afterwards, during direct instruction with contrasting solutions. The problem-solving phase serves as a preparation for the following instruction. In the words of Piaget (1977), students should assimilate intuitive ideas to the problem space to recognize that their solutions or concepts are not completely useful to solve the given problem. If they do so, the student's concepts can be accommodated during the instruction afterwards, but only if the instruction appeals to their concepts or solutions (see Loibl & Rummel, 2014). The point is that students only can recognize that their concepts do not fit as adequate problem solutions when there is an active confrontation during problem solving. For instance, students have to evaluate the solution by reflecting on the process of calculation or the produced result. Like socio-cognitive conflicts introduced by Mugny and Doise (1978), such a confrontation can emerge during interactive communication because students evaluate each other solutions against the background of their own preexisting concepts. This could also explain the interaction effect between interactive communication and the presence of solutions of category 1. When students produced solutions, which are near to the canonical solution (C2 and C3) or too obvious incorrect (C0), they cannot discuss the need or the absence of solution components. In contrast, solutions of category 1 do offer potential for discussion. Therefore, interactive communication may be much more productive or cause conflict when it refers to solutions of this category.

Further, the interactive participation does not only refer to solutions of category 1. A high amount of interactive communication means that the student was participating at the most in interactive sequences during the problem-solving phase. Therefore, high interactive participation of a student refers to almost every solution that was generated by the group. Interactive sequences may prepare for knowledge acquisition afterwards when the student contrasts solutions of category 1 with other solutions of different categories. However, the reason why high interactive communication without solutions of category 1 results in the worst posttest results concerning knowledge acquisition requires further investigation.

The most important limitation of our analysis is the operationalization of the ICAP-Modell (see Chi & Wylie, 2014), because we did not take into account the social learning activities in more detail. For instance, we did not distinguish whether interactive sequences included 'real' discussions (e.g., conflicts, etc.) or just synchronous contributions, which were not linked to each other.

Further, to analyze the presented data, we had to dichotomize the factor *Interactivity* (high / low). Therefore, the classification (high / low; $Md = .77$) was rather data driven rather than based on theoretical assumptions and experimental variation. Further, because of the small sample size (42 students; 14 groups), it was difficult to control for other important factors such as the quantity of solutions or group composition (see Loibl & Rummel, 2014). However, the analysis of various communication patterns during PF gives us the first insights into under which conditions interactive communication during problem solving could foster conceptual knowledge acquisition from PF by preparing students more effectively for the following instruction. Further studies have to take into account that interactive communication can emerge in different forms during problem solving and may exert influence only under certain conditions. One important condition seems to be that interactive communication occurs with those solutions that can be contrasted by the following instruction most effectively. Our findings provide a stepping-stone to investigate the role of social interactions for conceptual knowledge acquisition from PF in more detail.

References

- Chi, M. T. H. & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4), 219-243.
- Collins, A. (2012). What is the most effective way to teach problem solving? A commentary on productive failure as a method of teaching. *Instructional Science*, 40, 731-735.
- Deiglmayr, K., Rummel, N., & Loibl, K. (2015). The mediating role of interactive learning activities in CSCL: An Input-Process-Outcome model. In: O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.). *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Vol. 2* (pp. 518-522). Gothenburg, Sweden: International Society of the Learning Sciences.
- Dillenbourg, P. (1999). What Do You Mean by “Collaborative Learning”? In P. Dillenbourg (Ed.), *Collaborative Learning. Cognitive and Computational Approaches*. Amsterdam & Boston: Elsevier.
- Johnson, D. W., & Johnson, R. T. (2002). Social interdependence theory and university instruction. Theory into practice. *Swiss Journal of Psychology*, 61(3), 119-129.
- Johnson, D. W., & Johnson, R. T. (2009). An Educational Psychology Success Story. Social Interdependence Theory and Cooperative Learning. *Educational Researcher*, 38(5), 365-379.
- Kapur, M. & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences*, 21(1), 45-83.
- Kapur, M. & Kinzer, C. K. (2009). Productive failure in CSCL groups. *International Journal of Computer-Supported Collaborative Learning*, 4(1), 21-46.
- Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, 40(4), 651-672.
- Kapur, M. (2014). Productive Failure in Learning Math. *Cognitive Science*, 38(2014), 1008-1022.
- Loibl, K. & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction*, 34, 74-85.
- Mazziotti, C., Loibl, K., & Rummel, N. (2015). Collaborative or Individual Learning within Productive Failure: Does the Social Form of Learning Make a Difference? In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine & S. Ludvigsen (Eds.). *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Vol. 2* (p. 570-575). Gothenburg, Sweden: International Society of the Learning Sciences.
- Moshman, D. (1982). Exogenous, Endogenous, and Dialectical Constructivism. *Developmental Review*, 2, 371-384.
- Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performance. *European Journal of Social Psychology*, 8, 181-192.
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Oxford (GB): Viking.
- Sears, D. A. (2006). *Effects of innovation versus efficiency tasks on collaboration and learning*. Doctoral dissertation, Stanford University, California. Retrieved November 07, 2012, from <http://www.stat.auckland.ac.nz/~iase/publications/dissertations/06.Sears.Dissertation.pdf>.
- Stahl, G. (2013). Theories of Cognition in Collaborative Learning. In C. E. Hmelo-Silver, A. Chinn, C. K. K. Chan & A. M. O'Donnell (Eds.), *The International Handbook of Collaborative Learning* (pp. 74-90). New York, NY & London, UK: Routledge.
- Teasley, S. D (1997). Talking About Reasoning: How Important Is the Peer in Peer Collaboration. In L. B. Resnick, R. Säljö, C. Pontevorvo & B. Burge. *Discourse, Tools, and Reasoning. Essays on Situated Cognition* (pp. 361-384). Berlin: Springer.
- Van Boxtel, C., van der Linden, J. & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction*, 10(2000), 311-330.

Acknowledgments

We would like to thank Michael Wiedmann, Claudia Mazziotti and Jennifer K. Olsen for their valuable support in writing and revising this paper.

Competency-Based Digital Badges and Credentials: Cautions and Potential Solutions From the Field

Daniel T. Hickey, Indiana University, dthickey@indiana.edu

Abstract: This paper explores tensions that repeatedly surface alongside significant advances in instructional technology. One such recent advance concerns open digital badges that contain specific claims of competency and detailed evidence supporting those claims. Some assume that badges should only contain claims and evidence concerning specific measured competencies, while dismissing badges for participating in courses or activities that lack such evidence as “attendance badges.” Others assume awarding badges only for measured competencies bypasses very important forms of social and inquiry-oriented learning, ignores the limitations of assessment and measurement, and limits the transformative potential of web-enabled evidence-rich credentials. An extended study of the 29 projects in the 2012 *Badges for Lifelong Learning* initiative provided a unique opportunity to explore this enduring debate. None of six efforts to create competency-based open badge systems resulting in thriving ecosystems.

Keywords: digital badges, metadata, competency-based education, constructivism

Introduction

Each wave of advances in educational computing brings to the surface an enduring debate between two very different views of learning. Shortly after the graphical computer interface was introduced in 1961, the PLATO system was introduced at the University of Illinois. PLATO was primarily used to deliver programmed instruction in the behaviorist mastery-learning model that was widely accepted in era (e.g., Anderson, Kulhavy, & Andre, 1971). Six years after PLATO was introduced, Papert and colleagues at Bolt, Beranek and Newman introduced the LOGO programming language (Niemic & Walberg, 1989). LOGO was based on the cognitive developmental theories that were just coming to light among western educators and psychologists, and was intended to allow learners to discover programming concepts such as recursion and develop critical thinking skills. In Papert’s vision, “...the child programs the computer and, in doing so, both acquires a sense of mastery over a piece of the most modern and powerful technology and establishes an intimate contact with some of the deepest ideas from science, from mathematics, and from the art of intellectual model building” (1980, p. 5).

In practice, the divide between these approaches to educational computing was not always so clear, and many innovators resisted being essentialized. For example, Alpert and Bitzer (1970) insisted that PLATO was not limited to programmed instruction in basic skills but was also fostering inquiry and critical thinking. Nonetheless, moving forward into the personal computer era, publishers, parents, learners, and librarians had the choice between drill and practice programs like MathBlaster (Eckert & Davidson, 1987) and more discovery-oriented programs like the Logical Journey of the Zoombinis (Brøderbund, 1996) which was designed to support logical reasoning and critical thinking. While this tension simmered as online and hybrid instruction became more common (see Koszalka & Ganesan, 2004), it leapt to prominence with the emergence of massive open online courses (MOOCs). The initial “cMOOCs” emphasized knowledge construction and connections in emerging “connectivist” models of learning (Siemens, 2005). But the later “xMOOCs” like edX and Coursera featured streaming videos, structured practice problems, and multiple choice tests. The very scalable instructional model and limited peer interaction in xMOOCs allowed for the dramatic expansion of free online learning opportunities, leading the New York Times to dub 2012 “the year of the MOOC.” However, the static instructional model was a most frequently cited concerns in the “backlash” that followed (Kolowich, 2013).

This tension can be traced back to two very different views of cognition. PLATO, MathBlaster, and xMOOCs all assume disciplinary knowledge can be readily broken down into smaller associations that readily “reassemble” into more sophisticated higher order knowledge; such assumptions support instruction that efficiently trains and tests these specific associations. While some label such approaches as “behaviorist,” this is usually not appropriate. Many modern information processing models of human cognition also focus on specific cognitive associations as well. Hence, a more appropriate label for these approaches is “associationist.” Conversely, LOGO, Zoombini, and cMOOCs assume knowledge primarily consists of higher order conceptual “schema” that differ from learner to learner and are constructed via inquiry and problem solving. This calls for exploration and discovery that helps learners construct new knowledge. While this perspective is often labeled “cognitivist,” this is also generally inappropriate; the label “constructivist” or Papert’s term “constructionist” more accurately capture the focus on having students construct new conceptual knowledge; the term “connectivist”

captures the additional assumption in the cMOOCs that learners are doing so by digitally connecting with others. While these assumptions are tacit among many educators, the distinction between associationist and constructivist perspectives is widely appreciated in the Learning Sciences and among educational researchers more generally (e.g., Greeno, Collins, & Resnick, 1997) and has been the subject of vigorous debate (e.g., Kirschner, Sweller, & Clark, 2006 vs. Hmelo-Silver, Duncan, & Chinn, 2006; Tobias & Duffy, 2009)

Associationism vs. constructivism in new learning technologies

These tensions are taking on new importance as they surface within three related developments. The first is renewed interest in *competency-based education* (CBE). One proponent described CBE as “having a curriculum structured to demonstrate learning in clearly articulated competencies, is often self-paced, is agnostic as to the source of learning while maintaining clear and transparent learning standards” (Leuba, 2015, p. 1). While eschewing the radically reductionist approach of earlier mastery learning models (e.g., Guskey & Gates, 1985) with a focus on “authentic” and “real-life” demonstrations and projects, modern CBE models still focus on individualized mastery and demonstration of specific knowledge and skills, as opposed to completing activities or courses with other learners. Proponents of CBE put it in opposition with the prevailing credit-hour models, deriding the latter as “‘time-served’ rather than learning achieved” (Laitinen, 2012, p. 5). Another departure from earlier mastery learning models is that many current CBE models also “personalize” learning by giving students choices in how they gain and demonstrate their competencies (i.e., “agnostic as to the source of learning”). Within efforts to individualize learning with digital technologies, CBE has gained significant support from the US Department of Education, publishers, and major US Foundations. For example, the Gates Foundation’s *Project Mastery* initiative supports K-12 implementation of “proficiency-based pathways” which offer “opportunities for students to engage in a learning experience where they can demonstrate mastery of content and skills and earn credit towards a diploma, certificate, or some other meaningful marker” (Gates Foundation, 2014, p. 7). In the United States, although just taking hold in K-12 schools, CBE is already being used to structure entire universities (e.g., Southern New Hampshire University) and is entrenched in many professional vocational education sectors (e.g., healthcare and aviation). While concerns over “seat time” are mostly North American issues, a recent survey found noteworthy uptake of CBE around the world (Bristow & Patrick, 2014).

CBE is not without critics. A report from the Carnegie Foundation (the originators of the “Carnegie Unit”) provided perhaps the most comprehensive recent critique. Reflecting the underlying tensions introduced above, the report argued that “by focusing on the acquisition of discrete skills,” CBE “may make it more difficult to promote inter-disciplinary teaching, collaborative learning, and other instructional strategies that the latest research in learning science encourages—and the deeper, integrative learning that flows from those instructional strategies” (Silva, White, & Toch, 2015, p. 27). The report went on to summarize concerns that observers have raised regarding CBE, including: (a) CBE has the potential to widen achievement gaps, (b) the differentiation of instruction for individuals creates new challenges for educators, (c) it is expensive to implement, and (d) it overlooks the very real limitations of assessment and measurement.

A second recent development in learning technology is the advent of *open digital badges*. Launched by the MacArthur and Mozilla Foundations in 2011, this new web-enabled vision of credentialing has captured a good deal of innovation and attention (e.g., Carey, 2012). *Digital* badges can contain specific claims of learning, evidence to support those claims, and hotlinks to more evidence such as completed student work. *Open* digital badges can readily circulate in social networks and are interoperable with other learning management resources. The transparency associated with web-enabled credentials causes issuers, earners, and consumers to deliberate more about the nature of claims and the validity of the evidence than with conventional static credentials (whose credibility was derived from the issuing institution and its accreditation). This means that the introduction of open digital badges in educational ecosystems can be quite disruptive (Casilli & Hickey, 2016)

As with CBE, tensions that can be traced back to associationist and constructivist views of learning emerged around digital badges. Some viewed badges as an extension of CBE, insisting that the claims and evidence inserted in badges should contain measurable and (ideally) measured competencies. At the 2011 kickoff of the *Badges for Lifelong Learning* competition, US Secretary of Education Arne Duncan stated that digital badges “can help speed the shift from credentials that simply measure seat time to ones that more accurately measure competency, and we must do everything we can to accelerate that transition” (Duncan, 2011). Conversely, other influential commentators dismissed digital badges as behaviorist extrinsic incentives (e.g., Jenkins, 2012; Resnick, 2012). Badge proponents responded by pointing out that people could simply avoid using them as arbitrary extrinsic rewards for learning. Ravet (2014) raised this issue in the context of using badges to support a European effort to promote *Key Competencies for Lifelong Learning*. He argued against what he called “normative” badges awarded for mastery of specific competencies by articulating the more general worry that many constructivist researchers have with highly specific standards: “But there is one more fundamental problem

with standards, not with the standards as such, but with those who think that standards are the alpha and the omega of everything, letting standards be the proverbial tail wagging the dog” (p. 24). Rather than motivating compliance, and presumably undermining creativity and innovation, Ravet argued for “achievement badges” that are delivered *after* something has been achieved: “Achievements badges, contrary to key competency badges, do not have to be not normative. Created along the learning pathway, they can be designed with the learners rather than for them. As they can be created post-facto, they do not bear the stigma associated with the use of Open Badges as extrinsic motivators.” Ravet elaborated that the “beauty of achievement badges is that they capture the context of the achievement in the criteria: where, how, what resources, etc. And the collection of achievement badges creates a fabric of interwoven threads of narratives: one’s own story is interconnected to others’ stories through achievement badges (2014, p. 25).

This position directly follows from self-determination theory (particularly, Kohn, 1999). But the larger notion of badge-based pathways around more social forms of learning and recognition represent the underlying influence of contemporary sociocultural theories of learning (i.e., Yowell and Smylie, 1999) that helped shape the MacArthur Foundation’s introduction of open digital badges. In key ways, open digital badges nicely capture the tension between associationist and constructivist approaches, as well as the potential for the newer sociocultural perspective to point out a path forward. As shown below, a two year study of the 29 project funded in the aforementioned *Badges for Lifelong Learning* initiative was a useful context for exploring this potential.

A third relevant development in learning technologies is the ongoing effort to define *metadata* standards for educational information. Metadata is “data about data.” Efforts have been underway for years to come up with common standards for labeling and characterizing the vast quantities of information generated by e-learning systems (e.g., Bohl, Scheuhase, Sengler, & Winand, 2002). This work has taken on new importance with the use of learning management systems (LMSs) in nearly every college and university and many K-12 schools to support online, hybrid, *and* conventional classroom learning. Many in the Learning Sciences community are just beginning to appreciate the scope and implications of current efforts to develop educational information architecture standards. Proponents argue that these standards are necessary to support the inter-operability needed to allow learning technologies (including LMSs, applications, analytic tools, content, learner records, and credentials) to take full advantage of the Internet. As anyone who has tried to innovate within a comprehensive learning management system knows, the systems necessarily make assumptions that constrain options.

Much of the current information architecture work is being carried out by *IMS Global Learning*, a consortium of publishers, technology providers, schools, and policy makers. IMS Global’s new *Learning Tools Interoperability* (LTI) standards are already doing for LMSs what external apps do for smart phones. Consider, for example, digital badges. Issuing badges from within earlier LMS’s would have required massive amounts of programming and significant changes to the core program for each LMS. As of fall 2015, there are at least three LTI-compliant external applications for issuing digital badges (*Badgr* from the Badge Alliance/Concentric Sky, *BadgeSafe* from Accreditrust, and *BadgeOS* from Credly). Each is competing to offer desired functionality and serve different clients. Arguably, the existence of such standards and the ability to readily use and refine external applications are crucial for advancing the nascent digital badging community. IMS recently launched a project to develop a “currency framework” for digital badges that aims to “further the adoption, integration, and transferability of digital credentials within institutions, schools, and corporations” (IMS Global, 2015) (1). Likewise, similar efforts are underway by foundations and higher education associations to create common standards for “connected digital credentials” more generally. For example, one effort promoted by the American Council on Education argues that a “common DNA” is a critical ingredient in quality digital credentials. The effort cites research showing that that “all of the types of credentials in use—degrees, certificates, certifications, licenses, badges, etc.—can be described in the same language of competencies: the level of knowledge and specialized, personal, and social skills the credential represents” (Ganzglass & Good, 2015).

The tension between associationist and constructivist views of learning do not appear as obvious in efforts to standardize metadata for educational technologies and credentials. Given that education has primarily been organized around credits, the current standardization efforts include efforts to broaden technologies and credentials to support competency-based approaches. Leuba (2015, p. 1) describe a recent pilot effort to overcome these limitations, attempting to change the fact that “the products used to manage our institutions and the teaching and learning process are all deeply rooted in the credit-hour based, term-based, and course-based educational delivery model”. Initially, it seems, these efforts to standardize metadata and credentials will simply include things like “social skills” and “critical thinking” that are paramount to constructivist innovators as additional competencies to be categorized and measured and/or observed. As illustrated next, doing so within efforts to introduce digital badges systems may be more challenging than innovators expect.

Relevant findings from the Design Principles Documentation Project

The Design Principles Documentation (DPD) project was tasked with capturing the “practical wisdom” that emerged across the 29 diverse badges development efforts funded in the 2012 Badges for Lifelong Learning initiative, with the support of the John C. and Catherine T. MacArthur Foundation’s Digital Media and Learning initiative and the Gates Foundation’s Project Mastery initiative. In an elaborate competition, over 600 proposals were reviewed and the 29 winners were granted approximately \$200,000 and partnered with one of several badge system developer awardees to develop a badge system over a one-year period as the Open Badges Infrastructure standards and various supporting technologies were being developed by the Mozilla Foundation.

The DPD project first analyzed the content of 29 funded proposals to identify four categories of *intended practices* for digital badges: *recognizing* learning, *assessing* learning, *motivating* learning, and *studying* learning. The project then followed the 29 projects, conducting periodic interviews with project leaders to determine which of those practices each project was able to enact and what difficulties they ran into. In 2014 (after all of the new funds had been exhausted), the project carried out follow-up interviews to determine which practices were maintained, the final status of the badge system relative to the one articulated in the original proposal, and the final status of the larger badge-oriented “learning ecosystem” envisioned in the proposal. By clustering the more specific practices across projects into more general design principles, the DPD project was able to generate a wealth of knowledge about which principles were easier to implement and which principles were harder to implement. In 2015, the DPD project followed up to determine which ecosystems were still “thriving.”

The DPD project did not set out to study competency-based badge systems. However, seven of the 29 awardees (including three of the awardees funded by Gates’ Project Mastery) had proposed to develop badge systems for individualized self-paced mastery of specific competencies. As such, the DPD project was presented with a unique opportunity to examine the success of competency-based badge systems and related badge design principles. Before considering how these particular projects fared, it is worth noting that only 16 of the 29 projects succeeded in establishing the badge system they originally proposed, while 8 of the projects were judged to have established a different badge system than the one they proposed; 5 of the projects did not build any badge system.

Sustainable Agriculture and Food Systems (SA&FS) proposed to build its open badges system within a larger effort to develop a new competency-based interdisciplinary major in the College of Agriculture and Environmental Sciences at the University of California-Davis. As was widely reported in the educational media, SA&FS proposed to issue badges within a sophisticated custom e-portfolio system with comprehensive scoring rubrics, and was based on “a model of learning, participation, and assessment focused around high-level ‘core competencies’ that bridge classroom and real-world experiences, academic investigations and concrete skills” (2). As with many CBE projects, SA&FS carried out extensive research to document the specific competencies that employers who might hire their graduates were seeking. However, the competency-based initiative ultimately failed to gain the wider support of the university needed to sustain it and the project ultimately only managed to create a few test badges. The staff member who was spearheading the competency-based system left UC Davis and the new major was established as a conventional course-based program (3).

The *Pathways to Global Competence* badge system was proposed by the Asia Society, a New York-based non-profit that was also implementing its curriculum in dozens of US secondary schools with a *Project Mastery* grant. Like SA&FS, they attempted to build an open badge system around a sophisticated e-portfolio system; however, they proposed to do so in partnership with a commercial e-portfolio provider (ShowEvidence, Inc.). The project proposed to award *Global Leadership* badges to secondary students once they earned four global competency badges. For example, one of these four badges was for *Generating Global Knowledge*. Students were to earn this badge by submitting an e-portfolio that outside experts confirmed as evidence the student could “initiate investigations of the world by framing questions, analyzing and synthesizing relevant evidence, and drawing reasonable conclusions about globally focused issues.” (Asia Society, 2011). However the additional resources that were reportedly needed to establish the badges and the new e-portfolio system were not secured and both software development efforts encountered serious challenges, and the badge system was never developed. However, the competency framework and scoring rubric was implemented and is presumably still being used in most of their partner schools; an evaluation of the Project Mastery awardees found that many teachers reported having students develop their portfolios using Google Docs (Steele et al., 2014).

The *LevelUp* badge system was proposed by a partnership between the Adams County District 50 School System in Colorado, EffectiveSC, and Intific, Inc. Adams 50 was in the midst of a comprehensive effort to reform several of its underperforming schools using CBE with the support a Project Mastery grant. EffectiveSC was a non-profit that was developing the open-source *LevelUp* personalized competency tracking platform. Illustrating the technology challenges summarized by Leuba (2015), LevelUp was intended to streamline CBE by serving as “middleware” between the district’s existing *Educate* student information systems and online instructional resources where students could develop and demonstrate their competencies. Intific is a Texas software

development firm that was funded to develop four *Space Wolf* competency-based “learning progression games” that were to play a central role in the Adams 50 mathematics reforms. The project proposed to use the LevelUp platform to allow competencies that students developed playing the games to be automatically transferred to the *Educate* system. However, technology challenges with LevelUp and Educate and intellectual property issues kept that badges system from progressing beyond an initial pilot and it was suspended. According to participants from EffectiveSC, another obstacle was that the districts commitment to CBE declined sharply once the Project Mastery funding had expired and major technology challenges continued to frustrate teachers, students, and parents. These observations were generally confirmed in the external evaluation by Steele et al. (2014) which also reported statistically significant declines in math achievement in the participating schools (see also Sturgis, 2014).

The *Youth Digital Filmmaker Badge System* was proposed by the School District of Philadelphia, who partnered with the Youtopia’s commercial badging/gamification platform and the Philadelphia Youth Network. The badge system was part of a larger Gates Project Mastery initiative that also aimed to enhance the district’s *SchoolNet* LMS (from Pearson corporation) and its *Pathbrite* e-portfolio system to better support competency-based learning by allowing teachers to award course credits for competencies demonstrated in non-school projects. The badge system specifically aimed to foster afterschool “extended learning opportunities” around a new youth filmmaker program designed to “support academic credit attainment, anytime/anywhere learning, and skills mastery tied to Common Core State Standards for English and Language Arts.” More specifically, the project goal was “diversifying the ways and locations in which students can demonstrate mastery of critical reading, writing, and communication skills via multiple options to publish and produce films.” To support these goals, the project proposed to develop detailed scoring rubrics that would allow external agencies to endorse the badges, external experts to review storyboards, scripts, and videos for evidence of competencies, and teachers to award students formal academic credit for those competencies. However, significant technological challenges were encountered with all three new technologies. Additionally, the project failed to secure external endorsements for its badges and some of the teachers in the badges pilot project were reportedly reluctant to award formal course credit for the youth filmmaker badges. All three systems were paused after a single pilot.

The *National Manufacturing Badge System* was developed by the non-profit Manufacturing Institute in partnerships with *SkillsUSA*, a workforce development agency and with *Project Lead the Way* (PLTW), a national STEM curriculum-development initiative. The partnership with SkillsUSA aimed to issue badges for secondary vocational students in automated manufacturing programs who also attained passing scores on standardized performance-based assessments developed by SkillsUSA for industry-defined competencies. In this way, the project proposed to create “efficient competency-based pathways to careers” by integrating standardized assessments of industry-endorsed competencies into existing course-based educational programs. In this respect, the project has proposed the sort of education reform that many CBE proponents have been calling for. The project succeeded in creating a system for offering an Automated Manufacturing Technology badge for vocational students who passed the SkillsUSA assessments. But, the project abandoned its plan to offer “leveled” badges that recognized mastery of the more specific competencies (reportedly because they concluded that such a profusion of badges would confuse employers) (4). Instead the criteria for earning the final badge was attaining a passing score on all of relevant SkillsUSA assessments. Unfortunately, the project was unable to secure formal endorsements from manufacturers who would hire badge earners (reportedly because they wanted to first “see somebody who had earned the badge who could do the job”). This endorsement was necessary to convince educators to incorporate the assessments into their courses; without it, the project stalled at the pilot testing stage.

The other Manufacturing Institute badge project was a Computer Integrated Manufacturing badge for students at partner schools who completed PLTW’s standardized curriculum and attained a passing score on PLTW’s end of course assessment. However, other than the endorsement of the Manufacturing Institute, the badge did not contain any additional evidence beyond the formal credential issued by the schools. While the badge is still being offered by this partnership in 2014, the project leader reported that few, if any, of the potential earners were claiming the badge; while the badge was still being offered in 2015, we found no evidence it was thriving.

The *Young Adult Library Services Association (YALSA)* badge system was proposed to recognize mastery of YALSA’s Competencies for Serving Youth in Libraries (5). This included 48 specific competencies in seven areas, including social media, collection building, and public outreach. YALSA first attempted to create seven “pie badges” that displayed which of the sub-competencies that each earner had demonstrated. However, this proved technologically challenging and was set aside in lieu of badges that were issued once mastery of all of the competencies had been demonstrated. The project encountered significant challenges in establishing its website and was forced to scrap its initial system and state anew with a second web development team. They ultimately created a sophisticated system for peer assessment of the artifacts earners were asked to submit in order to earn the badges. However, in 2015, project leaders reported that few were attempting to earn the badges and concluded that the amount of work required to earn even a single badge was apparently too great given, given that there were

no specific incentive or opportunity associated with earning it. They reported that their advisors encouraged them to make sure that the badges were "really valuable."

The *Buzzmath* badge system was proposed by ScoLab, a small educational software firm in Montreal. They proposed to issue badges to recognize mastery of specific competencies as learners progressed through a drill and practice game for middle school mathematics (akin to MathBlaster). The firm used the grant to develop the badges as well as the larger Buzzmath platform. The project succeeded in building both the platform and the badges and aligning both to Common Core math standards. The system has proven to be a commercial success and continues to thrive; an independent evaluation showed that students and teachers believe that playing the games had a positive impact on math achievement and understanding (Morrison, Ross, & Lusiczka, 2015). However, privacy concerns precluded the use of web-enabled *open* and a planned peer tutoring system (the use of emails addresses as identifiers violate the stipulations of the US Children's Online Privacy Protection Act). They were also unable to secure external endorsements of their badges by schools or the organizations the established the educational standards their badges were aligned to.

Summary and conclusions

In summary, only none of the seven projects that attempted to build badge systems around self-paced mastery of highly specific competencies succeeded in creating a thriving open learning ecosystem--BuzzMath succeeded in building a system that used non-open badges as tokens in a drill and practice program. The reasons the other projects struggled were varied. Certainly, the blame cannot be placed entirely (or in some cases even partly) on the decision to implement a competency-based system. Put differently, the DPD project did *not* conclude that these projects would have been more successful had they attempted to issue "time-based" badges based on participation in courses or other education activities. Nonetheless, these projects suggest care and caution is needed when developing competency-based badge systems. In particular, it seems competency-based systems should anticipate the challenges that the DPD project uncovered as well as the tensions in CBE implementations reported in the separate evaluation of the three Gates' Project Mastery initiatives (Steele et al., 2015). These tensions included *equating evidence from anytime/anywhere learning with conventional criteria, determining who can authorize credit, maintaining a common definition of proficiency, building a sustainable model, technical barriers to efficiency, financial barriers to efficiency, logistical barriers to efficiency, and promoting equity*. In key ways, these conclusions bolster the concerns in the aforementioned Carnegie Foundation report, while also highlighting the challenges that student information systems present for CBE summarized by Leuba (2015).

Consider for example, the tensions over equating evidence and authorizing course credit. As mentioned previously, one of the three teachers in the Youth Filmmaker pilot was reluctant to accept the badges for course credit. The Rand evaluation reported project leaders were surprised by this reluctance given the efforts they had taken to align their badges to the Common Core standards. The Rand Report explained that the teacher was not convinced "that the persuasive writing skills appropriate for script development of a documentary film were the equivalent of what he expected students to achieve in a persuasive essay, especially in terms of issues like essay structure and sentence structure." The report went on to state that the teacher "believed that preparing a short nonfiction film and preparing a persuasive essay tapped different skills, both applicable to the real world, but not interchangeable" (Steele et al., 2014, p. 42). After examining the curriculum and the relevant standards, the report concluded the teacher's position "seems reasonable." They also reported concerns such as privacy and validity led all three of the Project Mastery projects to abandon their plans to use external expert arbiters of credit; in the case of the Youth Filmmaker project, this eliminated the intended expert feedback, a key element of the envisioned learning ecosystem. The Rand evaluation reported that Asia Society teachers "found it difficult to get district approval to turn school-led travel experiences into course credit" (p. 42) and that some resorted to substantial measures to bypass their school districts course-based policies and information systems.

The tensions over equating and authorizing gets at the heart of the CBE's "agnosticism" regarding the source of learning. For many assessment researchers and validity experts, this assumption is untenable. This is because an assessment system *must* take into account for the way individuals prepared for that assessment if the resulting evidence is to provide valid evidence to support claims of competency. This is because two different individuals can attain the same score on a given test or produce portfolios or other artifacts that earn the same score via very different paths. Compare, for example, an individual who has prepared narrowly for a particular assessment by memorizing the specific associations (correct and incorrect) represented by the various items with an individual who has completed a broader course or learning pathway that was not focused on those associations. If both individuals get the same score, the second individual is almost certainly more knowledgeable than the first, because of all of the other new knowledge that the assessment could not capture. Likewise, consider an individual who creates an e-portfolio following very detailed guidelines and has access to the scoring rubric and

individualized feedback on drafts against the rubric. This individual almost certainly is less competent than an individual who submits a similarly-scored e-portfolio developed without access to this information and feedback.

This concern with “teaching to the test” is what the validity expert Samuel Messick (1994) labeled “construct-irrelevant” easiness. Hickey and Zucker (2013) argued that this phenomenon has always been more insidious and much more complex than many educators and innovators realize. Caution seems called for given (a) the compromise and cheating made possible by new digital devices and social media and (b) the disdain for multiple-choice achievement tests and embrace of performance-based and portfolio assessment among CBE and digital badge proponents. Take, for example, the seemingly sensible practice of awarding badges and credit for “anytime/anywhere” learning for passing “scenario-based” performance assessments. Once such information is circulating freely in professional social networks, it becomes substantially easier for vocation educators and potential test takers to figure out precisely what scenario is used in the assessment and potentially the specific questions asked. Unless the assessment includes dozens of scenarios and hundreds of questions, it cannot possibly cover the entire range of competencies. Such concerns certainly give credence to the concerns constructivist educators and proponents of time-based credentials have for strict competency programs. Messick introduced raised this issue of construct-irrelevant easiness in the context of the large-scale K-12 performance and portfolio assessment reforms that came (and largely went) in the 1990s. Concerns over validity and lack of promised positive consequences for teaching and learning were major reasons for the rollback of this earlier wave of assessment reforms. As a kind of assessment reform, it seems that CBE should attend to this issue as well.

Arguably, newer sociocultural perspectives that provided much of the impetus for introducing digital badges offers a potential path forward. As argued by Hickey (2015), sociocultural approaches to assessment and validity tend to focus primarily on communal participation in social practices, and only secondarily on assessment of individual knowledge and skill. This allows these practices to treat all forms of individual assessment as “secondary” kinds of evidence—special cases of primarily social learning. When coupled with contemporary design-based research methods that guide iterative refinement of communal participation, this makes it possible to treat individual performance on the entire range of assessments and tests as evidence of the success of those refinements. An initial examination of two project whose badge systems are particularly thriving (the *Support to Reporter* youth journalism project and the *MOUSE* youth network manager mentoring program) supports this belief. Both issued what might best be described as “role-based” badges. While the badges included claims of specific competencies, the evidence of these competencies was primarily the earner’s participation in workshops and other activities with peers and endorsements by experts and peers. Rather than creating a comprehensive list of competencies and assessments in advance, both projects gradually and iteratively refined networked social learning activities for cohorts of participants to maximize the quality of and quantity of disciplinary interactions among learners (see O’Byrne, Schenke, Willis, & Hickey, 2015). Additional efforts now underway are exploring this question and attempting to derive a more comprehensive characterization of how these perspectives guide the design of successful open badge systems while reconciling the tensions between different perspectives.

Endnotes

- (1) Disclosure: the author is participating in this activity as an advisory board member.
- (2) Unless indicated otherwise, all of the quotes from are from the submitted proposals or DPD Project interviews which can be accessed at the project website and reported in Author (in preparation).
- (3) <https://www.pltw.org/pltw-engineering-curriculum>
- (4) <http://asi.ucdavis.edu/programs/safs>
- (5) <http://www.ala.org/yalsa/guidelines/yacompetencies2010>

References

- Alpert, D., & Bitzer, D. L. (1970). Advances in computer-based education. *Science*, 167(3925), 1582-1590.
- Anderson, R. C., Kulhavy, R. W., & Andre, T. (1971). Feedback procedures in programmed instruction. *Journal of Educational Psychology*, 62(2), 148-156.
- Asia Society (2011). *Graduation portfolio system global leadership performance outcomes*. New York.
- Bill and Melinda Gates Foundation (2014). *Supporting students: Investing in innovation and college*. College Ready Work Monographs. Seattle WA.
- Bristow, S. F., Patrick, S. (2014). *An international study in competency education: Postcards from Abroad*. Washington, D.C.: International Association for K-12 Online Learning.
- Bohl, O., Scheuhase, J., Sengler, R., & Winand, U. (2002, December). The sharable content object reference model (SCORM)-a critical review. Proceedings from the *International Conference on Computers in Education*. 950-951.
- Brøderbund (1996). *Logical Journey of the Zoombinis (Computer software)*. Novato, CA: Brøderbund.

- Carey, K. (2012 November 12). Show me your badge. *The New York Times*.
- Casilli, C., & Hickey, D. (2016). Transcending conventional credentialing and assessment paradigms with information-rich digital badges. *The Information Society*, 32(2), 117-129.
- Duncan, A. (2011, September 1). Opening remarks and DML 2012 competition event transcript. Retrieved from: <http://dmlcompetition.net/Competition/4/opening-event-transcript.php>.
- Eckert, R., & Davidson, J. (1987). Math blaster plus [Computer software]. *Torrance, CA: Davidson & Associates*.
- Ganzglass, E., & Good, L. (2015). *Rethinking credentialing*. Washington DC: American Council on Education, Center for Education Attainment & Innovation.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1997). Cognition and learning. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-47). New York: Simon & Schuster Macmillan.
- Guskey, T. R. & Gates, S. L. (1985). A synthesis of research on group-based mastery learning programs. Proceedings from the 69th Annual Meeting of the American Educational Research Association.
- Hickey, D. T. (2015). A situative response to the conundrum of formative assessment. *Assessment in Education: Principles, Policy & Practice*, 22(2), 202-223.
- Hickey, D. T., & Zuiker, S. J. (2012). Multilevel assessment for discourse, understanding, and achievement. *Journal of the Learning Sciences*, 21(4), 522-582.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark. *Educational Psychologist*, 42(2), 99-107.
- IMS Global Learning Consortium (2015, April 21). *IMS Global Announces Initiative to Establish Digital Badges as Common Currency for K-20 and Corporate Education* [Press Release].
- Jenkins, H. (2012, March 5). *How to earn your skeptic badge*. [Web log post].
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, 41(2), 75-86.
- Kohn, A. (1999). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston: Houghton Mifflin Harcourt.
- Kolowich, S. (2013). Faculty backlash grows against online partnerships. *Chronicle of Higher Education*.
- Koszalka, T., & Ganesan, R. (2004). Designing online courses: A taxonomy for strategic use of features available in course management systems (CMS) in distance education. *Distance Education*, 25(2), 243-256.
- Laitinen, A. (2012). *Cracking the Credit Hour*. Washington DC: The New America Foundation. [Policy Paper].
- Leuba, M. (2015, October 12). *Competency-based education: Technology challenges and opportunities*. EDUCAUSE Review, [Online].
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Morrison, J. R., Ross, S. M., & Lesiczka, J. K. (2015). *Report for iZone NYC Department of Education: Results of Buzzmath short cycle evaluation*. Baltimore MD: Johns Hopkins University Center for Research and Reform in Education.
- Niemiec, R. P., & Walberg, H. J. (1989). From teaching machines to microcomputers: Some milestones in the history of computer-based instruction. *Journal of Research on Computing in Education*, 21(3), 263-276.
- Ian O'Byrne, W., Schenke, K., Willis, J. E., & Hickey, D. T. (2015). Digital badges: Recognizing, assessing, and motivating learners in and out of school. *Journal of Adolescent & Adult Literacy*, 58(6), 451-454.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc..
- Ravet, S. (2014). #Openbadges for key competencies. *Learning Futures: Reflections on Learning, Technologies, Identities, and Trust* [Web log post].
- Resnick, M. (2012, February 27). *Still a badge skeptic*. [Web log post].
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Education*, 2(1).
- Silva, E., White, T., & Toch, T. (2015). The Carnegie Unit: A century-old standard in a changing educational landscape. *Carnegie Foundation*.
- Steele, J. L., Lewis, M. W., Santibanez, L., Faxon-Mills, S., Rudnick, M., Stecher, B. M., & Hamilton, L. S. (2014). *Competency-Based Education in Three Pilot Programs: Examining implementations and outcomes*. Santa Monica, CA: The Rand Corporation.
- Sturgis, C. (2014, March 12). *A conversation with Adams 50*. [Web log post].
- Tobias, S. & Duffy, T. M. (Eds.). (2009). *Constructivist instruction: Success or failure?* New York: Routledge.
- Yowell, C. M., & Smylie, M. A. (1999). Self-regulation in democratic communities. *The Elementary School Journal*, 469-490.

Scaling Up Productive Disciplinary Engagement With Participatory Learning and Assessment

Daniel T. Hickey, Suraj Uttamchandani, and Joshua D. Quick,
dthickey@indiana.edu, suddamch@indiana.edu, jdquick@indiana.edu
Indiana University Learning Sciences Program

Abstract: This paper presents the results of a three-year program of design research that aims to scale highly interactive social learning in open online courses. This research started with simple features that had previously been refined in conventional online courses in order to foster more productive forms of disciplinary engagement. These features reflect five design principles for fostering productive forms of disciplinary engagement, motivating and assessing that engagement, and finally assessing the resulting individual understanding and achievement. This paper shows how these features were scaled up for use by hundreds of students in an open course and summarizes the impressive levels of disciplinary engagement and achievement that resulted. It also presents two new features that were introduced to allow self-paced participatory learning with little or no instructor involvement.

Keywords: personalized learning, learning analytics, assessment, analytic approaches

Introduction

MOOCs and other open online courses have dramatically expanded opportunities to learn by significantly reducing the added incremental cost of additional students once the course is established. The so-called “xMOOCs” (“eXtended”) at edX, Udacity, Coursera, and elsewhere typically feature streaming videos, online readings, automated problem sets and quizzes, and peer discussion forums. The explosive growth of MOOCs led the *New York Times* to deem 2012 “the year of the MOOC” while also leading to new scrutiny. Studies showed that many were disappointed in the more limited interaction in xMOOCs (e.g., Khalil & Ebner, 2013), and one effort to include more interaction and group projects in Coursera was widely cited for going “laughably awry” (Oremus, 2013). Some observers had already commented on the difficulty of connecting with others in the “connectivist” “cMOOCs” that were designed specifically to support social interaction at scale (Mackness, Mak, & Williams, 2010). One study found that engagement in Coursera discussion forums declined significantly over time among completers due to information overload as discussion threads become unnavigable, and that instructor involvement actually worsened participation (Brinton et al., 2014). The lack of peer and social interaction was a prominent concern in the widely-cited “backlash” against MOOCs in 2013 (e.g., Kolowich, 2013).

MOOC proponents responded that the social experience in typical MOOCs was actually quite similar to what many students experience in the large lecture courses that are common at many college and universities. Indeed, the peer discussion forums available for many MOOCs are similar to the informal study sessions that many students organize themselves into for conventional lecture courses. Nonetheless, many MOOC organizations and instructors began investing significantly in improving the usefulness of their discussion forums by adding features, trained volunteers, and paid discussion leaders. By 2014, a significant progress was already underway towards more social interaction and peer learning in MOOCs and in learning at scale more generally. For example, a team at Stanford was developing a MOOC platform (now called NovoEd) which is explicitly based on social learning theory (Ronaghi, Saberi, & Trumbore, 2015). A major program of research at the Open University resulted in FutureLearn which supports “discussion-in-context” and “community-supported learning” (Parr, 2013), while a promising new strand of research supporting social interaction around peer assessment emerged (Kulkarni, Socher, Bernstein, & Klemmer, 2014). It is particularly noteworthy that some of these efforts to scale social learning are drawing insights from the Computer-Supported Collaborative Learning and Learning Sciences communities. This includes work at Carnegie-Mellon using intelligent conversation agents and social recommendation technology to enhance edX discussion forums (Rosé, Goldman, Sherer, & Resnick, 2015) and work at the University of Toronto to support inquiry-oriented project-based learning and user-contributed content in edX (Håklev, Slotta, & Najafi, 2015). Indeed, commentators have recently pointed to increased social learning as the “next challenge” in supporting learning at scale (e.g., Bryant, 2015; Parr, 2014).

Participatory social learning and the Assessment BOOC

This paper is concerned with scaling a particular type of social learning. This type of learning builds on theories of situated cognition that emerged at the Institute for Research on Learning starting in the late 1980s. In particular, the social learning described in this paper builds on the “situative synthesis” advanced by Greeno and colleagues (Greeno, 1998). This perspective diverges from prior theories of cognition by focusing primarily on knowledge that resides in social and cultural practices (i.e., is “situated”) and then framing individual behavior and individual cognition as secondary “special cases” of this more socially defined knowledge. Building on Greeno’s notion of “engaged participation” in the communal construction of socially-defined knowledge, these perspectives value participation *as* learning and are thus often referred to as “participatory.”

The instructional framework that is the focus of this paper is currently referred to as Participatory Learning and Assessment (PLA). The PLA framework emerged from an extended program of design-based educational research using situative theories of cognition to uncover new ways of engaging learners, assessing learning, and evaluating programs. The PLA framework as it is presented here initially emerged in design studies of secondary language arts instruction (Hickey, McWilliams, & Honeyford, 2011) and educational videogames (Hickey & Jameson, 2013). Key features used to enact this framework online emerged in conventional graduate courses on *Assessment in Schools* and *Learning and Cognition* (Hickey & Rehak, 2013). This paper concerns a new series of design studies that explored how these features could be streamlined and automated to allow the same interactive social learning while making fewer demands on the instructor. This was intended to allow “big” open online courses with hundreds of students in the near term, working towards massive courses with thousands of students in the longer term.

A grant from Google was used to offer a “big open online course” (BOOC) in the fall of 2013 based on the first author’s existing *Assessment in Schools* course. The grant supported customization of Google’s *Coursebuilder* platform to scaling “wikifolios” and peer interaction features that had previously delivered in small courses using *Sakai*. As summarized in Hickey, Kelly, and Shen (2014), over 500 students (mostly practicing teachers and faculty) registered for the 12-week course and 160 completed the first assignment. Over 60 students (including eight for-credit students) completed all three modules (assessment *practices*, *principles*, and *policies*) to earn a web-enabled digital credential, a “badge” that could contain all completed work and interactions. The “Assessment BOOC” was again offered in the summers of 2014 and 2015. With the grant funding depleted, the open course was not as widely promoted and the efforts were primarily supported by the for-credit students completing the course as a part of their graduate degree. Most of the programming effort was directed at further automating course features to allow learners to experience interactive peer learning while placing minimal demands on the instructor. Additional features were added to allow self-paced learners. This ability to support self-paced learners while still offering them the interactive peer learning experience has been a major goal of this program of research. The weekly deadlines and relatively demanding assignments made it difficult for many open learners to keep up with the course. More generally, these features should eventually make it possible to provide entirely peer-to-peer experiences requiring instructor input, eventually leading to an entirely automated massive course featuring much (but likely not all) of the interaction describe below.

PLA design principles and corresponding features

The PLA design principles coordinate activity across different kinds of interactions that afford different kinds of learning. Drawing on Hall and Rubin’s (1998) studies of mathematics, the principles distinguish between interactions that are *public* (visible to every participant), *local* (in public but between specific individuals), or *private* (between individuals). A fourth kind of interaction, *discreet* (i.e., unobtrusive), was added to highlight the core PLA assumption that achievement tests may be appropriate (and are indeed necessary to support some claims of competency), but should be used judiciously and presented inconspicuously.

These principles are all oriented towards supporting what Engle and Conant (2002) called *productive disciplinary engagement* (PDE). This framework presumes that engagement that is *disciplinary* involves both the declarative knowledge of the discipline (what experts *know*) as well as the social and cultural practices in which disciplinary experts engage (what experts *do*, in professional contexts). They further argued that disciplinary engagement that is *productive* generates numerous connections between that declarative knowledge and the learner’s experiences engaging in disciplinary practices. The first two PLA design principles directly follow from Engle and Conant’s design principles for fostering PDE. Central to this extension is the manner in which PLA uses engagement in disciplinary *practices* (which are more concrete and meaningful) to foster engagement with disciplinary *knowledge* (which is more abstract and challenging to learn). A central insight from the prior design research was that peers more readily question and discuss characterizations of disciplinary practices because they are generally not “factual” like disciplinary knowledge (Hickey & Rehak, 2013).

PLA principle #1: Use public context to give meaning to knowledge tools

This first principle embodies the core situative assumption that the context in which disciplinary knowledge is learned and used is a fundamental part of that knowledge (Greeno, 1998). Students' own prior experiences, current interests, and future aspirations (i.e., their nascent disciplinary practices) are used to publically "problematize" the disciplinary knowledge of the course. This is consistent with Engle and Conant's first design principle: *problematize subject matter from the perspective of the learner* (Engle & Conant, 2002).

Public wikifolios

All interaction is organized around public (to the class) "wikifolios." These are intact pages where each set of instructions can be hidden or revealed, and where participants respond to instructions in simple WYSIWYG editing boxes. The wikifolios engage participants in more open-ended disciplinary practices (which naturally foster discussion) rather than more specific declarative knowledge. The wikifolios generally use personalized disciplinary practices to engage students with the more specific disciplinary knowledge. In each wikifolio, disciplinary engagement is fostered with "ranking" features, which have proven to be a simple and scalable way of fostering social PDE. Students rearrange text boxes summarizing 3-5 elements of declarative knowledge that are elaborated in the textbook and/or open educational resources. Participants make sense of that information by ranking the relevance of each element relative to their curricular aim (discussed below) and/or professional role, and then justify that ranking. A particularly important aspect of this feature is that even when students lack the experience or understanding to rank something, they must engage with the knowledge to reach that conclusion. This prepares them to readily appreciate the rankings and rationales of peers with more experience. The expectation which has been repeatedly borne out in practice is that learners develop a routine whereby they do an initial ranking and rationale before looking at examples of peers who they interact with regularly.

Each wikifolio included other activities that reflected the textbook chapters. For example, each of the wikifolios in the "assessment practices" unit had students create example assessments using item writing guidelines before ranking the relevance of those guidelines. Other features had students simply summarize the big ideas of a text chapter or external resource. Each wikifolio also included a number of "optional" elements that the for-credit students were required to complete. This open format and the various features described next succeed in motivating participants to write a great deal. In 2013 the eight credential students wrote an average of 1398 words per wikifolio across their 11 wikifolios, while the 60 open learners who completed the course averaged 1207 words. Most impressive was that the 100 open learners who started but did not complete the course averaged 1137 words per wikifolio. In 2014, the average number of words per wikifolio doubled for the 12 credential students to 2820, while the averages for the 10 open completers (1377) and the 54 who dropped (1080) stayed roughly the same. The 2015 course maintained these levels of engagement for the 23 credential students (2374) and the 2 open completers (1783), while the 5 who dropped averaged 843 words per wikifolio.

Self-contextualization at registration

When participants register for the Assessment BOOC, they are asked to select a primary educational *role* (teacher, administrator, etc.) and *setting* (secondary, college, etc.) and asked to define an initial *curricular aim* (a learning goal for a particular class they have taught or might teach) that embodies that role. The role and setting are used to automatically assign students to a networking group on the participant locator page described below. The curricular aim is automatically inserted into the first wikifolio assignment, which guides students through the process of using a text chapter and open educational resources to further refine that aim. This registration feature highlighted the personalized approach that the course would take, presumably discouraging registrants who were not serious or did not like that approach. Reflecting the core situative assumption that disciplinary practice and disciplinary knowledge reciprocally define each other, learners were instructed to continue refining their curricular aim in each wikifolio as their knowledge of classroom assessment expanded. Put differently, the refined curricular aim embodies each learner's growing understanding of their own professional practice; as their knowledge of assessment grows, their knowledge of relevant aspects of their professional practice grows as well. This growing understanding their own professional experience provides the context which grounds their learning of new and otherwise more abstract disciplinary concepts as the course progresses.

Peer networking groups and peer location tools

A challenge for fostering peer learning at scale is helping participants find ideal peers to interact with. The participant locator display is a page that lists each participant along with hotlinks to their published wikifolios. Depending on what kinds of peers one is looking for, participants can be displayed according to networking groups, primary role, recent updates, and number of peer promotions. In 2014, the third wikifolio assignment was modified to invite students to extend their usernames and thus project additional identities that were not included

in the original networking groups (e.g., *librarian, unemployed math teacher*). More than half of the students now do so. In 2015, a new archiving feature let participants archive their wikifolios in a way that reflected their wishes for interacting with subsequent participants after they had completed the course. One setting displayed the archived work in green. This setting caused subsequent comments to trigger an email to the author and indicated the author *would* respond to comments. The yellow archive setting also caused subsequent comments to trigger an email but indicated that author *might* respond to comments. The red archive setting did not trigger email notification (but subsequent students were still free to endorse, promote, and comment. While students also have the option to delete their wikifolios when they complete the course, very few do so.

PLA principle #2: Reward productive disciplinary engagement

This second principle assumes that productive forms of disciplinary engagement with resources, peers and instructors should be facilitated and rewarded. This principle uses situative theories of motivation and incentives (Hickey, 2003) to enact Engle and Conant's second and third PDE design principles: *give students authority over their disciplinary engagement* and *hold students accountable for their disciplinary engagement*. Engle and Conant's study, and the various subsequent studies that they and others carried out to further refine these principles, were conducted in conventional classrooms (where PDE could be modeled by the teachers and fostered in conversation). Although several features had been refined for enacting these two principles online in the prior courses, enacting them at scale in the BOOC required significant innovation. It is worth noting that students were instructed to engage in peer commenting, promotion, and endorsement, but there was no requirement or accountability of any sort beyond the collaborative reflection described below.

Peer commenting

A key feature for supporting PDE online was having students and instructors interact with each other via threaded comments posted directly to wikifolios. This contrasts with discussion forums, which are relatively removed in time and space from the completed student work, and which routinely veer from the topic of the assignment and sometimes veer from the topic of the course. The scale of the BOOC and the introduction of self-paced learning called for a way for participants to readily locate (a) new comments on their own wikifolios, (b) replies to their comments on other participants' wikifolios, and (c) new course announcements and feedback. A hotlinked notification feature was added at the top of every page alerting each participant to all such developments and linking directly to the new activity.

When coupled with the peer locator tools, this feature made it quite simple for participants to efficiently engage with their peers and allowed the instructor to efficiently highlight (and therefore reward) exemplary wikifolios and discussion threads via the participatory feedback described below. In 2013, the for-credit students and open completers averaged 5.6 and 3.0 comments per wikifolio, respectively; in 2014, these two groups averaged 4.2 and 3.4 comments per wikifolio. In 2015, these two groups averaged 4.0 and 3.6 comments per wikifolio; across groups and years, comments averaged around 100 works. Coding of a representative subsample of comments from 2014 revealed that 92% of the comments referenced the topic of the chapter directly (Hickey, Quick, & Shen, 2015). While we lack a ready comparison, this seems like much more disciplinary engagement than is typical of many conventional online courses and most open courses.

Peer endorsement and promotion

These features assume that conventional peer assessment is usually awkward and not particularly productive, because it focuses on "known answer" questions about disciplinary knowledge. While students dislike assessing whether peers "know," something, they do not seem to mind assessing whether peers "did" something. The peer endorsement feature allowed participants to endorse wikifolios as "complete." Peers simply clicked one of two buttons if the author had completed (a) all of the required elements or (b) all of the required elements plus the optional elements. The peer promotion feature allowed to students to promote one (and only one) wikifolio each week as being "exemplary." Peers were required to include a warrant for the promotion and the warrant was displayed alongside the name of the peer who awarded it. In 2013, 56% of the students who posted a wikifolio also promoted a peer; this proportion increased to 60% in 2014 and to 86% in 2015. These seem like very promising levels, particularly given that there was no accountability for either type of participation.

Participatory feedback

A major challenge in scaling up wikifolios is that most useful examples and interactions are "buried" within individual wikifolios. In the 2013 and 2014 Assessment BOOCs, the cohorted format made it possible to have two types of announcements around the weekly wikifolio deadlines. The *early posts* announcements went up after the first 3-5 participants posted their wikifolios each week, invariably including some of the most experienced and ambitious participants. The instructor would reward the early posters with extensive comments, elaborations,

and pointed questions used to introduce more advanced content. The early posts announcement included hotlinks and encouraged other students to examine the early posts and comments, but only *after* getting started on their own wikifolio. The assumption is that the experience of starting the wikifolio would provide the other learners the personalized context that would help them make sense of the examples and additional content.

The *post-deadline* announcement went up a day or two after each weekly deadline. This highlighted and linked to the most widely promoted wikifolios and important discussion threads. The post-deadline announcement also included one of the most novel and potentially far-reaching features of the course. Each week the instructor and research assistants would examine the aggregated rankings by networking group for illuminating patterns. For example, when completing the *Validity* wikifolio in the *Principles* module, most of the educators concluded that *content-related* validity evidence was most relevant (because it concerned the content of an assessment), while most of the administrators concluded that *criterion-related* validity evidence was most relevant (because it concerned the relationship between scores and external criteria such as promotion). A bar graph displaying this information and hotlinks to examples helped participants review and revisit difficult concepts. For example, this explained that the few students who found *construct-related* validity evidence most relevant were doctoral students who were interested in constructs like self-efficacy. Comments on the announcements page and reflections confirmed that this feedback helped the educators and administrators more fully understand the (highly abstract) notion of psychological “constructs.” We believe that this innovation is particularly productive and a good example of how situative theories lead to useful features that help learners make sense of the most challenging elements of disciplinary knowledge.

While participatory feedback seems very promising, it was also one of the most laborious aspects of the Assessment BOOC. Even though the ranking by group data was downloaded to a spreadsheet, it was still necessary to locate, summarize, and link to good examples. In 2015, the two forms of feedback were combined into one weekly post to ease the workload. Comments on the announcement pages and interviews confirmed that many participants (a) found the feedback and examples interesting and useful, (b) were motivated by the possibility of getting mentioned, (c) wanted to see who got mentioned each week, and (d) used the feedback to help review for exams. Providing participatory feedback has proven challenging in moving to self-paced courses. Currently, self-paced learners are encouraged to examine current or archived examples of the existing early poster-feedback and the current wikifolios that received the most peer promotions. Current efforts will automatically display graphs of the rankings by groups in real time with links to promoted examples.

Digital badges

One additional goal of this program of research was exploring the ways that open digital badges (evidence-rich web-enabled credentials) could motivate and recognize productive disciplinary engagement and achievement. Completing each of the three modules generated a badge in which learners could choose to include any and all of their work and interactions. Earning the three module badges generated the course badge (which contained the three module badges). These badges employed the new Open Badges Infrastructure metadata specifications. This meant that learners could readily share them over email or social media, which would display a hotlinked impact that would redirect the badge viewer to our course site and the archived work.

Each module badge can display the number of comments, endorsements, and promotions posted on each wikifolio, along with a link to the wikifolio as well as the content of those comments and promotions. In 2013-2015, one member of each networking group was awarded a *Leader* version of the badge for earning the most peer promotions. The Leader badges clearly served to motivate some students to post early and engage more deeply with their peers. Efforts are now underway to determine if and how leader badges might be awarded in a fully self-paced courses without cohorts of students in which to make this judgement.

PLA principle #3: Evaluate artifacts via local reflections

Happily, the features used to enact the first two principles generate lengthy wikifolios and extensive peer commenting. Particularly regarding the credential students (and the corresponding expectations for accountability), this creates a new challenge of evaluating and grading all of these artifacts and interactions. The third PLA principle eschews any formal summative evaluation of the content of public artifacts and local interactions. This principle thus builds on existing assessment research that suggests “no marks” (i.e., ungraded) feedback (Harlen, 2007) and concerns about excessively detailed portfolio and performance assessment rubrics.

These prior assessment guidelines were reframed using sociocultural approaches to portfolio assessment (Batson, 2011; Habib & Wittek, 2007). This perspective leads to the assumptions that artifacts themselves primarily show what learners *did* (not what they *can* do in the future). This is because the many different routes to producing a given artifact means that it is very difficult to use artifacts themselves to support claims of proficiency. This principle also reflects the corresponding assumption that formal summative evaluation of

artifacts in order to make such claims undermines the formative goal of building individual and collective knowledge around the creation and discussion of those artifacts (Gitomer & Duschl, 1995). Thus, instead of diminishing learners' engagement with the disciplinary content by marking it as "right" or "wrong," wikifolios were graded through learners' reflections. The features used to enact this principle have been refined and studied extensively (Hickey, 2015). This is a crucial feature for portfolio assessment because it helps resolve the tensions between formative and summative functions described by Barrett (2010).

Contextual, collaborative, and consequential reflections

One of the optional wikifolio elements (required for credentials) consisted of three carefully worded reflection prompts for learners to answer once they had posted their wikifolio and interacted with peers. Building on Gresalfi, Barab, Siyahhan, and Christensen (2009), participants are instructed to reflect on their *contextual engagement* ("How suitable was your context for learning this knowledge?"), *collaborative engagement* ("Who else's work and whose comments helped you learn this new knowledge?"), and *consequential engagement* ("What will you do differently in your context and beyond as a consequence of learning this knowledge?").

The assumption here is that students who had not engaged productively and socially with the disciplinary knowledge of the course will be unable to draft a coherent and convincing reflection. In this way the reflections summatively assess one kind of learning (prior engagement) while formatively assessing another kind of learning (understanding the relationship between new disciplinary knowledge and a variety of disciplinary practices). The ultimate intention of these reflections is rooted in the anthropological notion of *prolepsis* (the way anticipated future events shape present activity; Cole, 1993); learners know they will need to reflect and so engage more deeply than they otherwise might. Put differently, we assume that because learners know that they will have to reflect on these nuanced – but important – aspects of engagement, they are motivated to think about these prompts in advance of the reflections as part of completing their own wikifolios.

This feature helped limit both time-intensive private feedback to students and extended review of completed student work. So long as the for-credit students posted their wikifolios on time and included a coherent reflection, they were given all of their points for the wikifolios (which counted for 55% of the grade). In practice, the process of privately awarding points to the for-credit students was nearly automated. Points were awarded and unless the student work was particularly weak, a boilerplate feedback statement was added.

PLA principle #4: Let individuals assess their understanding privately

PLA assumes that the more formal assessments that efficiently generate valid evidence of *prior* learning inevitably frame that knowledge in ways that limit the assessment's value for directly supporting *new* learning (Author, 2013a; 2015). This leads to a second assumption that public and local interactions should not take place around the more static representations of knowledge (i.e., known answer questions) in formal assessments, and thus that any formal assessment of knowledge should be carried out privately. A third assumption is that well-designed "curriculum-oriented" assessments are uniquely suited for letting participants figure out for themselves how much declarative knowledge that they have taken away from their prior engagement with the resources and peers. A fourth assumption is that such curriculum-oriented assessments that are of a reasonable length can only cover a fraction of the declarative knowledge that engaged learners should take away from each assignment. A fifth assumption is that such assessments can't really assess the extent to which students connected that knowledge to their own disciplinary practices, because such knowledge is so highly contextual.

Ungraded self-assessments

In 2014, ungraded quizzes featuring six to eight open-ended assessment items were added to each wikifolio. Students had to enter a response to each item in order to see the scoring key for the item. These formative assessments were entirely voluntary and students were encouraged to attempt the items from memory. The instructions recommended that students who were unable to answer more than one item from memory should re-engage with their classmates (starting with the public feedback) and the text before taking the module exam. These instructions have been repeatedly refined in an effort to maximize the formative benefit for engagement and to discourage students from memorizing the answers to those questions to prepare for the exam.

PLA principle #5: Measure aggregated achievement discreetly

This principle encourages using externally developed multiple-choice achievement test items for very specific purposes. It has proven to be one of the most controversial aspects of the PLA framework due to widespread concerns that such tests narrow curriculum and focus on shallow factual learning. The "distal" items are "standards oriented." The principle assumes that as long as the items are not "cherry picked" to tap into topics of the specific curriculum, they can be used to create an achievement test that is largely independent of the way a particular

course was designed. As such they are useful (and indeed necessary) for measuring learning within courses, comparing learning across different versions of the same course, and accurately documenting course improvement over time. By “discreet” this principle means unobtrusive and ephemeral; course assignments should never be directly aligned to achievement tests. In most cases students should only see their overall score. Most importantly, little if any course time should be devoted to instructing students on how to answer multiple-choice items. Such tests should feature items that go beyond factual knowledge and the items should be analyzed using the item analysis routines that are widely available.

Time-limited multiple-choice achievement tests

Each of the three modules in the Assessment BOOC included a timed multiple-choice exam consisting of items selected from multiple assessment textbooks’ item banks. The exams included many “best answer” (rather than “correct answer”) items which would be impossible to look up and figure out with the limited time available. Test takers only saw their score, and not the correct answer for each item. Item analysis was used to replace misbehaving items. Scores averaged around 85% with one or two perfect scores and a normal distribution.

Starting in 2015, participants who did not attempt the exam or did not attain a score of 80% only earned badges for *Assessment Practices*, *Assessment Principles*, and *Assessment Policies*; Participants who took each exam and attained at least 80% earned the *Expertise* version of the module badge and could choose to include that information in their badge. Participants who earned all three badges also earned the *Educational Assessment* Badge which contained the other three badges. To earn the *Educational Assessment Expertise Badge* participants had to earn two out of three expertise badges and attain a score of at least 80% on a comprehensive final exam. Credential students were required to complete all of the exams, and those scores counted towards 45% of their grade in the course (10% for each module exam and 15% for the final).

Conclusions

The impressive levels of engagement and achievement with hundreds of learners in 2013 and with dozens of learners and minimal instructor involvement in 2014 and 2015 support several conclusions. Most generally, these findings suggest significant progress was made in scaling up participatory learning. While we are still not prepared for massive scale courses, we are unaware of any other effort to scale up open learning that resulted in these levels of disciplinary engagement, understanding, and achievement. Second, these findings support the conclusion that *some scaling should be done gradually*. In order to quickly scale up to massive numbers of users, most MOOCs and MOOC platforms were forced to sacrifice interaction and personalization; because the code behind them is already so complex, those platforms are now finding it challenging to incorporate new features to support social learning. Third, the steady pace of improvement over years supports the conclusion that *scaling should be done iteratively*. These efforts were directly shaped by newer design-based research methods that emphasize the development of “local” theories in the context of reform efforts. Furthermore we conclude that such iterative refinements should be done within a coherent theoretical framework. Our commitment to situativity allowed us to draw directly from other research in that tradition to generate useful insights and solutions. In particular we found the notion of PDE particularly helpful, because it let us evaluate our innovations in terms of their presumed or actual impact on disciplinarity and productivity of interactions.

A final point is that this approach seems likely to have a much more profound impact on learners’ professional identities than most existing scalable instructional models. We conclude with our conviction regarding the value of insistently connecting the learning of new disciplinary knowledge with a growing understanding of one’s disciplinary practices *and* the practices of one’s peers. We believe that these approaches represent an efficient and scalable way of achieving the “joint accomplishment of identity” as recently described by Hand and Gresalfi (2015). This conclusion in turn points to a major question going forward regarding the extent to which this approach can be used in other domains and contexts. It is currently being used successfully in English, social studies, and biology Courses at the fully online *Indiana University High School* (Itow & Hickey, 2015). Further, pilot studies in Secondary Algebra and Freshman Calculus have confirmed that the framework requires substantial revision for use in mathematics (Uttamchandani & Hickey, 2015).

References

- Barrett, H. (2010). Balancing the two faces of ePortfolios. *International Journal of ePortfolio*, 3(1), 6–14.
- Batson, T. (2011). Situated learning: A theoretical frame to guide transformational change using electronic portfolio technology. *International Journal of ePortfolio*, 1(1), 107–114.
- Brinton, C. G., Chiang, M., Jain, S., Lam, H. K., Liu, Z., & Wong, F. M. F. (2014). Learning about social learning in MOOCs: From statistical analysis to generative model. *IEEE Transactions*, 7(4), 346-359.
- Bryant, T. (2015, June 22). Bringing the social back to MOOCs. *EDUCAUSE Review*. [Online]

- Cole, M. (1993). Remembering the future. In G. Harmon (Ed.) *Conceptions of the human mind. Essays in honor of George A. Miller* (pp. 247–265). New York: Psychology Press
- Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learner’s classroom. *Cognition and Instruction*, 20(4), 399-483.
- Gitomer, D. H., & Duschl, R. A. (1995). Moving toward a portfolio culture in science education. *Learning science in the schools: Research reforming practice*, 299-326.
- Greeno, J. G., & Middle School Mathematics through Applications Project Group. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5–26.
- Gresalfi, M., Barab, S., Siyahhan, S., & Christensen, T. (2009). Virtual worlds, conceptual understanding, and me: Designing for consequential engagement. *On the Horizon*, 17(1), 21-34.
- Håkløv, S., Slotta, J. & Najafi, H. (2015, Oct). *It Wouldn't Be the Same without You: MOOC Design that Supports User-Contributed Content, Interaction and Teamwork*. Columbia University: Learning with MOOCs II,
- Habib, L., & Wittek, L. (2007). The portfolio as artifact and actor. *Mind, Culture, and Activity*, 14(4), 266-282.
- Hall, R., & Rubin, A. (1998). There’s five little notches in here: Dilemmas in teaching and learning the conventional structure of rate. *Thinking practices in mathematics and science learning*, 189-235.
- Hand, V., & Gresalfi, M. (2015). The joint accomplishment of identity. *Educational Psychologist*, 50, 190–203.
- Harlen, W. (2007). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation*, 33(1), 15-28.
- Hickey, D. T. (2003). Engaged participation versus marginal nonparticipation: A stridently sociocultural approach to achievement motivation. *The Elementary School Journal*, 401-429.
- Hickey, D. T. (2015). A situative response to the conundrum of formative assessment. *Assessment in Education: Principles, Policy & Practice*, 22(2), 202-223.
- Hickey, D. T. & Jameson, E. (2012). Designing for participation in immersive educational videogames. In D. Ifenthaler, D. Eseryel, X. Ge (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 401-430). New York: Springer
- Hickey, D. T., Kelly, T. A., & Shen, X. (2014). *Small to big before massive: Scaling up participatory learning and assessment*. Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, Indianapolis, IN (pp. 93-97) <http://dx.doi.org/10.1145/2567574.2567626>.
- Hickey, D. T., McWilliams, J. T., & Honeyford, M. A., (2011). Reading Moby-Dick in a participatory culture: Organizing assessment for engagement in a new media era. *Journal of Educational Computing Research*, 44 (4), 247-273
- Hickey, D. T., & Rehak, A. (2013). Wikifolios and participatory assessment for engagement, understanding, and achievement in online courses. *Journal of Educational Media and Hypermedia*, 22 (4), 229-263.
- Itow, R. C., & Hickey, D. T. (2015, April 20). *Impacting online teacher practice, epistemic frames, and agency with design-based implementation research*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Khalil, H., & Ebner, M. (2013). “How satisfied are you with your MOOC?” A research study on interaction in huge online courses. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Vol. 2013, pp. 830–839). Retrieved from <http://www.edlib.org/p/112057/>
- Kolowich, S. (2013, May 1). Faculty backlash grows against online partners. *Chronicle of Higher Education*.
- Kulkarni, C. E., Socher, R., Bernstein, M. S., & Klemmer, S. R. (2014, March). Scaling short-answer grading by combining peer assessment with algorithmic scoring. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 99-108). ACM.
- Mackness, J., Mak, S., & Williams, R. (2010). The ideals and reality of participating in a MOOC. In *Proceedings of the 7th International Conference on Networked Learning 2010*, 266-275.
- Oremus, Will. (2013, February 3). Online class on how to teach online classes goes laughably awry. *Slate*.
- Parr, C. (2013, May 23). FutureLearn reveals big plans to deliver MOOCs on the move. *Times Higher Education*, 23, 9.
- Parr, C. (2014, November 13). Making MOOCs social is the next challenge. *Time Higher Education*.
- Ronaghi, F., Saberi, A., & Trumbore, A. (2014). NovoEd, a social learning environment. In Paul Kim (Ed.), *Massive Open Online Courses: The MOOC Revolution* (96-105). New York: Routledge.
- Rosé, C. P., Goldman, P., Zoltners Sherer, J., & Resnick, L. (2015). Supportive technologies for group discussion in MOOCs. *Current Issues in Emerging eLearning*, 2(1), 5-20.
- Uttamchandani, S. L., & Hickey, D. T. (2015, September 11). Calculus PLANet: Promising first steps in participatory supplemental instruction in mathematics. *Remediating Assessment Blog*: <http://remediatingassessment.blogspot.com/2015/09/calculus-planet-promising-first-step-in.html>

Providing Adaptive Scaffolds and Measuring Their Effectiveness in Open Ended Learning Environments

Satabdi Basu, Vanderbilt University, satabdi.basu@vanderbilt.edu
Gautam Biswas, Vanderbilt University, gautam.biswas@vanderbilt.edu

Abstract: Open ended learning environments (OELEs) offer students powerful learning opportunities, but managing the available tools and choices is challenging for novice learners. Adaptive scaffolding can support learners, but it requires understanding and responding to learner actions and strategies. We discuss a generalized adaptive scaffolding framework for OELEs based on a task and strategy model, and effectiveness and coherence measures for evaluating learner proficiencies for tasks and strategies. We apply this framework to support learners in CTSiM (Computational Thinking using Simulation and Modeling), a learning-by-modeling OELE for synergistic learning of science and computational thinking (CT). The effectiveness of our approach is demonstrated by a classroom study with two conditions. Students in the adaptive scaffolding condition showed a better understanding of science and CT concepts, built more accurate models, used better modeling strategies, and transferred modeling skills better to new scenarios than students in the control condition who received no adaptive scaffolding.

Keywords: open ended learning environments, learning by modeling, computational thinking, modeling and simulation, science education, learner modeling, adaptive scaffolding

Introduction

Open-ended learning environments or OELEs (Land, 2000) are learner centered computer environments designed to support thinking-intensive interactions with limited external direction. They typically provide a learning context and a set of tools to help students explore, hypothesize, and build solutions to authentic problems. The complex nature of the problems requires students to develop strategies for decomposing their problem solving tasks, developing and managing the accompanying plans, and monitoring and evaluating their evolving solutions. Thus, OELEs offer powerful learning opportunities for developing metacognitive and self-regulation strategies (Bransford & Schwartz, 1999). However, learning in OELEs is challenging for novices who may lack proficiency in using the system's tools, resulting in adoption of suboptimal learning strategies. Adaptive scaffolding may help learners overcome these difficulties (Puntambekar and Hubscher, 2005).

Many OELEs provide non-adaptive supporting tools like guiding questions, argumentation interfaces, workspaces for structuring tasks, and data comparison tools. As Puntambekar and Hubscher (2005) point out, such scaffolding tools support student learning, but they neglect important features of adaptive scaffolding such as ongoing diagnosis, calibrated support, and fading. Even in OELEs with adaptive scaffolding, few of them provide scaffolds that target students' understanding of domain knowledge, cognitive processes, and metacognitive strategies in a unified framework. MetaTutor (Azevedo, 2005) measures student behaviors using factors, such as the number of hypermedia pages visited and the length of time spent on each page, to decide when to provide adaptive scaffolds, e.g., "*You should re-read the page about the components of the heart*". In Ecolab (Luckin and du Boulay, 1999), the scaffolding agent intervenes when students specify an incorrect relationship in their models and provides a progression of hints, each more specific than the previous one, with the final hint providing the answer. In Co-Lab (Duque et. al., 2012), the system provides feedback on students' models and work processes, but is limited to reminding students about model building and testing actions not taken.

In this work, we have developed a task- and strategy-based modeling framework combined with coherence analysis to interpret and analyze students' actions (Segedy et. al. 2015) in Computational Thinking using Simulation and Modeling (CTSiM) – a learning-by-modeling OELE that we have developed to support synergistic learning of science and computational thinking (CT) in middle school science classrooms (Basu et. al., 2014; Sengupta et. al., 2013). We use the framework to determine need for scaffolding in CTSiM, and report results from a classroom study where a group of students used CTSiM with adaptive scaffolding and another group used it without the adaptive scaffolding. The effectiveness of our scaffolding approach is demonstrated in terms of students' science and CT learning, online modeling performance, learning strategies employed, and transfer of modeling skills outside the CTSiM OELE.

The CTSiM learning environment

The CTSiM environment (Basu et. al., 2014; Sengupta et. al., 2013) adopts an agent-based, learning-by-modeling

approach where students' model building activities are supported by two linked representations for conceptual and computational modeling. In the abstract conceptual model representation, students use a visual editor to identify the primary agents and environmental elements in the domain of study, along with their relevant properties. Students also identify agent behaviors and represent them using a sense-act framework by specifying which properties need to be sensed in order for the behavior to occur, and which properties will be acted upon in the behavior. For example, in one of the activities where students model a fish tank, 'fish' represents an agent with properties like 'hunger' and 'energy' and behaviors like 'feed' and 'swim', while 'water' is an environment element with properties like 'cleanliness' and "dissolved oxygen." The 'fish-feed' behavior senses the properties 'fish-hunger' and 'duckweed-existence', and acts on properties like 'fish-energy'. However, this representation abstracts details like how and when the different properties are acted on. These details are captured in the computational models, where students use a visual programming environment, and add and arrange provided blocks from a palette to create their models. The programming blocks can be domain-specific (e.g., "speed-up" in kinematics, "feed" in biology) or domain-general (e.g., conditionals and loops). The properties specified in the sense-act conceptual model for a behavior determine the set of domain-specific blocks available in the palette for the behavior. This dynamic linking helps students gain a deeper understanding of the representations and their relations. For example, the 'wander' block is available in the palette of available blocks for the 'fish-swim' behavior only if 'fish-location' is specified as an acted on property for the behavior.

Figure 1 represents the "Build" interface for modeling agent behaviors ('fish-feed' in this case). The leftmost panel depicts the sense-act conceptual representation, the middle panel shows the computational palette, and the right panel contains the student-generated computational model. The side-by-side placement of the representations is deliberate to emphasize their connectedness. To further aid the integration, the red/green coloring of the sense-act properties provides visual feedback about the correspondence between students' conceptual and computational models for an agent behavior. Initially, all the properties are colored red. As students add 'sense and act' blocks corresponding to the properties, the properties change color from red to green. For example, Figure 1 specifies O₂-amount as a sensed property for the fish-feed behavior. However, the computational model does not include O₂-amount and hence the property is colored red. In such cases, students can verify individual agent behaviors and decide how to refine their computational and/or conceptual models.

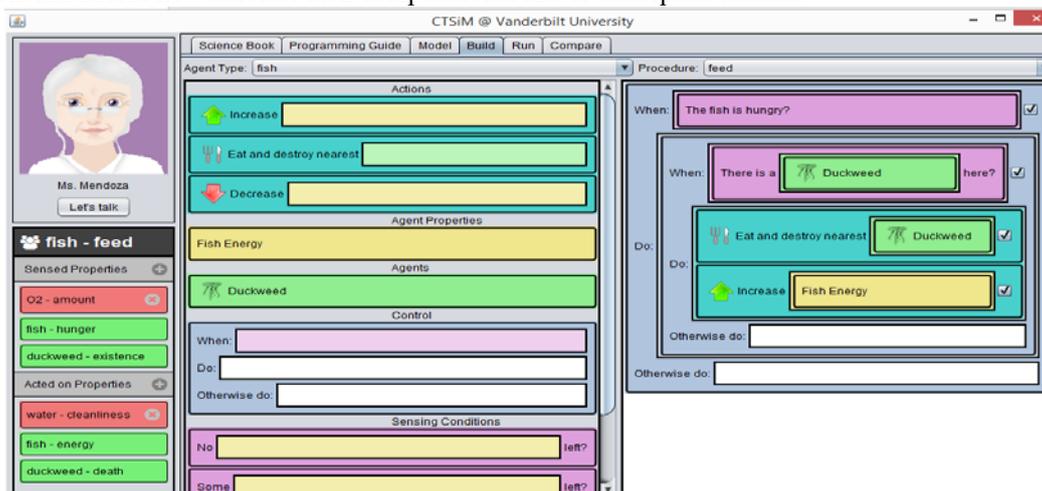


Figure 1. The linked conceptual-computational "Build" interface for modeling agent behaviors.

As students construct their models, they can visualize their model behaviors as NetLogo simulations (Wilensky, 1999), and verify their evolving models (the entire model or a subset of agent behaviors) by comparing the model behaviors against a matched 'expert' simulation. They do not have access to the expert computational model, but can analyze the differences between the simulation results to guide them in improving their models.

CTSiM also provides two sets of searchable hypertext resources, one with information about the science topic being modeled, and the other with information about agent-based conceptual and computational modeling. Students can also check their understanding of science and CT concepts by taking formative quizzes administered by a mentor agent in the system named Ms. Mendoza. The mentor grades students' responses to the multiple-choice type quiz questions and suggests resource pages to read in case of incorrect student responses.

A generalized scaffolding framework for OELEs

We develop a theoretical framing for our generalized scaffolding framework for OELEs and describe how we

apply it to the CTSiM environment. Some of the distinguishing features of our framework are as follows:

1. Tracking and interpreting learner behaviors using a task and strategy model
2. Determining the effectiveness of tasks and strategies using coherence metrics based on the relatedness and relevance between the tasks performed.
3. Providing adaptive scaffolding on students' strategies when their task performances are below par.
4. Providing contextualized feedback as mixed-initiative conversations initiated by a scaffolding agent.

At the core of our adaptive scaffolding approach is a task model that provides a hierarchical breakdown of the primary OELE tasks into their component subtasks and observable actions. The top layer of the task model defines tasks common across a class of OELEs; the middle layer defines related subtasks specific to a particular OELE; and the lower levels map onto observable actions performed using the tools provided in the OELE. As a specific example, the CTSiM task model in Figure 2 breaks down the OELE tasks of information acquisition (IA), solution construction (SC) and solution assessment (SA) into CTSiM specific subtasks and actions. IA is linked to identifying and interpreting science and CT information by reading and searching through the science and CT resources and taking formative quizzes, SC covers using identified information to build conceptual and computational models of science topics, and SA covers verifying the models in their entirety or in parts by observing their simulations or comparing their simulations against expert simulations.

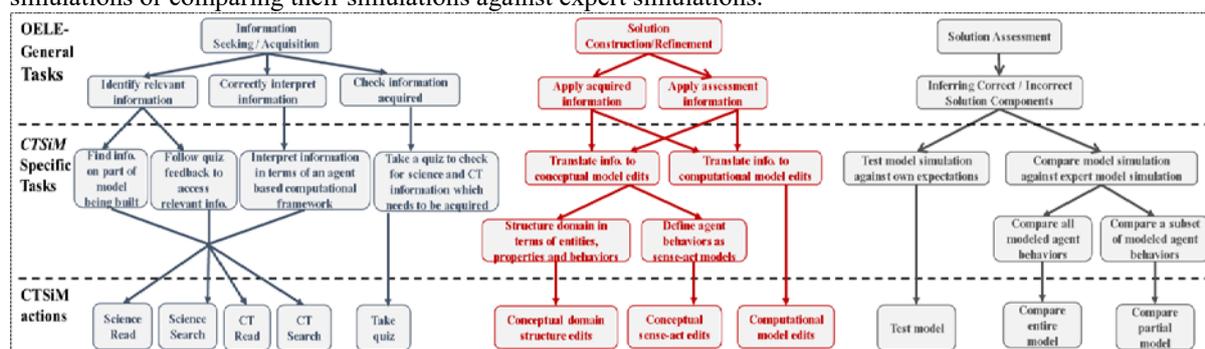


Figure 2. The CTSiM task model.

The task model does not specify any ordering or relations between sub-tasks or actions. These relations are represented by a strategy model that defines meaningful sequences of actions, subtasks, and tasks for accomplishing model building and learning goals specified in the OELE. For example, a set of action sequences that characterize features of individual actions (unary relations) and relationships between two or more action sequences (binary and higher-order relations) specify a ‘strategy model’ for CTSiM. In this work, we use a unary measure called ‘*effectiveness*’, where effective actions move the learner closer to their corresponding task goal. For example, effective SC actions bring the learners’ conceptual and computational models closer to a desired ‘correct’ model, and effective SA actions generate information about the correctness (and incorrectness) of individual agent behaviors modeled by the learner. Similarly, we adopt binary ‘*coherence*’ metrics for defining effective strategies, where two temporally ordered actions or tasks ($x \rightarrow y$), i.e., x before y , exhibit the coherence relationship ($x \Rightarrow y$) if x and y share contexts, i.e., the context for y contains information contained in the context for x . The context for an action comprises the specifics of the action, such as the specific science or CT page read, the conceptual or computational components edited, or the agent behaviors compared. A general strategy definition can be hierarchically linked to more detailed versions that represent desired or suboptimal variants. By tracking student’s activities, the system can compare strategy matches to desired versus suboptimal variants to estimate the student’s proficiency versus need for scaffolding with respect to the strategy. The need for scaffolding is determined based on a combination of suboptimal action sequences and low modeling performance.

While several useful strategies can be defined using different combinations of tasks and actions from the CTSiM task model, we chose a set of five desired strategies (S1–S5) based on our previous observations of students’ difficulties. We analyzed students’ actions to detect deficiencies in these strategies.

S1. *Desired*: SC followed by coherent IA action ($SC \Rightarrow$ Science Read)

Suboptimal: (ineffective $SC \rightarrow$ Science Read), i.e. ineffective SC action followed by an incoherent science read

S2. *Desired*: SA followed by coherent IA action ($SA \Rightarrow$ Science Read)

Suboptimal: (effective SA detecting incorrect agent behaviors \rightarrow Science Read), i.e. a SA action testing the model in parts and detecting incorrect agent behaviors followed by an incoherent science read action

S3. Desired: IA prior to solution construction or assessment strategy (Science Read => SC|SA)

Suboptimal: Lack of a Science Read action or an incoherent Science Read action before an effective SA action

S4. Desired: Test in parts strategy (Effective Compare action)

Suboptimal: ineffective Compare action

S5. Desired: Conceptual sense-act model building actions followed by coherent computational model building actions (Sense-act build => Computational build)

Suboptimal: incoherent (Sense-act build → Computational build) action sequence, or lack of the action sequence

S1, *S2*, and *S3*, link SC and SA actions to IA actions, implying the usefulness of seeking information about the part of the model they are building or assessing. *S4* describes a strategy for testing the model in parts to help isolate errors. *S5* pertains to SC, and how to effectively use multiple linked representations to build science models. When frequency counts for suboptimal uses of a strategy exceed a predetermined threshold (in the range of 2-5), scaffolds are triggered to provide feedback on use of the strategy. A local history of learners' conceptual and computational modeling skills are maintained by comparing different aspects of their models against the corresponding expert models to detect aspects of the modeling tasks they are struggling with. Separate 'missing' and 'extra' measures are maintained for different conceptual model components like agents, environment elements, properties, and behaviors chosen, as well as the sensed and acted-on properties specified for each agent behavior. Similarly, computational modeling skills are captured in terms of the number of missing and extra blocks, and whether all actions in a behavior occur under the right set of conditions. Students are scaffolded on their modeling tasks if their modeling skills do not improve between successive model assessments.

The task and strategy oriented scaffolds are all delivered in the form of a mixed-initiative conversational dialog initiated by Ms. Mendoza, and linked to students' recent actions and available information (e.g. simulation information or domain information). This conversation format engages students in a more authentic social interaction, and allows them to control the depth and direction of the conversation within the space provided by the dialogue and response choices. Our scaffolding approach helps students with a task or strategy only we detect that they are persistently facing problems, instead of correcting them every time we detect a problem. Scaffolds offer suggestions and reminders of good strategies and help point out possible sources of errors, but never provide 'bottom-out-hints' by telling students exactly what to correct in their models.

Method

We conducted an experimental study with 98 students (average age = 11.5) from four 6th-grade sections in a Tennessee middle school. The science teachers assigned students from two sections to the control group ($n = 46$) which used a version of CTSiM without adaptive scaffolding, and students from the other two sections to the experimental group ($n = 52$) which received adaptive scaffolding.

The study was run daily over a span of three weeks during students' science periods (one hour daily for each section). All students worked on the same learning progression across two domains - Kinematics and Ecology. On Day 1, students took three paper-based tests that assessed their knowledge of (1) Kinematics, (2) Ecology, and (3) CT concepts. On day 2, students were introduced to agent based modeling concepts, and the whole class worked together on an introductory single-agent shape drawing activity using simple CT concepts like iterations. From Day 3, students worked individually. On days 3 and 4, they worked on generating growing and shrinking spiral shapes, which emphasized the relations between distance, speed, and acceleration. This activity was for practice only; students were allowed to seek help from their science teacher or from the research team if they had difficulties. From Day 5, students worked on the three primary modeling activities, and were not provided any individual help external to the system. On days 5 and 6, they worked on the first modeling activity, where they modeled the speed of a roller coaster (RC) car moving along different segments of a track. This required use of more complex CT constructs like conditionals. After Activity 1, students took paper-based Kinematics and CT post-tests on Day 7. On days 8-12, students progressed to modeling multiple agents with multiple behaviors in a fish tank system. In Activity 2, students built a *macro*-level, semi-stable model of a fish tank with two types of agents: fish and duckweed, and behaviors associated with the food chain, respiration, locomotion, and reproduction of these agents. Since the waste cycle was not modeled, the build-up of toxic fish waste caused the fish and the duckweed to gradually die off. In Activity 3, students addressed this problem by introducing micro-level entities, i.e., Nitrosomonas and Nitrobacter bacteria, which complete the waste cycle by converting the ammonia in the fish waste to nutrients (nitrates) for the duckweed. Students took their Ecology and CT-final post-tests on Day 13. Finally, on Day 14, they worked on a paper-based transfer activity where they started with a detailed textual

description of a wolf-sheep-grass ecosystem and constructed conceptual and computational models of the ecosystem using modeling primitives specified in the question. Unlike the CTSiM environment, students did not have access to any of the online resources or tools.

All students' actions in the CTSiM system were logged to answer the following research questions:

1. Do the adaptive scaffolds make students' better science and CT learners?
2. Do the adaptive scaffolds improve students' modeling skills, and do these skills transfer beyond the CTSiM environment?
3. How do the adaptive scaffolds impact students' use of effective and suboptimal strategies?

We measured students' learning gains for kinematics, ecology, and CT by calculating their pre- and post-test scores. The Kinematics test assessed students' understanding of relations between speed, acceleration and distance. Students interpreted and generated speed-time and position-time graphs to explain motion in a constant acceleration field. The Ecology test focused on students' understanding of interdependence and balance in an ecosystem, and how a change in the population of one species in an ecosystem affects the other species. We assessed CT skills by asking students to predict program segment outputs, model scenarios and develop meaningful algorithms using CT constructs like conditionals, loops, and variables.

We assessed students' modeling skills using metrics similar to those used online for determining need for scaffolding (see Section 3). We computed the 'distance' between students' conceptual and computational models and the corresponding expert models in terms of the missing and extra model components, and normalizing it by the size of the expert model (i.e., the sum of the number of elements of each type of model component) to make the 'distance' measure independent of the size of the expert model (Basu et. al., 2014). We described students' modeling progress during an activity by calculating the model distances at each model revision and then characterized the model evolution using 3 metrics: (1) *Effectiveness*- the proportion of model edits that bring the model closer to the expert model; (2) *Slope* – the rate and direction of change in the model distance as students build their models; and (3) *Consistency* – How closely the model distance evolution matches a linear trend. In order to study students' effective uses of strategies, we matched their logged action sequences with the definitions of desired strategies specified in Section 3. Suboptimal uses of the strategies were counted in terms of the frequencies of strategy feedback received by the experimental group and that would be received by the control group. The logged action sequences for students in the control group were matched against the suboptimal strategy variants, similar to the online strategy matching done for students in the experimental condition.

Findings

Science and CT learning gains and modeling proficiency

Our results show that students who received adaptive scaffolding had higher learning gains on both science and CT content. Table 1 summarizes students' pre-post learning gains for kinematics and ecology science content, and CT concepts and skills based on the CT test administered at the end of the study. Students in the experimental group had higher pre-test scores, hence we computed ANCOVAs comparing the gains between control and experimental conditions taking the pre-test scores as a covariate. Both groups had significant learning gains with medium to high effect sizes (Cohen's d), but the gains were higher in each case for students in the experimental group: kinematics gains ($F = 18.91, p < 0.0001, \eta_p^2 = 0.17$); ecology gains ($F = 52.29, p < 0.0001, \eta_p^2 = 0.36$); CT gains ($F = 40.69, p < 0.0001, \eta_p^2 = 0.31$). We also assessed students' performances on the first CT post-test at the end of kinematics unit, and found that students in the experimental group showed higher learning gains from the pre-test to the first post-test ($F = 18.16, p < 0.0001, \eta_p^2 = 0.16$), and gained further from the second to the third CT post-test administered at the end of the ecology unit ($F = 18.85, p < 0.0001, \eta_p^2 = 0.17$).

To answer our 2nd research question, we compared the accuracy (distance to expert-model) of the final conceptual and computational models built by the control and experimental group of students. Figure 3 shows that students in the experimental condition built more accurate conceptual and computational models for all the activities (the final model distance scores were significantly lower) when compared to students in the control condition. Further breaking down the aggregate distance scores revealed that both missing and extra model constructs were significantly lower for the experimental condition for both conceptual and computational models. This implies that the experimental group's models included more model components from the expert model (lower missing score) and fewer redundant and incorrect components (lower extra score) than the control group's models.

In addition to building more accurate final models, the experimental group's progress towards the final conceptual model was significantly better than the control group as evidenced by three metrics: (1) higher percentage of effective (i.e., correct) conceptual edits in all three activities; (2) conceptual model accuracy improved with time in each activity, i.e., the slope for model distance over time was negative, whereas the distance slope

for the control group was positive. (This was because the control group kept adding unnecessary elements to their models, and their conceptual models became more inaccurate in each activity as time progressed); and (3) modeling consistency was higher for the experimental group in the fish-micro unit. Also, the experimental group's computational model progressions within each unit were more consistent and improved more rapidly. Both conditions had negative computational model evolution slopes, i.e., their model accuracy improved over time in each of the activities. However, the rate of improvement was significantly higher for the experimental group in all the activities. We also analyzed students' performances on the transfer task and separately scored their conceptual and computational models of the wolf-sheep-grass ecosystem. We found that students in the experimental condition were better at applying their modeling skills and built more accurate conceptual ($p < 0.0001$, *Cohen's d* = 1.53) and computational ($p < 0.0001$, *Cohen's d* = 1.46) models compared to students in the control condition.

Table 1: Science and CT learning gains for students in the control and experimental conditions

		Pre	Post	Pre-to-post gains	Pre-to-post <i>p</i> -value	Pre-to-post Cohen's <i>d</i>
Kinematics (max = 45)	Control	12.52 (6.32)	15.55 (5.72)	3.03 (4.78)	<0.0001	0.55
	Experimental	16.65 (6.61)	22.38 (6.39)	5.72 (5.62)	<0.0001	0.88
Ecology (max = 39.5)	Control	7.40 (3.90)	16.19 (8.35)	8.78 (7.17)	<0.0001	1.35
	Experimental	9.39 (4.47)	27.91 (6.70)	18.53 (6.31)	<0.0001	3.25
CT (max = 60)	Control	16.49 (5.68)	22.53 (5.70)	6.04 (5.44)	<0.0001	1.06
	Experimental	22.72 (7.68)	32.24 (5.86)	9.52 (5.23)	<0.0001	1.39

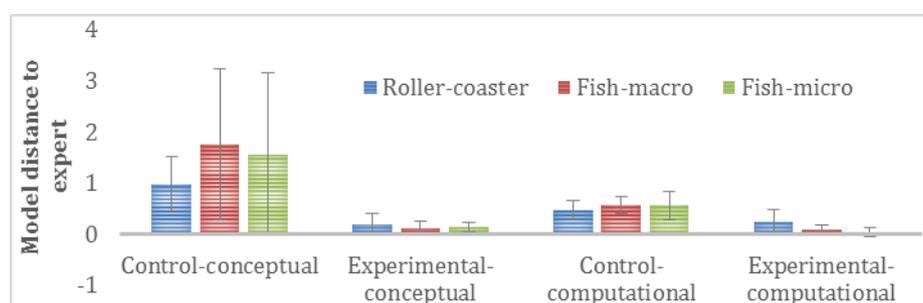


Figure 3. Modeling performance across conditions.

Effective and suboptimal uses of desired modeling strategies

To address our 3rd research question, we first computed the average number of times each of the five strategies was used effectively in each modeling activity, as well as the percentage of students who used the strategy effectively at least once in each activity (see Table 2). We note two general trends: (1) the fraction of students in the experimental group who used the strategies effectively was always greater than or equal to that in the control group, and (2) the average effective use of the strategies was also higher in the experimental group. While most of the differences had low to medium effect sizes (*Cohen's d* in the range of 0.2 to 0.7), the differences in use of the *Model-Build* strategy had much larger effect sizes in all three modeling activities (*Cohen's d* in the range of 1.36 to 1.75). Effective uses of this strategy were also strongly correlated with science learning ($p < 0.0001$).

We also studied the effect of our adaptive scaffolds on students' suboptimal uses of strategies. Since the strategy oriented scaffolds were triggered based on the suboptimal strategy uses, we counted the feedback received in the experimental group and calculated the feedback that would be received by the control group. For each type of strategy feedback, Table 3 provides for each activity: (1) *n*, which represents the number of students who receive the feedback at least once in the activity, (2) *min-max*, which represents the lowest and highest number of times feedback is received by any student during the activity, and (3) *mean (s.d.)* represent the average number of times (and standard deviation) the feedback was received during the activity. We see that the experimental group students need significantly lower amount of strategy feedback than the control group would have needed, especially for the *Model-Build* strategy, the *test-in-parts* strategy, and the *IA-SC/SA* strategy. This implies that the adaptive scaffolds helped improve effective uses of the strategies, and reduced their suboptimal uses.

We also performed more fine grained analysis of effects of the scaffolds on effective uses of strategies by counting the number of effective uses before and after feedback instances. Our results show a general trend for students who needed scaffolding, their effective uses of strategies became more frequent as they received feedback for their suboptimal uses. For example, for *S4* (the test-in-parts strategy) in the fish-macro unit, 10 of the 52

experimental group students never received feedback on *S4* and made 0.8(1.5) effective uses of *S4* on an average. 15 students received feedback exactly once, and made an average of 2.0 (4.7) partial model comparisons before receiving feedback, which increased to 2.73(6.24) after receiving feedback. The other 27 students received feedback on *S4* two or more times; they used *S4* an average of 0.93(2.4) times before receiving any feedback, 1.93(4.2) times between the first and second feedback instances, and 4.7(7.43) times after receiving feedback twice.

Table 2: A comparison of the use of desired strategies across conditions (Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

Strategy		RC		Fish-macro		Fish-micro	
		Fraction of students	Mean (s.d.)	Fraction of students	Mean (s.d.)	Fraction of students	Mean (s.d.)
S1. SC action followed by relevant science reads	C	37%	1.33 (2.99)	54%	2.43 (4.8)	70%	1.93 (2.05)
	E	63%	2.23 (4.71)	83%	4.75 (4.97)*	85%	3.4 (4.51)*
S2. SA actions followed by relevant science reads	C	4%	0.07 (0.33)	26%	0.76 (1.66)	26%	0.85 (9.31)
	E	38%	1.37 (2.69)**	44%	1.66 (2.29)*	44%	1.06 (0.24)
S3. Fraction of assessed behaviors that were read about before being assessed	C	80%	.73 (.42)	93%	.5 (.33)	83%	0.89 (0.27)
	E	92%	.86 (.28)	96%	.77 (.32)***	100%	0.96 (0.16)
S4. Number of partial-model comparisons	C	0%	na	48%	2.65 (5.79)	15%	0.57 (1.98)
	E	0%	na	58%	5.42 (7.16)*	19%	1.97 (3.22)*
S5. Fraction of sense-act properties removed or followed by a coherent computational edit	C	100%	0.67 (0.27)	100%	0.69 (0.31)	98%	0.59(0.31)
	E	100%	0.97 (0.1)***	100%	0.99 (0.03)***	100%	0.98 (0.06)***

Table 3: Comparing suboptimal uses of strategies in terms of feedback received or would be received (Note: * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0001$)

		RC			Fish-macro			Fish-micro		
		<i>n</i>	Min-max	Mean(s.d.)	<i>n</i>	Min-max	Mean(s.d.)	<i>n</i>	Min-max	Mean(s.d.)
S1: SC-IA strategy	C	4	0-1	0.09 (0.28)	19	0-28	1.93 (4.46)	15	0-7	1.13 (1.98)
	E	3	0-1	0.06 (0.2)	22	0-4	0.69 (1.02)	8	0-1	0.15 (0.36)**
S2: SA-IA strategy	C	0	0	0(0)	0	0	0(0)	3	0-10	0.28 (1.5)
	E	0	0	0(0)	0	0	0(0)	0	0	0(0)
S3: IA-SC/SA strategy	C	16	0-57	8.43 (15.8)	41	0-62	18.8 (15.7)	8	0-14	1.11 (2.94)
	E	18	0-15	1.37 (3.11)**	19	0-14	1.81 (3.27)***	4	0-3	0.13 (0.52)*
S4: Test-in-parts strategy	C	0	0	0(0)	46	1-26	9.57 (6.77)	37	0-30	3.85 (5.27)
	E	0	0	0(0)	42	0-9	2.23 (2.13)***	23	0-6	0.83 (1.26)***
S5: Model-Build strategy	C	41	0-32	7.17 (6.19)	45	0-130	34.83 (28.87)	36	0-150	18.85 (26.95)
	E	32	0-8	1.79 (2.17)***	35	0-10	2.04 (2.43)***	30	0-10	1.33 (1.82)***

Finally, we also looked at how the task and strategy based scaffolds needed by the experimental group varied with time. We found that students needed a combination of task and strategy feedback in all the activities. In the initial RC activity, students received more task-oriented feedback than in the other two activities. In the more complex fish-macro activity, students needed more strategy feedback than in the RC activity, but less task feedback than in the RC activity, implying that the effects of the task feedback persisted across units. However, students found it challenging to manage and integrate the different tasks in a complex modeling activity involving a new domain. Finally, in the fish-micro activity, the task feedback received was further reduced, and the strategy feedback also decreased (to a smaller number than in the initial RC activity). This provides preliminary evidence that our scaffolding effects persisted, and, therefore, a fading effect occurred naturally as students worked across units. Further, the conceptual and computational models in the fish-micro activity were the most accurate of any activity, even though the students received less feedback in each category of scaffolds than in the earlier activities.

Discussion and conclusions

In this paper, we have presented a generalized adaptive scaffolding framework for OELEs and a specific example of its application in the CTSiM environment. Learning in CTSiM is based on an iterative model-test-refine cycle with a well-defined goal state. Representing a science topic using conceptual and computational representations helps students understand the science concepts underlying the topic, and offers them a chance to iteratively refine their understanding of the science concepts as they refine their models. Hence, our scaffolding approach does not provide students with the correct model at any point. Unlike several learning-by-modeling environments for science, which provide students with hints on incorrect relationships modeled, we combine students' modeling behavior and performance to determine their need for scaffolding. A study run with control (no adaptive scaffolding) and experimental (adaptive scaffolding) conditions demonstrates the effectiveness of our approach. The experimental group scored higher than the control group on science and CT assessments which were designed to test students' understanding of science processes and their reasoning and problem-solving skills as opposed to rote or inflexible knowledge. They also outperformed the control group in the ability to construct correct models, frequent use of effective strategies and infrequent use of suboptimal strategies. Further, we noticed a fading effect of our scaffolds - students in the experimental condition required less scaffolds across activities.

This work also contributes to the field of CT in K-12 education where few successful systems have been developed, assessments of students' learning are lacking, and scaffolds are limited to automatic assessments of students' computational artifacts based on the CT primitives and patterns contained. Our work provides an example of how CT principles can be operationalized and integrated with science curricula, and how scaffolds contextualized in science content can help students learn important CT concepts like sequences, loops, conditionals, and variables, and become more proficient in vital CT practices like decomposing complex tasks, testing and debugging, and abstracting and modularizing. Our results demonstrate significant correlations ($p < 0.05$) between CT and science learning gains, as well as between important CT practices and science learning (Basu et. al., 2016). As future work, we plan to continue verifying our scaffolding framework with different OELEs and also in CTSiM with a more comprehensive set of strategies and global measures of students' performance and behaviors.

References

- Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, 40(4):199–209.
- Basu, S., Biswas, G. & Kinnebrew, J.S. (2016). Using multiple representations to simultaneously learn computational thinking and middle school science. In Thirtieth AAAI conference on Artificial Intelligence. Phoenix, Arizona, USA.
- Basu, S., Dukeman, A., Kinnebrew, J., Biswas, G., & Sengupta, P. (2014). Investigating student generated computational models of science. In *Proceedings of the 11th International Conference of the Learning Sciences* (pp. 1097-1101). Boulder, CO, USA.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of research in education*, 61-100.
- Duque, R., Bollen, L., Anjewierden, A., & Bravo, C. (2012). Automating the Analysis of Problem-solving Activities in Learning Environments: the Co-Lab Case Study. *J. UCS*, 18(10), 1279-1307.
- Land, S. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*, 48(3), 61-78.
- Luckin, R. and du Boulay, B. (1999). Ecolab: The development and evaluation of a vygotskian design framework. *International Journal of Artificial Intelligence in Education*, 10(2):198–220.
- Mitrovic, A.: 2011, 'Fifteen years of Constraint-Based Tutors: What we have achieved and where we are going'. *User Modeling and User-Adapted Interaction*.
- Puntambekar, S. and Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, 40(1):1–12.
- Segedy, J.R., Kinnebrew, J.S., and Biswas, G. (2015). Using coherence analysis to characterize self-regulated learning behaviours in open-ended learning environments. *Journal of Learning Analytics*, 2(1), 13–48.
- Sengupta, P., Kinnebrew, J.S., Basu, S., Biswas, G., & Clark, D. (2013). Integrating Computational Thinking with K-12 Science Education Using Agent-based Computation: A Theoretical Framework. *Education and Information Technologies*, 18(2), 351-380.
- Wilensky, U. (1999). *NetLogo*. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. (<http://ccl.northwestern.edu/netlogo>).

Acknowledgments

We thank participating teachers and students. This work funded by NSF Grant #1441542.

Negotiation Towards Intersubjectivity and Impacts on Conceptual Outcomes

Catherine L. Dornfeld and Sadhana Puntambekar
cldornfeld@wisc.edu; puntambekar@education.wisc.edu
University of Wisconsin–Madison

Abstract: To achieve a collective goal, students collaborating in groups must share their interpretations of the goal to negotiate a shared understanding, or establish intersubjectivity, before making progress towards that goal. To investigate negotiation towards intersubjectivity along with students' conceptual outcomes, we studied the discourse of two groups of middle-school students who participated in a 12-week science curriculum. We also evaluated differences in scores on a physics assessment. We found differences in how groups negotiated shared understandings of tasks and concepts and how they participated in conceptual discourse. Group A required significantly more instances of negotiation to establish intersubjectivity compared to Group B, indicating that Group A struggled with establishing shared understandings. Students in Group A also demonstrated greater variance in conceptual discourse and assessment scores, whereas Group B demonstrated more aligned conceptual outcomes. These results indicate that effective negotiation for establishing intersubjectivity is not only an important first step for students' participation in conceptual discourse, but also for achieving balanced learning gains across group members.

Keywords: collaboration, intersubjectivity, conceptual discourse, learning outcomes, CSCL

Introduction

Small groups participating in coordinated activities must establish intersubjectivity, or shared understanding of the current activity; this often occurs through group discourse (Rogoff, 1990; Wertsch, 1979; Jarvela, 1995; Rowe, 2011). Upon initiating group work, each group member may have different interpretations of the shared goal and its sub-components, even when group members receive identical instructions, content, and/or materials. As a result, each group member may describe the current activity differently in terms of the goal itself, necessary tasks, relevant domain knowledge, intentions, and affective responses (Rummel, Dieglmayr, Spada, Kahrimanis, & Avouris, 2011; Jarvela, 1995). For example, students collecting physics data with a rollercoaster simulation may have different understandings of physics concepts; different ways to use classroom tools to collect, record, and organize task-related information; and different connections between collected data, initial hypotheses, and final conclusions. In this example, we can see multiple opportunities for different interpretations – and thus a need to negotiate shared understandings. Different interpretations of the goal and its components necessitate dedicated discourse for establishing intersubjectivity, particularly during collaborative decision-making (Barron, 2003). Students must reveal their reasoning through sharing, negotiating, and jointly creating understandings of the goal in order to make progress.

Group members' interpretations or understandings may be described as *situation definitions* (Wertsch, 1984). During group work, students actively create mental representations of the current situation or activity. However, different students may have different representations. Differences among representations may arise from differences in individual zones of proximal development (ZPD) within the group, especially when considering prior knowledge (Vygotsky, 1978; Wertsch, 1984). Students may define and decompose a situation (e.g., goal, task, or concepts) differently based on their actual levels of development. Additionally, situation definitions are located in time and subject to change. Students may alter and adapt their representations as they encounter others' perspectives and co-construct knowledge (Park & Moro, 2006). These evolving situation definitions may be described as *fluctuations*, which are developed as co-construction "through the process of conflict, negotiation, and renegotiation about the activity during interactions" (Park & Moro, 2006, p. 113) occurring as utterances over time (Rowe, 2011; Hall, 2011; Rommetveit, 1976). Fluctuations help the group to overcome rigid (and possibly inaccurate) representations. They also allow the group to develop a collective situation definition for the activity based on their negotiated understandings.

As group members share and iterate on their situation definitions, they also engage in negotiation of shared conceptual understanding. Collaborative learning involves mutual construction and negotiation of cognition through interactions, such as patterns in discourse (Roschelle, 1992; Rummel et al., 2011; Barron, 2003). Group members collaboratively co-construct meaning as they negotiate conceptual situation definitions. Through

a progression of contributing and listening, group members build mutually shared cognition (Baker, 1994; Miyake & Kirschner, 2014; Barron, 2003; Roschelle, 1992). Beyond ensuring that students have similar conceptual understandings, mutually shared cognition is also associated with greater group effectiveness (Miyake & Kirschner, 2014).

In this study, we investigated how two small groups engaged in negotiation of situation definitions of tasks and concepts as they participated in a science curriculum. We used a comparative case-study approach to investigate two research questions: *How did negotiation of situation definitions move students towards establishing intersubjectivity, and how did negotiation towards intersubjectivity impact conceptual outcomes over time?* By investigating these questions, we aim to reveal potential relationships between collaborative group dynamics and meaning-making outcomes. We also aim to identify opportunities for interventions within group discourse that increase the likelihood of greater learning outcomes for *all* group members.

Methods

To study negotiation of situation definitions and students' conceptual understandings, we examined the discourse of two small groups over a 12-week science curriculum. The groups participated in the CoMPASS project (Puntambekar, Stylianou, & Goldstein, 2007), which investigated how students developed science literacy as they interacted with digital textbooks and other distributed scaffolds. CoMPASS was designed as a 12-week design-based physics curriculum that applied physics concepts such as forces, work, energy, and motion to roller coaster design. Students collaborated in small groups and used a computer simulation, a digital textbook (CoMPASS), and scientist's journals throughout the unit. The experiments involved students' manipulation of variables within the simulation to discover relationships between concepts. In this study, we examine how students negotiated shared understandings of goals, activities, and concepts during four of the experiment sessions using a combination of cognitive ethnography and comparative case studies (Puntambekar, 2013; Barron, 2003).

Participants

For our two cases, we selected two groups of four sixth-grade students ($N=8$) from the larger study sample. These groups were selected because they were from the same classroom, instructed by the same teacher, and attended the same CoMPASS sessions. Participants also had similar prior knowledge based on pre-test scores (see Data Sources). Their class consisted of 22 students divided into six groups. Group A's students included Simon, Ben, Ali, and Mallory. Group B's students included Zach, Lucas, Morgan, and Ana. They attended a large suburban public middle school in the U.S. Midwest.

Data sources

We collected video and audio data of the groups' experiment sessions over a 12-week curriculum (14 sessions total). We selected four sessions to analyze discourse over time: two sessions (1 & 3) from the beginning of the unit, and two (13 & 14) from the end of the unit. All group members were present for each session.

We also used a test designed by the project that assessed understanding of physics concepts. This test was administered prior to and after the 12-week implementation. The test consisted of 29 items related to relationships between forces, work, energy, friction, and Newton's Laws. Correct answers received one point while incorrect answers received zero points; the maximum score was 29 points. The groups' pre-test scores were not significantly different, indicating that students in both groups had similar prior knowledge.

Analysis

Qualitative analysis

Overall, students in both groups contributed 1579 total turns of talk over the four sessions. After reviewing several sessions, we designed a grounded coding scheme that focused on conceptual and procedural and negotiation in small-group discourse. This two-dimensional coding scheme (see Table A1 in Appendix) focused on how students established common ground through negotiation in discourse. Each turn of talk was coded for content (Purpose Code) and progression towards shared understanding, or intersubjectivity (Negotiation Code). Turns that did not fit these categories were coded as Off-Task (19.7% of turns) or N/A (7.0% of turns). Multiple codes were permitted for turns of talk. For inter-rater reliability, we achieved a Cohen's kappa value of 0.904 with an external coder for a subset of the data. Differences were resolved through further discussion.

Quantitative analysis

We used quantitative analyses to compare differences between and within groups, using turns of talk as our unit of analysis. We calculated frequencies and proportions of coded turns of talk for groups and individual students.

To compare Groups A and B, we performed two-tailed tests of homogeneity for proportions of coded turns (critical z -score = ± 1.96 , $\alpha = 0.05$). To study differences over time, we used nonparametric chi-square tests to compare Purpose (2 x 4) and Negotiation (3 x 4) coded contributions over the four sessions for each group. To assess conceptual outcomes, we compared proportions of individual students' contributions to conceptual talk over the four sessions. We also calculated learning gains on the physics conceptual assessment (as differences in pre-post scores), along with the mean, standard deviation, and range of learning gains for each group to understand how groups' conceptual understandings developed over time.

Findings

Group differences in talk and changes over time

To see how the two groups negotiated situation definitions, we compared both groups' coded contributions to talk over four sessions. Table 1 shows that Group A contributed significantly more suggestions (0.124 vs. 0.094, $z = 2.607$, $p = 0.009$), agreements (0.043 vs. 0.025, $z = 2.778$, $p = 0.005$), and disagreements (0.046 vs. 0.029, $z = 2.485$, $p = 0.013$) than Group B over the unit. Group A engaged in cycles of suggesting and disagreeing with ideas significantly more often than Group B. In contrast, Group B engaged in more off-task talk (0.136 vs. 0.075, $z = -5.366$, $p < 0.001$) than Group A over the unit. Considering that both groups had the same amount of time for experiment sessions, Group A may have struggled to establish shared understandings, while Group B readily established shared understandings and thus had time available to engage in unrelated talk. Overall, Group A engaged in negotiation of situation descriptions significantly more often than Group B.

Table 1. Group comparison z -scores for proportions of talk. Positive values indicate greater contributions from Group A, while negative values indicate greater contributions from Group B. *Significant result at $p < 0.05$.

Session	Purpose Code		Negotiation Code			
	Task	Learning	Suggesting	Agreeing	Disagreeing	Off-Task
1	0.138	2.434*	-0.814	3.430*	-2.022*	-2.764*
3	-0.616	-0.223	-0.290	0.084	2.306*	2.223*
13	2.111*	0.548	3.526*	1.586	2.497*	-6.138*
14	-0.373	1.170	2.018*	0.497	2.436*	-2.059*
Overall	0.671	1.427	2.607*	2.778*	2.485*	-5.366*

We also evaluated contributions of talk over time. We used chi-square tests to detect session-based differences in groups' turns of talk related to Purpose (2 x 4 test) and Negotiation (3 x 4 test). Both groups contributed significantly more task-based and learning talk between sessions over time (Group A: $\chi^2 = 9.919$, $p = 0.0193$; Group B: $\chi^2 = 14.166$, $p = 0.0027$), indicating that both groups engaged in more turns of talk more over time. Also, both groups demonstrated significant differences in the frequency of suggestions, agreements, and disagreements over time (Group A: $\chi^2 = 16.718$, $p = 0.0104$; Group B: $\chi^2 = 16.928$, $p = 0.0096$), indicating that the types of their contributions varied significantly over time.

Conceptual understanding over time

This unit was designed to facilitate learning of concepts through collaboration and interaction with distributed scaffolds. We would ideally see improved conceptual outcomes for all group members (i.e., all students benefited from the unit). To assess conceptual outcomes, we studied group members' contributions to conceptual discourse over four experiment sessions (as seen in Table A2) along with performance on the conceptual assessment. We calculated pre-post differences for each student, or learning gains, in assessment scores along with the mean, standard deviation, and range of learning gains for each group.

To see how groups' contributions to learning-based discourse varied over the unit, we examined differences between group members' proportions of conceptual discourse (from highest to lowest proportions). Figures 1 and 2 (below) show individual students' contributions to conceptual discourse over time. We see that Group A demonstrated greater differences between students ranging from 0.041 (Session 1) to 0.114 (Session 14; see Table A3 in Appendix for all differences). Group A's contributions to conceptual discourse were highly varied between students over time, indicating that some group members participated in conceptual discourse more often than others. In contrast, Group B demonstrated differences between students ranging from 0.000 (no difference; Session 1) to 0.054 (Session 14). This indicates that students in Group B contributed relatively similar levels of conceptual discourse over time than Group A.

Overall, students in Group B were more aligned in their individual contributions to conceptual discourse than students in Group A. This fits with both groups' learning gains on the conceptual assessment. Table 2 shows that Group A's scores showed greater variance ($SD = 5.377$, range = 13), while Group B's showed less variance ($SD = 1.291$, range = 3). Overall, students in Group A demonstrated greater variance in conceptual outcomes while Group B demonstrated relatively similar conceptual outcomes.

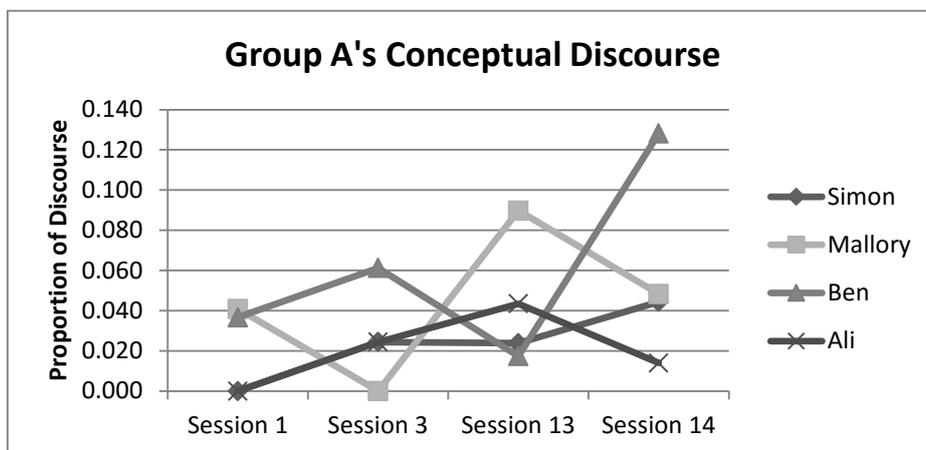


Figure 1. Conceptual discourse contributions (as proportions of talk) by students in Group A.

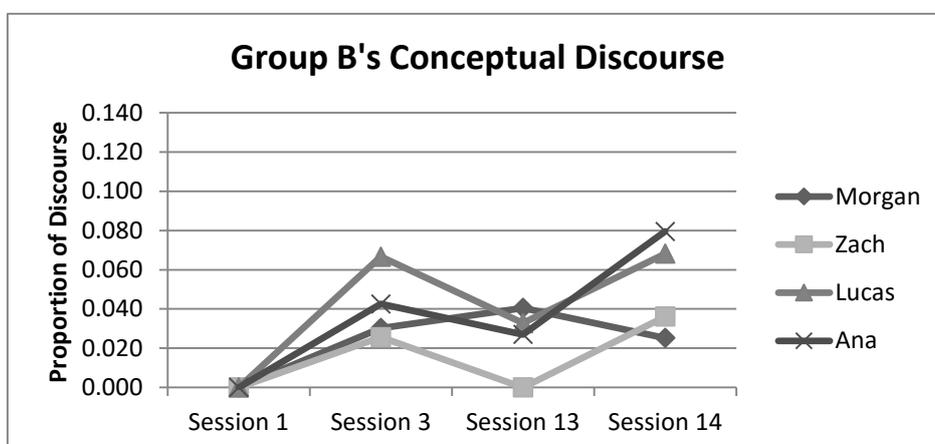


Figure 2. Conceptual discourse contributions (as proportions of talk) by students in Group B.

Table 2. Student learning gains and descriptive statistics for the physics conceptual assessment (max. score = 29 points).

Group	Student	Pre-test Score	Post-test Score	Learning Gains	Range	Mean Gain	Standard Deviation
A	Simon	17	10	2	13	2.75	5.377
	Mallory	18	20	2			
	Ben	13	23	10			
	Ali	18	15	-3			
B	Morgan	15	21	6	3	5.50	1.291
	Zach	14	21	7			
	Lucas	12	17	5			
	Ana	16	20	4			

Here we present excerpts of conceptual discourse as qualitative evidence of group differences. Group A's discourse frequently involved negotiation as initial suggestions followed by disagreements and/or several alternative suggestions. In this excerpt, Group A reached a standstill in their decision for a rollercoaster simulation

variable. The goal for this activity was to stop a rollercoaster car in a way that was quick and efficient but did not put excessive force on the riders. Mallory and Ben had different interpretations of the lesson goal; Mallory prioritized shorter stopping distance, while Ben prioritized rider safety. They attempted to explain their reasoning to each other, but their conversations were cut short by Ali and Simon, who prioritized completing the task over reasoning through debate (turns 5 and 10).

- Ali: I think we should do 3 anyway.
- Ben: I think we should do 2.5. That works.
- Mallory: It's not going to help if we take a vote. It's going to be two-on-two.
- Ben: 'Cause 3 just stopped it too fast. No, no. It's not fair. It's not scientific-y.
- Ali: Let's just do it.
- Ben: No, it's not fair. Why do you guys think it's 3?
- Ali: Because, um, it takes, um, less [inaudible] and less track to stop it, and it's still safe and efficient.
- Ben: 2.5 was more efficient. Uh, 2.5 is more safe. And it still doesn't take that much track.
- Ali: It took --
- Simon: Let's just do this right now and worry about that when we're at the simulation.

In contrast, Group B's discourse typically involved brief instances of negotiation in which students made decisions based on earlier experiments. In this excerpt, students in Group B considered different rollercoaster hill height values, car masses, and friction levels for a complex simulation experiment with multiple input variables. We see evidence of students actively listening to suggestions and repeating them to confirm. Students were able to finish utterances and justify ideas without interruption.

- Teacher: So if this is your hill height, you don't have a -- you can still change the value. You just don't have a ton of wiggle room.
- Lucas: 80, 85, and 90.
- Morgan: For this you can do 0.2, 0.4, 0.6.
- Ana: Yeah, for our car masses we're doing 0.2, 0.4, 0.6.
- Lucas: 0.2 K, 0.4, 0.6 K.
- Ana: I think our friction level should be four.
- Zach: What page is that on?
- Morgan: Well, for our friction level we agreed because that was the one that was, like, the most safe.

Overall, we see that Group A's discourse included unproductive debate and frequent interruptions during explanations, which resulted in incomplete reasoning in decision-making. Group B's discourse included repetition of suggestions to indicate confirmation and productive debate involving active listening and full explanations. While both groups debated how to set up their experiments, Group B's communication and negotiation practices facilitated clarity in suggestions, explanations, and decision-making over the unit.

Discussion

In this study, we investigated how students in small groups negotiated shared understandings, or established intersubjectivity, through their discourse moves. We focused on two questions: *How did negotiation of situation definitions move students towards establishing intersubjectivity, and how did negotiation towards intersubjectivity impact conceptual outcomes over time?* We found that Group A engaged in significantly more negotiation discourse than Group B. Group A required more negotiation of tasks and relevant concepts than Group B, which

may be explained by Group B's better externalization and negotiation habits, such as sharing mental models of concepts and actively listening. Also, Group B demonstrated relative ease in reaching consensus on shared situation definitions compared to Group A; thus, they did not require as much negotiation. Overall, Group B established intersubjectivity as shared situation definitions more effectively than Group A.

We evaluated group meaning-making by examining contributions to conceptual discourse. Students in Group A contributed to conceptual discourse in highly varying levels over the unit, while students in Group B contributed to discourse at similar levels over the unit. Group A's varied levels of conceptual discourse aligned with their varied learning gains; Ben gained 10 points from pre- to post-test, while Ali lost 3 from pre- to post-test. Group B's similarity in conceptual discourse contributions aligned with their similar learning gains; Ana gained 4 points while Zach gained 7 points. We believe that Group A's increased instances of negotiation may have negatively impacted their developing conceptual understandings; they never seemed to be "on the same page." In 50-minute classroom sessions, increased time spent on debating situation definitions, especially as unproductive argumentation, limited engagement in conceptual discourse and resulted in differential learning gains for group members. However, students who effectively negotiated shared understandings of the task along with relevant concepts showed similar conceptual outcomes for all group members, possibly explained as reaching mutually shared cognition.

Another lens for study could investigate social dynamics, such as control, in each group. One student maintained control of the shared computer in Group A, while the students in Group B rotated control of the computer. Control over resources can impact adoption of discrete roles, such as technology manager, and alter participation in discourse (Dornfeld & Puntambekar, 2015). This could also impact negotiation of actions and shared understandings. Following studies would involve investigation of relationships between control of shared resources and constraints on negotiation towards intersubjectivity. Students who control the computer take on a role with power and responsibilities, which may impact group dynamics and participation in discourse. Additionally, examination of specific sequences of discourse moves may provide insight into productive (or unproductive) negotiation for establishing intersubjectivity. Sequential pattern analysis of coded turns of talk may reveal ideal or problematic negotiation patterns as intervention targets. For example, if teachers observe a known problematic sequence, such as repetitive arguments, they may intervene and help students to externalize situation definitions and reach consensus. Finally, we plan to develop interventions that evenly distribute and increase learning gains for group members.

The main limitation to this study is the small sample size of students ($N = 8$) and limited diversity in the sample, which decreases generalizability to other populations. However, investigation of this small sample permitted fine-grained analysis of group discourse over the unit. Replication of this study with a larger sample could provide evidence to support our finding that effective negotiation towards intersubjectivity may help group members achieve similar conceptual gains.

Implications and conclusion

This study indicates that establishing shared situation definitions, or intersubjectivity, is a precursor to optimal conceptual outcomes for group collaboration. Collaboration necessitates that students negotiate shared representations of their goals and relevant concepts. In this science curriculum, students needed to share interpretations of their collective goal and its sub-components to effectively make decisions, carry out experiments, and understand the conceptual relationships at work. The theoretical implication of this study involves the importance of intersubjectivity in achieving similar conceptual understanding for group members. This study reveals another piece of how collaboration processes vary between groups and impact how students' conceptual outcomes. Group negotiation of shared understandings is not always easy. For example, students with different zones of proximal development may define situations very differently and thus require additional negotiation towards shared understandings. For practical implications, this study allows identification of opportunities for instructor-based interventions involving effective negotiation of situation definitions. For example, teachers may model ideal discourse practices or class norms. Also, teachers can monitor groups for ineffective discourse patterns (such as when groups appear to be "stuck"), and then intervene with suggestions or scaffolds for building intersubjectivity and shared meaning-making.

Overall, negotiation towards intersubjectivity appears to have consequences for students' conceptual outcomes when participating in small-group work. Groups that are able to co-construct shared situation definitions may show similarities in understanding domain areas, such as physics concepts. However, groups with ineffective negotiation practices may fail to co-construct shared situation definitions and show differential learning outcomes. If we aim for balanced collaborative learning outcomes in small-group work, we must identify and encourage effective negotiation practices, such as externalization and meaning-making strategies within group discourse.

References

- Baker, M. (1994). A model for negotiation in teaching-learning dialogues. *Journal of Artificial Intelligence in Education*, 5(2), 199-254.
- Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12, 307-359.
- Dornfeld, C. & Puntambekar, S. (2015, June). Emergent roles and collaborative discourse over time. In O. Lindwell, P. Hakkinen, T. Koschman, P. Tchounike, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: Computer-Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (pp. 380-387). Gothenburg, Sweden: The International Society of the Learning Sciences.
- Hall, B.M. (2011). *Digging in: Designs that develop intersubjectivity in course room discourse*. Proceedings of the 34th annual international convention of the Association for Educational Communications and Technology (Vol. II, p. 384-387). Bloomington, IN: AECT.
- Jarvela, S. (1995). The cognitive apprenticeship model in a technologically rich learning environment: Interpreting the learning interaction. *Learning and Instruction*, 5, 237-259.
- Miyake, N. & Kirschner, P. A. (2014). The social and interactive dimensions of collaborative learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Park, D. & Moro, Y. (2006). Dynamics of situation definition. *Mind, Culture, and Activity*, 13(2), 101-129.
- Puntambekar, S. (2013). Mixed methods for analyzing collaborative learning. In C.E. Hmelo-Silver, C. A. Chinn, C. Chan, & A. M. O'Donnell (Eds.), *The International Handbook of Collaborative Learning*. New York, NY: Routledge
- Puntambekar, S., Stylianou, A., & Goldstein, J. (2007). Comparing classroom enactments of an inquiry curriculum: Lessons learned from two teachers. *The Journal of the Learning Sciences*, 16(1), 81-130.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, NY: Oxford University Press.
- Rommetveit, R. (1974). On the architecture of intersubjectivity. In L. Strickland (Ed.), *Social Psychology in Transition*. New York. New York: Springer US.
- Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences*, 2, 235-276.
- Rowe, S. (2004). Discourse in activity and activity as discourse. In R. Rogers (Ed.), *An Introduction to Critical Discourse Analysis in Education* (1st ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Rummel, N., Dieglmayr, A., Spada, H., Kahrimanis, G., & Avouris, N. (2011). Analyzing collaborative interactions across domains and settings: An adaptable rating scheme. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing Interactions in CSCL* (p. 367-390). New York, NY: Springer US.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wertsch, J. V. (1979). From social interaction to higher psychological processes. *Human Development*, 22, 1-22.
- Wertsch, J. V. (1984). The zone of proximal development: Some conceptual issues. *New Directions for Child and Adolescent Development*, 1984(23), 7-18.

Acknowledgments

We thank participants in the CoPASS project along with Nicole D. Martin for her assistance with this study. This work was supported by the Bill & Melinda Gates Foundation, EDUCAUSE NGLC grant, and NSF SAVI grant #1258471.

Appendix

Table A1: Two-Dimensional Coding Scheme for Negotiation of Procedural and Conceptual Understanding

Purpose Code	Description	Example
Task-based (Procedural)	Statement or question concerned with shared task goals, task completion, or procedural decisions (i.e., completing workbook pages)	"We have to do the height."
Learning-based (Conceptual)	Statement or question concerned with shared understanding of concepts or meaning, science content, or conceptual decisions	"Do you think PE will always be the same?"
Negotiation Code	Description	Example
Questioning	Asks for others' observations, opinions, hypotheses, or other information	"What are we doing?"
Reporting	Makes a statement about observations, opinions, hypotheses, or other information	"0.49 J" "Whoa, it's bigger."
Suggesting	Gives a suggestion for the next group action or how a concept can be explained	"Let's do the height." "Okay, and then we hit play."
Agreeing	Agrees with description or interpretation of task/concept through affirmatives, repetition of descriptions, or expansion of descriptions of tasks/concepts by extending the line of thinking (i.e., finishing another's thought)	"Okay." "Good."
Disagreeing	Disagrees with description or interpretation of task/concept through short statements or with explanations of why the suggestion isn't a suitable course of action	"No, but you have to change that. Change that."
Deciding	Decides on a course of action that is carried out by the group	"Whoa whoa whoa whoa. Okay, I want to make sure. Oh wait, we need to play it again."
Off-Task	Discusses topics not related to the task or physics concepts	"I have French and gym."
N/A	Unintelligible, inaudible, random noises, or unclear meaning	"So..."

Table A2. Students' contributions to conceptual discourse over time (as proportions of overall talk).

Session	Group A				Group B			
	Simon	Mallory	Ben	Ali	Morgan	Zach	Lucas	Ana
1	0.000	0.041	0.037	0.000	0.000	0.000	0.000	0.000
3	0.024	0.000	0.061	0.024	0.030	0.026	0.067	0.043
13	0.024	0.090	0.017	0.043	0.041	0.000	0.033	0.027
14	0.044	0.048	0.128	0.014	0.025	0.036	0.068	0.079

Table A3. Differences between highest and lowest proportions of conceptual discourse contributed by individual students within each session.

Session	Group A			Group B		
	Highest	Lowest	Difference	Highest	Lowest	Difference
1	0.041	0.000	0.041	0.000	0.000	0.000
3	0.061	0.000	0.061	0.067	0.026	0.041
13	0.090	0.017	0.073	0.041	0.000	0.041
14	0.128	0.014	0.114	0.079	0.025	0.054

Supporting Elementary Students' Science Learning Through Data Modeling and Interactive Mapping in Local Spaces

Kathryn A. Lanouette, Sarah Van Wart, and Tapan S. Parikh
kathryn.lanouette@berkeley.edu, vanwars@berkeley.edu, parikh@berkeley.edu
University of California, Berkeley School of Education and School of Information

Abstract: Our work is based on three premises: (a) children's experiential, everyday knowledge about local spaces is a rich resource for science learning, (b) involvement in data modeling contributes to children's conceptual understanding, and (c) interactive maps can support children in leveraging their experiential, everyday knowledge about local spaces into their scientific reasoning. In this paper, we report on the results of a design-based research study, in which 4th grade students used a web-based tool called *Local Ground* to collect data, model, and collaboratively analyze what's underground in the local soil ecology around their school. We examine the affordances of situating children's data modeling in local spaces and using *Local Ground* as a representational medium. We find that this instructional design supported children in integrating their experiential knowledge to reason about organisms in relation to the environment.

Keywords: elementary science education, ecological systems, data modeling, data visualization, GIS

Introduction

Creating, testing and revising models are central to science (Giere, 1997), reflected in a shifting emphasis within science education research (Lehrer & Schauble, 2015) and recent consensus documents (NRC, 2012). Yet constructing and interpreting representational forms is challenging, with extensive literature documenting children's difficulties in moving between the world and its symbolized forms. Children's familiarity with the phenomena in question is an important but often overlooked element of deciphering representations, wherein attaining a "fluid" reading involves bidirectional movement between phenomena and various symbolic forms (Roth, Pozzer-Ardenghi, & Han, 2005; Latour, 1999).

Existing science instruction, especially at the elementary school level, makes this representational interpretation process more challenging in several ways. First, the phenomena of study is often decontextualized (Metz, 2008) and separated from children's everyday lived experiences (Rivet & Krajcik, 2008), with science instruction seldom engaging children's local spaces as sites for scientific inquiry. As a result, children's related knowledge sources are rarely utilized, missing a key opportunity to integrate children's extensive everyday forms of knowing with more scientific ones (Vygotsky, 1978). Second, children are often presented with "final-form" representations, offering a distorted view of the development and purpose of representational models in science and limiting children's ability to participate in data transformation processes. Children's engagement in the modeling process, through which certain aspects of the phenomena are amplified while others are reduced (Latour, 1999), is crucial to both supporting children's conceptual understanding (Lehrer & Schauble, 2012) and to considering uncertainty of the data (Metz, 2004). Yet in working with increasingly abstracted forms, representations inherently lose the locality and materiality of the phenomena, making it challenging for children to connect these abstracted forms of data to the original phenomena and immediate context.

Our work is based on the following design principles: (a) local spaces can serve as rich resources in children's science learning, (b) children's involvement in data modeling can play an integral part in supporting children's conceptual understanding, and their ability to raise questions about the scientific process and its results, and (c) interactive digital maps can support children in leveraging their experiential, everyday knowledge about local spaces into their reasoning about relationships between organisms and their environment.

Within the science education research field, recent generative work has used local spaces as a context for children's data modeling (Lehrer & Schauble, 2012; Manz, 2012, 2015), with a handful of projects using paper mapping as part of elementary children's sensemaking. Other projects have used interactive digital maps, often for young people to reflect on social issues in their community (Taylor & Hall, 2013; Van Wart, Tsai & Parikh 2010; Enyedy & Mukhopadhyay, 2007). However, this work has not considered the specific cognitive affordances of spatial representations (i.e. maps) in supporting children's reasoning. Additionally, while some of this work has centered on data modeling within children's school or neighborhood environments, it hasn't looked specifically

at how children's local knowledge of phenomena is integrated into scientific reasoning. Moreover, most of this work has been done with older students, usually high-school level.

Work in the emerging area of Citizen Science has also investigated how to expand participation (especially local) in scientific inquiry. Much of this work has focused on citizens as data collectors - leveraging their local knowledge and access to contribute to aggregated data sets that support larger-scale analysis and reasoning (Bonney et al., 2009). Other projects have also looked at involving users as data analysts - leveraging their prior knowledge to explain and discuss interesting patterns and trends in the data (Viegas, 2007). While these projects have focused on adult participation, others have investigated involvement of youth and novices (Willett, Aoki, Kumar, Subramanian, & Woodruff, 2010). Yet this work thus far hasn't focused on how and what forms of learning are enabled by these various forms of participation, and whether we need new tools to support these learning outcomes.

In this paper, we report on the results of a design-based research study, in which 4th grade students used a web-based tool called *Local Ground* to collect data, model, visualize and engage in discussion about the local soil ecology around their school. We analyze the affordances of situating children's data modeling in local spaces, and of using this map-based representational medium to support children in integrating their experiential knowledge into their reasoning as they investigate organisms in relation to the environment. Specifically, we explore the following research questions:

1. In what ways do children reason about organisms in relation to the environment, as they engage in data collection, modeling, and interactive mapping of their local soil ecology?
2. As they reason, in what ways do children draw on their context-specific knowledge and experiences to support their claims?
3. How do children use *Local Ground's* interactive map representation in constructing and communicating these explanations?

We begin by describing the instructional design and context of this research project, followed by our analytical methodology and lastly, our results and implications for future research.

Instructional design and context

Our research was conducted in an urban public elementary school (K-5) in the Western United States (40% free or reduced lunch). We worked with one fourth-grade class of 21 students that met twice a week in the Science lab for periods of 45 and 90 minutes. Working closely with the science teacher, we designed a ten-week instructional unit focused on exploring soil ecosystems found on the school grounds. Researchers were involved in all aspects of the design - including curriculum design, developing supporting tools, delivering instruction, designing classroom materials and activities, and leading discussions.

The curriculum was framed around four central questions: (1) What is underfoot? (2) Is it different in different places? (3) How can we find out? and (4) Why might these differences exist? These questions were intended to anchor instruction by problematizing (Dewey, 1929) the physical space, raising questions of method selection and data uncertainty, and encouraging children's explanations of this local ecosystem. Within children's take-up of these what, how, and why questions, we were particularly interested in the children's integration of locally situated knowledge sources into their reasoning, and their use of *Local Ground's* interactive map interface in constructing and sharing their explanations.

During the first few lessons the researchers elicited the class's ideas about what was underfoot around their school, grouped these ideas into biotic and abiotic "parts" of these underground places, and supported discussion about ways to gather more information about them. Children then worked in pairs to choose sites of interest to explore these "parts" further, selected by the researchers from students' initial ideas to represent a range of distinct microenvironments within the schoolyard. In these groups, children then collected data on both biotic (total invertebrate counts, earthworms, roots) and abiotic (soil moisture, soil compaction, soil color, soil composition) indicators. Children's paper notes took the form of written field notes, sketches, and photographs.

Back in the classroom, each pair worked together to add their data and field notes to *Local Ground's* spreadsheet interface (Figure 1). The teacher-researcher then used the interactive map interface (Figure 2) to engage children in thinking about similarities and differences at the ten sites, and to reason about why these relationships might exist. Children also physically returned to the ten original sites, to see if any additional site information would be useful in making sense of the relationships (or lack thereof) in their data.

Local Ground

Local Ground is a software tool designed to support children's data collection, modeling, and analysis (Van Wart, Tsai & Parikh 2010). This software allowed the research team to upload and geo-reference the various forms of children's soil data (pictures, audio, video, drawings). It also allowed children to enter and edit their own measurements and notes, and ultimately browse and visualize everyone's data, using either a spreadsheet (Figure 1) or interactive map (Figure 2) interface.

team_photo	site_photo	worms	moisture	percolation	pitfall	composition	notes	soil_sketch_1
		1	moist	16	9	grass, roots, she	most of the roots were green and I think there might have been more roly polies than we saw.	
		2	dry	18	2	Twigs, pebbles, leave	lots of twigs, pebbles, leaves, glass, and tiny roots	

Figure 1: Spreadsheet interface, which students used to enter/edit data into the system.



Figure 2: Interactive map interface, visualizing select variables from students' data collection.

The interactive map allowed children to assign specific symbols to discrete values or ranges for each variable, and to turn on/off any combination of symbols, allowing for the exploration of multivariate relationships across the schoolyard. We mapped four variables – earthworm count, soil moisture, percolation time (in seconds) indicating soil compactness, and pitfall trap count (number of invertebrates sampled). Figure 2 shows two of these indicators, “worm count” and “soil moisture”, displayed for each of the ten sites.

Methods

In this paper, we focus on the final three class sessions of the design experiment, which consisted of one day of teacher-led whole classroom discussion and two days of student-led informal presentations. In the teacher-led discussion, children explored relationships between the worm count, soil moisture, soil compactness, and total invertebrate counts, looking at one, two, and then all four variables on the interactive map interface. The teacher-researcher prompted the students by asking them to attend to similarities and differences across variables and sites and to identify “puzzling” relationships. In the student-led presentations, each site group (consisting of two

children) explored the interactive map interface together on a laptop, with the same four variables as above, to find “interesting” similarities or differences across sites and reasoning about why these relationships might exist. Then, in a whole-class discussion, each group took turns explaining their findings and related explanations to their peers using a large projected version of the interactive map. Children used this map to explore and discuss one- and multi-variable relationships in their data and to link their observations back to the original data collection sites.

We selected these three lessons because they show how children talked about their data observations at the end of the design experiment, how they leveraged their earlier field and data transformation experiences into this reasoning, and how *Local Ground’s* interactive map interface supported this process. Because we wanted to understand how the instructional design, data collection activities, and software mediated student reasoning about the local environment, we created a coding scheme consisting of (1) the different forms of reasoning that children used and (2) the resources children drew on as they reasoned. Using video transcriptions of the last three class sessions, these codes emerged through an iterative process (Miles, Huberman, & Saldaña, 2013) of identifying the types of local, experiential, and map-based evidence that the students marshaled to support their what, why and how reasoning. We started by coding the first lesson together, collaboratively identifying and defining what, why, and how forms of children’s reasoning in each turn of child or teacher talk. Using this initial coding scheme, two researchers independently coded the remaining two lessons. We then met to resolve discrepancies and refine the codes further to reach complete agreement.

Next, we coded for the different forms of evidence that children used, working together to identify the different contextually-situated resources children integrated into their reasoning in one lesson and then independently coding the two other lessons, reaching agreement through additional discussion and revisions of these codes (see Table 1). We included the use of the interactive map within these resources, noting indexical utterances (e.g. “here”) and gestures involving the map to indicate use. With these forms of reasoning and resources defined and indexed (see Figure 3), we then examined each form of reasoning and the resources together, noting any repeating patterns between these two groups of codes (e.g. children’s data collection experiences were often leveraged during how reasoning), and how the interactive map functioned in mediating this process.

Table 1: Coding Definitions and Examples

	Code	Definition	Example
Forms of Reasoning	What	A simple reporting of data	“They had wet soil and we had dry soil”
	Why	Establishing connections between biotic and abiotic variables (and sites)	“Wouldn't no worms be in compact soil because it would be hard to get into the soil if it's all crammed together?” [connecting worms to soil compaction]
	How	Reflecting on whether the data collection and presentation methods could be used to make a particular claim	“Maybe we didn't dig deep enough to find the worms” [questioning certainty of data sampling techniques]
Resources for Reasoning	Local Data Collection	Referencing specific aspects of the data collection experience	“We both had rollie pollies but theirs were alive and ours were dead” [referring to invertebrate collection experience]
	Local Knowledge	Referencing previous experiences in the schoolyard and local neighborhood	“Those garden beds get watered all of the time”
	Interactive Map	Using the map representation to talk about data points and / or context-specific locations	“I don't think you're right because one of the planters isn't showing here [pointing to the map]”
	Prior Knowledge	Reasoning using other forms of knowledge that were applied to a specific contexts but seemed to be generalizable across settings	“Worms can't climb up planters”

Data sources include video recordings of all classroom sessions and children’s data collection activities; videos of students’ computer screens and corresponding conversations (using the *SnagIt* screen capture tool); students’ written and illustrated work; and semi-structured interviews conducted during and after the final class session with select students. In this paper, we report on an analysis of our findings from the final three classroom videos.

Results and discussion

In this section, we first report on the general forms of children’s reasoning that we observed in the three lessons. Next, we discuss the different resources children drew up in their reasoning. We conclude with a vignette illustrating how several different sources of evidence and the *Local Ground* interactive map were used together to support children’s reasoning about select variables in relation to the surrounding environment.

Different forms of children’s reasoning

Children’s reasoning took many forms throughout the three lessons (see Figure 3): describing biotic or abiotic data points (coded as What), locating, considering, and refuting data relationships and explanations (coded as Why), and raising critical questions about methods and measurement issues (coded as How). On multiple occasions, children offered simple descriptions of the data at the ten sites (e.g. Lety: “Our site was dry”). More often, children considered and contested possible explanations about these relationships across several variables and sites. With instructional prompts that encouraged children to look for similarities and differences in the data and explain emergent relationships, they often considered one variable across several sites (e.g. John: “They had wet soil and we had dry soil”) and several variables within one site (e.g. Toby: “I noticed Keanu and Lety’s site had ... um moist soil and one worm and 30 seconds or less for percolation time”). At times, several children even considered multiple variables across several sites (John: “They both had no worms, moist soil, and between 20 and 200 seconds [for percolation time] and ... less than 10 invertebrates”). Given the complexity of the data (e.g. multiple biotic and abiotic variables at ten different sites), it is striking how children identified numerous covariate relationships, a challenging undertaking for children and adults alike when using canonical representational forms like scatterplots (Lehrer & Schauble, 2012). Children raised considerations of uncertainty multiple times as well, often in response to puzzling or unexpected relationships in their data. Children posed questions related to sampling errors (e.g. not digging deep enough to find worms, or percolation measurement techniques) and sampling variability (e.g. changing environmental conditions due to the sequencing and timing of data collection).

elapsed time (minutes):	L17: Whole Class Discussions							L18: Peer Presentations				L19: Peer Presentations					Total
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	0-5	5-10	10-15	15-20	0-5	5-10	10-15	15-20	20-25	
Forms of Reasoning																	
What Reasoning				1		1		3		3	5	5	2	1	1	2	24
Why Reasoning	4	1	2	4		3		3	3	2	2	5	3		2	6	40
How Reasoning		1		1		1	2				1				1	1	8
Resources for Reasoning																	
Local Knowledge	1	2							1		5		2	1			12
Local Data Collection	3	2					1		1			5	2				14
Interactive Map	1		1	1		6		1	1	3	5	2	1	3	2	6	33
Prior Knowledge							1		1	1	4	2			1		10

Figure 3: Counts of children’s scientific reasoning and the resources they draw upon, over time. Note students’ frequent use of the interactive map, integrating multiple resources as they reason about schoolyard ecosystem.

Resources integrated to support children’s reasoning

When looking at the resources children engaged during their reasoning, notable patterns emerge. Children drew from several different sources: experiences stemming from the actual data collection process, local knowledge drawn from daily activities around the school, and prior knowledge not directly related to the local context. In the following paragraphs, we describe the different resources in more detail and the ways in which they were used in children’s reasoning.

Children’s data collection experience

In most instances, children drew on their data collection experiences to raise questions about methodology and accuracy. Interestingly, these moments occurred in response to confusing or puzzling patterns in their data. For example, as the class reasoned about why worms were unexpectedly found at some of the ten sites and not others, Amir interjected, “Actually, I was going to say...some people didn’t find earthworms, they found these other weird worms.” Here, Amir is referring to his group’s and another group’s discovery of a different species of worm

during data collection, a finding that was discussed during data collection and recorded in both groups' scientific sketches. In light of Amir's contribution, the class continued to reason about the puzzling relationships, considering that different worms might indeed seek out different kinds of soil.

In a few instances, children also used their first-hand data collection experience to describe specific data points in more detail to better support their own observations or explanations. Elaborated visual or tactile descriptions of the data points were used to describe biotic and abiotic variables in greater detail. For example, in a conversation about different soil moisture levels in two planters – and their relationship to different worm counts – Tomas described the soil at one site as “pretty wet soil and it was shiny”, combining both a description of the soil moisture data point and referencing the visual qualities of that original phenomena. Here, students' data analysis reflects a “fluid” movement between the symbolized form and students' memories of their own data collection experiences and those of their peers.

Children's local knowledge about the area

Children's local knowledge was used in more diverse ways to reason about biotic or abiotic variables and the relationship between them, most frequently leveraged in the formation of and argumentation about explanations. This knowledge included human use patterns within the school grounds, as well as “natural” context-specific attributes such as shade and sunlight patterns. Take a discussion about “puzzling” patterns in the class data, emerging from the disconnect between children's expectation that worms prefer soil conditions with more moisture and the children's data on the map reflecting a different pattern:

Nadia: I think it's weird because, like, worms like wet areas and rain. But we didn't find any worms [in our wet site].

Researcher 2: Do you have any ideas about why?

Nadia: I think that me and Vanessa [interrupted briefly] because... there's a lot of children who play there and they might just stomp...And like the worms will go away. Or maybe the worms are just deep down.

Here, Nadia brings in her local knowledge about children's use of the space during recess to explain why worms might leave the area. She also raises the possibility that worms were actually at her site but the sampling methods didn't accurately capture them.

Children also leveraged their local knowledge of the space to consider additional site attributes influencing biotic and abiotic indicators. In a class discussion highlighting similarities and differences across sites and variables, one child, Eric, notices that two sites located far apart – one in the garden and one in a planter next to a classroom – both had no worms. Eric says, “So I think why me and Tomas didn't find any worms [*points to his site and the other group's site*] ... I think we didn't find any worms because mine was right next to the fence where it gets a lot of sunlight, and Tomas's is just...just sitting out in the sun.” Here, he argues that this might be caused by the amount of sunlight each location receives, a variable not originally considered by the class in their data collection and field notes.

Attention to sunlight and shade also emerged in several other children's comments, wherein children attended to shadows created by trees (e.g., Keanu: “Because this one is more in the sunlight and that one's in the shade where the trees are.”) and buildings (e.g., Sam: “No, it was covered by the shady part of the building.”). By considering these additional site attributes, children were able to reason about soil moisture and worm counts *in relation to* other factors within the original environment. Making a connection between data relationships and the original environment is a conceptually challenging yet crucial step in children's understanding of organisms' “fit” within an ecosystem (Lehrer & Schauble, 2012).

Children's prior knowledge not connected to specific sites

In several instances, children reasoned about relationships in their data using other forms of knowledge that were applied to the specific context, but likely originally developed in a different setting. For example, in trying to explain why worms would likely not climb into the planters (which were surrounded by wide expanses of blacktop), Tomas reasoned: “Well they could...but worms don't really like sunlight. They try...try to stay in the soil so they probably could but wouldn't because it would be really risky.” Attending to the plants growing at two sites, one child commented that worms were found at her site and another site because “they [worms] like to munch on leaves...and it would make sense that there are worms there.” These reasons tended to be drawn from other experiences, yet brought to bear to reason about the specifics of their local soil ecosystem.

Use of interactive map interface

Throughout the three class sessions, students used the interactive map representation to construct and support their arguments (see Figure 3). In many instances, the children simply used this interface to refer back to individual and aggregate data points. Other times, it functioned like a map, wherein children pointed and gestured toward specific locations in the original space (as in Keanu & Lety pointing to the interactive map interface, saying: “The apple tree's right over here.”) Interestingly, at many instances, children used the representation to talk about their biotic and abiotic site data in relation to the specific locations, collapsing the spatial, biotic, and abiotic forms of information together. In these moments, children considered spatially anchored phenomena (like sunlight/shade and schoolchildren’s use of these spaces over time) in relation to their abiotic and biotic data points.

On several occasions, the map interface supported students integrating multiple forms of contextually situated knowledge as they explained relationships between organisms and their environment. These moments often occurred during sustained engagement with particular data points and while exploring potential relationships, where explanations were being offered, taken up, and contested by several students. For example, in response to Eric’s prediction that the class’s data would not show worms in the planter sites “because I don’t think worms can climb up planters”, Tomas replied:

- Tomas: Uhhh, Eric...I don’t think you're right because one of the planters [*site #2*] is not showing right there.
- Researcher 1: What do you mean?
- Tomas: [*Pointing to the two sites on the interactive map*] There are two planters and one of the planters was a bit to the left...like around there... yeah [*in response to Researcher 1 pointing at site #2*]
- Researcher 1: Cool, so you think there should be worms in this spot? [*pointing to site #2*].
- Tomas: I think its...because the planters...mine and Heather’s [*pointing to his data and site #2*]...it's because ours is dry and theirs [the other planter group] is really wet. We both had rollie pollies but theirs were alive and ours were dead.
- Researcher 1: Here, right? [*pointing to the two sites*] Ah...so you are saying that you think there are differences in the worms based on soil moisture?
- Tomas: I think it was, it was because we had dry soil and they had pretty wet soil and [the wet soil] was shiny.
- Researcher 1: Ah, and you were both planters, right?

In this example, Tomas argues that differences in worm counts are related to soil moisture levels, not his classmate’s mechanistic explanation of worm’s climbing abilities. He uses the map-based representation to identify two sites (his and another’s groups) that were both in large soil planters. In addition to the worm count data and soil moisture data showing on the interactive map representation, Tomas also recalls and references the invertebrate pitfall trap data collection experience, wherein both groups found rollie pollies (sowbugs) yet his group’s sowbugs had all died. In this instance and several others like it, the teacher and students use the interactive map to establish a shared understanding of the data points being discussed (worm count and soil moisture) as well as the original context where the data was gathered (two planters on opposite sides of the schoolyard).

Conclusions

Our findings suggest that situating children’s engagement in data design, collection, and analysis within surrounding local spaces and using interactive data maps fostered generative conditions for children to reason about complex biological relationships. Within this design study, children marshaled diverse sources of evidence in reasoning about their data to elaborate descriptions of biotic or abiotic data points, consider and refute often “puzzling” data relationships, and to raise critical questions about methods. Details such as sunlight against a fence, shadows cast by buildings, Kindergarten-favored places to play, and a tree’s falling apples were accessible and integrated into children’s reasoning, supporting children in sharing and considering their peers’ complex explanations. In reasoning this way, children were able to leverage their vast experiential and everyday ways of knowing in grappling with the complexity of a local soil ecosystem.

Local Ground’s interactive map appears to be a potentially powerful representational form that warrants further exploration - in both its ability to establish shared references about complex data points and of the physical

context from which they come and to facilitate children's "fluid" movement between the world and its symbolized forms. Within science education modeling research, there has been careful attention to children's movement between base and target forms to support knowledge formation. It seems fruitful to explore the potential of this movement across the local contexts of children's daily experiences and within the representational forms themselves, building on the potential affordances of interactive spatial representations to help bridge material and symbolic worlds.

References

- Bonney, R., Cooper, C., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K., & Shirk, J. (2009). Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11), 977-984.
- Dewey, J. (1929). *The quest for certainty*. New York: Minton, Balch & Co.
- Enyedy, N., & Mukhopadhyay, S. (2007). They don't show nothing I didn't know: Emergent tensions between culturally relevant pedagogy and mathematics pedagogy. *Journal of the Learning Sciences*, 16(2), 139-174.
- Giere, R. (1997). *Understanding scientific reasoning* (4th ed). New York: Harcourt Brace.
- Latour, B. (1999). *Pandora's hope: Essays on the reality of science studies*. London: Cambridge University.
- Lehrer, R., Schauble, L., & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development*, 23(4), 512-529.
- Lehrer, R., & Schauble, L. (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. *Science Education*, 96(4), 701-724.
- Lehrer, R. and Schauble, L. (2015). The development of scientific thinking. *Handbook of Child Psychology and Developmental Science*, 2(16), 1-44.
- Manz, E. (2012). Understanding the codevelopment of modeling practice and ecological knowledge. *Science Education*, 96(6), 1071-1105.
- Manz, E. (2015). Resistance and the development of scientific practice: Designing the mangle into science instruction. *Cognition and Instruction*, 33(2), 89-124.
- Metz, K. E. (2004). Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations of their own design. *Cognition and Instruction*, 22(2), 219-290.
- Metz, K. E. (2008). Narrowing the gulf between the practices of science and the elementary school science classroom. *The Elementary School Journal*, 109(2), 138-161.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook*. Thousand Oaks, CA: SAGE Publications, Incorporated.
- National Research Council. (2012). *A framework for k-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.
- Rivet, A. & Krajcik, J. S. (2008). Contextualizing instruction: Leveraging students' prior knowledge and experiences to foster understanding of middle school science. *Journal of Research in Science Teaching*, 45(1), 79-100.
- Roth, W.M., Pozzer-Ardenghi, L., & Han, J. (2005). *Critical graphicacy: Understanding visual representations in school science*. Dordrecht, The Netherlands: Springer.
- Van Wart, S., Tsai, K., & Parikh, T. (2010). Local Ground: A paper-based toolkit for documenting local geo-spatial knowledge. In A. Dearden (Ed.), *Proceedings of ACM DEV*. London, UK: ACM.
- Taylor, K. H., & Hall, R. (2013). Counter-mapping the neighborhood on bicycles: Mobilizing youth to reimagine the city. *Technology, Knowledge and Learning*, 18(1-2), 65-93.
- Viegas, F., Wattenberg, M., Van Ham, F., Kriss, J., & McKeon, M. (2007). Manyeyes: a site for visualization at internet scale. *Visualization and Computer Graphics*, IEEE Transactions on, 13(6), 1121-1128.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Willett, W., Aoki, P., Kumar, N., Subramanian, S., & Woodruff, A. (2010). Common sense community: Scaffolding mobile sensing and analysis for novice users. *Lecture Notes in Computer Science*, 6030 LNCS, 301-318. doi:10.1007/978-3-642-12654-3_18

Acknowledgements

This work is based in part upon work supported by the National Science Foundation under Grant Numbers IIS-1319849 and IIS-1054332. We also acknowledge the support of a Sloan Foundation fellowship and a Hellman Faculty fellowship.

The Epistemology of Science and the Epistemology of Science Teaching

Carl Bereiter, University of Toronto, carl.bereiter@utoronto.ca

Abstract: Paul Kirschner has criticized constructivist science educators for confusing epistemology with pedagogy. He has also criticized them for a retrograde conception of science. This paper starts by agreeing with his second point and elaborates it into a fuller critique of naïve scientific epistemology, arriving at the conclusion that a better epistemology points to a better pedagogy. The key is a programmatic view of science, which sees the scientific enterprise as working continually toward stronger explanations of natural phenomena. This is a program that can be joined by people at all levels of sophistication, including young children, the goal always being to make today's ideas superior in some demonstrable way to yesterday's. Engaging students in explanation improvement has numerous advantages over atheoretical inquiry and project-based activities. Knowledge Building is highlighted as an approach that gives idea improvement a central place in its pedagogy.

Keywords: epistemology, nature of science, scientific method, Knowledge Building

Introduction

In a series of papers Paul A. Kirschner (1992, 2009; Kirschner, Sweller, & Clark, 2006) has argued that the epistemology of disciplines is a poor basis for the design of education in those disciplines. Although presented as a general criticism of constructivist, inquiry, and discovery educational approaches, Kirschner's argument is a direct challenge to approaches that take disciplinary practice as a model for education in the disciplines. Kirschner's position most obviously challenges Jerome Bruner's proposal (1960) that the "structure of the disciplines" should shape instruction—the pedagogical implication of which was that the way to learn a discipline is to practice it: "The schoolboy learning physics *is* a physicist, and it is easier for him to learn physics behaving like a physicist than doing something else" (Bruner, 1960, p. 14). A number of curriculum innovations in the 1960s, mostly but not exclusively in science, followed this line of thought. Perhaps the most extreme version of bringing disciplinary practice into the heart of educational practice, however, is Knowledge Building (Scardamalia & Bereiter, 2006, 2014; Scardamalia, Bereiter, & Lamon, 1994) where "knowledge creation," as practiced in research laboratories and innovative companies, becomes the principal means of education in the disciplines (Bereiter & Scardamalia, 2014). Thus Knowledge Building, at least in Scardamalia's and my formulation of it, takes a diametrically opposite position to Kirschner's. Despite this, I agree with almost everything Kirschner says in advancing his argument. How, then, do I arrive at an opposite conclusion?

In accord with many contemporary cognitive scientists and philosophers, Kirschner uses the term epistemology broadly to refer not only to how knowledge claims are justified but also to how they are produced. He draws the following distinction between epistemology and pedagogy:

Modern curriculum developers and instructional designers confuse the epistemological nature of a domain with the psychological bases of learning and the pedagogic bases for teaching. Epistemology refers to how knowledge is acquired and the accepted validation procedures of that knowledge; pedagogy refers to how something is taught. (Kirschner, 2009, p.151)

Kirschner criticizes constructivists not only for failing to honor this distinction but also for promoting an inductivist epistemology that he calls "basically flawed" (1992, p. 275), citing a number of authorities who make this point: e.g., "Gardner (1975) calls the notion that 'scientists patiently gather piles of evidence which they can put together inductively to form a law is absurd . . . inductive reasoning may be involved in the checking stage of a law, but not at the formulation stage' (p. 17)." In his 2009 paper, Kirschner makes the same point more briefly, but calls it a digression from his main point about the difference between epistemology and pedagogy. This is where we part company. I see the "flawed epistemology" on which much of constructivist pedagogy is based as a critical factor in what is wrong with it, and this leaves open the possibility that pedagogy based on a better epistemology could perform better and might be preferable to an "instuctivist" pedagogy. The purpose of the present paper is to explore this possibility and what it entails for educational design.

Can students carry out disciplinary knowledge production?

Kirschner does not object in principle to education based on the epistemology of disciplines. His objection is a practical one: It's too hard; students aren't equipped for it. Although he draws on theoretical backing from Piaget and Vygotsky in his 2009 paper and on cognitive load theory in Kirschner, Sweller, and Clark (2006), the question of whether young students can actually practice disciplinary knowledge creation is an empirical, not a theoretical question. Claims that they cannot do so stand to be refuted by empirical demonstration. I believe the accumulated research on Knowledge Building provides such refutation (Bereiter & Scardamalia, 2010; Scardamalia & Bereiter, 2014; Scardamalia, Bereiter, & Lamon, 1994). Some of the research involves experimental comparisons, usually of a before-and-after nature, but for the present purpose even anecdotal data can carry weight. It is beyond the scope of this paper to present evidence supporting this claim; the following example may, however, at least serve to clarify what is being claimed. The example comes from a grade one (age approximately six years) class reported in Scardamalia (2002) and based on a teacher's report:

After a fall outing, the children came up with the question: Why do leaves change color in the fall? The students contributed notes proposing explanations. The teacher or an aide provided help with spelling as needed. Some examples of student contributions:

"Because the sap cant get to the leaf because of a plug. Then the chlorophyll dies and the leaf changes colour." (Illustrated with drawing of sap running through the branches.)

"fall – i think the chlorophyll goes into the tree to keep warm for the winter."

"i think leaves change colour because when the leaf falls down I think that the chlorophyll goes to the outside of the leaf so it leaks off the leaf."

"Because it's too cold for the chlorophyll to make food for the tree."

Notice that although the theories are faulty they do hypothesize physical processes and they are potentially testable. To test whether cold alone was sufficient to explain leaves changing color, the first-graders put green leaves into a freezer. More remarkable is something that occurred months later during a field trip to a maple-tree farm to see how maple syrup is made. One child, watching the sap flow from the tree, remarked that her theory regarding chlorophyll must be wrong, because the sap she saw was not green. Others raised many other issues about what they saw, and how the flow of sap gave them new ideas about the internal structure of a tree, and the relation of its internal structure to their theories. A new theory emerged to the effect that there are two paths for sap to flow, one near the outside of the tree trunk for colorless sap and one deeper inside for the chlorophyll-bearing sap. In rudimentary form, all the essential elements of scientific theory-building are evident among these six-year-olds.

Some common limitations of children's scientific thinking are evident in this example: a lack of system (Vygotsky, 1934/1962, p. 116), a penchant for single-cause explanations, and a susceptibility to animism (the chlorophyll trying to keep warm). Except for the animism, these limitations may be found in abundance in the comments by adults on internet news sites. But on the positive side, the children demonstrated the fundamentals of authentic scientific thinking: the creation of explanatory hypotheses and the testing and revision of these on the basis of evidence. It has seemed, on the basis of numerous examples of young children's explanation-building, that their implicit epistemology is more advanced than the "scientific method" they will be taught in later years. It seems these six-year-olds intuitively understood that the job of theories is to explain facts, not to achieve the status of facts. I have seen evidence of this in even younger children. They display a level of scientific thinking that apparently gets dragged down by later instruction.

Kirschner agrees that students need a better epistemology—a better understanding of disciplinary practice—but says they should not be left to discover this for themselves or to develop it intuitively through the usual constructivist diet of "guided discovery" exercises and hands-on projects. Teaching scientific epistemology to school students is not a walk in the park, however. Carey and Smith (1993) identified three levels of student understanding of the nature of science and scientific knowledge. At the lowest level, as noted previously, students see science as the unproblematic accumulation of facts. At level 2 the uncertainties of knowledge claims and the importance of evidence are recognized. At level 3 the role of theories in both the formulation and interpretation of scientific research is recognized. Through Carey and Smith's own research and through much subsequent research, it has been found that moving students from level 1 to level 2 is achievable, but that movement to level 3 is rare (Chuy, Scardamalia, Bereiter, et al., 2010). The bulk of research on student epistemologies, starting with that of Robert Perry and continuing through a steady flow of descriptive and occasionally experimental studies, has for the most part not recognized level 3 at all but has instead concerned itself with level 1 and variations within level 2—generally variations between solipsistic relativism and recognizing that some beliefs can be better justified than others. According to Chinn and Malhotra (2002), theory plays virtually no role in forms of inquiry

common in school science education. Level 3 is evidently rare among teachers as well. Windschitl (2004) found that teachers' own beliefs about the nature of science overlooked its creative, theory-focused character.

The reduction of science to atheoretical fact gathering and hypothesis testing coupled with an emphasis on the uncertainties of scientific knowledge make for a pedagogical brew that provides students little motivation to seek scientific careers and leaves "it's just a theory" as a reason to reject climate change, the descent of *homo sapiens* from earlier life forms, or any other unwelcome truth. Students may still find things of interest in science, sometimes because of inherent interestingness of topics such as dinosaurs and space travel, sometimes because the hands-on activities are intriguing. But this is a far cry from the central motivation of science, which is to understand how the world works.

The importance of understanding the nature of science ("NOS") is by now well recognized in curriculum planning. However, failure to understand the nature and role of theory is failure to understand NOS—and on that basis failure can be said to be endemic. Failure to understand the nature of theory does not only affect those who might pursue careers in science but has social, political, and personal consequences. Political controversies often involve theories and theoretical models, current examples being controversies about climate change and economic policy. Public support of basic research requires some appreciation of the value of theory, of the theoretical value of research on organisms and phenomena of little obvious practical significance, and the often crucial importance of understanding a problem before plunging ahead with possible solutions. All of these real-life functions of scientific knowledge require understanding science as an effort to construct a coherent and testable understanding of the world—in short, as theory building. This implies an epistemology that goes well beyond the shallow constructivism of contemporary school science.

The crucial importance of progressivity in knowledge

If it is agreed that students should be gaining a better understanding of the nature of science, then the next question is what *is* the nature of science? Any perusal of contemporary philosophy of science will make it clear that this is by no means a settled issue. There are a number of opinions, which may be roughly categorized as no-difference, methodological, and programmatic. Arguments for these positions are all plausible enough that I do not see any virtue in educationists trying to decide which is right. A pragmatic approach holds more promise (cf. Resnik, 2000)—in this case an approach based on consideration of what position is best for the design of school science education.

The no-difference position does not hold that there is no discernible difference between scientific and other practices, only that there is no clear line of demarcation. Whatever is singled out as distinctive in scientific practice can also be found in some form in everyday thought or in pseudo-sciences such as astrology and dianetics. As applied in school, however, the no-difference position would seem to legitimate low-level relativism: whatever belief feels right to you is okay. Methodological views of the nature of science hold that there are certain practices for arriving at conclusions that are different from other ways and that these practices are what make conclusions "scientific." From a pedagogical pragmatic viewpoint, the question is not whether methodological criteria for what constitutes science are correct or adequate but whether they represent a good basis for educational design. I have already signaled my main objection: Education based on a conception of science as method represents it as a joyless, pedestrian occupation with its whole bold, creative, frontier-advancing character bled out of it.

The programmatic view of the nature of science is probably the least recognized in education, although it has strong philosophical backing. It requires viewing science over time spans and at a systemic level, which goes against the educational tradition of treating science as first and foremost a matter of individual knowledge and the testing of isolated hypotheses. Although beginnings of a programmatic view may be found among logical positivists, it was Imre Lakatos (1970) who explicitly argued that the proper object of scientific evaluation is not particular theories or knowledge claims but rather research programs. At the root of Lakatos's argument was the observation that scientists rarely abandon a theory because of disconfirming evidence (as schooling in the "scientific method" would have us believe). Instead they try to improve the theory or modify the knowledge claim in order to accommodate the troublesome evidence. What he called "progressive" research programs deal with anomalous data in ways that lead to stronger and stronger theories, whereas non-progressive programs deal with it in a variety of other ways, as documented by Chinn and Brewer (1998). Evolutionary theory has had to deal with numerous discoveries that challenge existing theory and has grown progressively stronger as a result, so that inexplicable findings become increasingly rare. By contrast, Intelligent Design, as a research program, keeps getting weaker as its knowledge claims in the form "Evolution cannot explain..." succumb to evidence that evolutionary theory can in fact account for the phenomenon at issue. Intelligent Design researchers have to keep looking for new things they believe evolutionary theory cannot explain, and in recent times have shifted their ground from the evolution of species and structures to the origin of the life and the universe—which are matters that evolutionary theory does not claim to deal with.

As a basis for educational design this programmatic conception of science has significant advantages over both no-difference and methodological conceptions:

- It allows the curriculum to start with whatever students currently (rightly or wrongly) understand and move on from there in a progressive fashion.
- It solves the teacher's problem of how to respond to students' inventive but wrong ideas. By actively promoting the working assumption that all ideas are improvable, the teacher is free to respond with "That's a wonderful idea. How could we make it even better?"
- It allows ample room for the creative aspect of science, which tends to be suppressed in methodological conceptions (i.e., the traditional "scientific method"), but also avoids the "anything goes" implication of no-difference conceptions.
- It sets out a realizable goal for students—idea improvement—rather than a discouragingly lofty one.
- It invites students to work collaboratively at idea improvement and to take mutual responsibility for it.
- It allows students to work progressively within their developmental and information processing capacity limits. Students can do productive work with ideas even though they lack "conservation of substance," "control of variables," and other elements of mature scientific thinking.
- It supports a progression from self-centered to idea-centered. Alice's idea and Jaime's idea (which may or may not be the same idea and which may be so vague that it is impossible to tell) are eventually absorbed into a collectively produced theory or solution that becomes community rather than individual property.
- Efforts to improve ideas so that they explain more facts or solve more problems tend to move beyond singular ideas and combinations of ideas to integrated structures of ideas—that is, toward theories.
- Although Lakatos was mainly concerned with theoretical programs, the idea of progressive programs applies equally to ones that address practical problems or social issues—and to some educators these are the most important objectives of science education for the masses.
- It provides a basis for students' seeing their work as part of a larger project of knowledge creation and becoming enculturated into a society of knowledge creators, with its long history and high prospects.

Students as practitioners of progressive science

The programmatic view of science has clear advantages for developing student understanding of the nature of science, but is it applicable to students' own scientific inquiry, as carried out under normal school and laboratory course conditions? Its applicability is severely limited if inquiry is confined to brief and unconnected episodes. Continuity and progressive development of explanations are essential according to the programmatic view. This means that, in order to be practitioners rather than only assimilators of progressive science, students need to be engaged not only in explaining but in building stronger explanations. Drawing mainly on Thagard's theory of explanatory coherence (2000), but with echoes of work by others such as Karl Popper and David Deutsch, we can stipulate that an explanation is getting stronger to the extent that

1. it explains more facts
2. it excludes more false statements
3. it connects to more other explanations
4. it explains things in more detail
5. parts of the explanation interlock so that it becomes increasingly difficult to modify parts without altering the whole
6. it is able more clearly to identify what it fails to explain
7. it generates better predictions
8. it explains *how* identified causal factors work, rather than only identifying and quantifying their effects.

An explanation that satisfies all these criteria is not something student explanation building is likely to achieve very often. However, the criteria themselves are not difficult to understand and thus represent something students can work toward collaboratively. In doing so, they will be engaged in theory building (Bereiter, 2012). Theory development, theory testing, and theory improvement are what basic research in science is all about. This is not a high-flown notion beyond the grasp of young students. It is perfectly comprehensible to them. They take

naturally to producing theoretical explanations of natural phenomena and judging and modifying them in the light of evidence (Chuy, et al., 2010). It is schooling in the so-called “scientific method,” with its often pointless testing of predictions, that may be responsible for students later losing their grip on this understanding. However, if Newton’s laws are taken as an exemplar, theory building is apt to seem forbiddingly difficult and definitely not for everyone. Newton’s laws and most physical theory building in the physical sciences represent what Bruner (1986) defined as the “paradigmatic mode,” grounded in mathematics and logic and typically universal in scope. Young students can function in paradigmatic mode, as indicated by their ability to work out rules accounting for numerical or geometric progressions (Moss & Beatty, 2006). But most of their theorizing about natural and social phenomena is going to take place in what Bruner called the “narrative mode.” In this mode, a process or phenomenon is explained by a series of events, with one causing the next. Building on children’s narrative competence affords a way of getting students into scientific explanation without running up against the cognitive demands of paradigmatic explanation.

Of particular interest for science education at the school level is the “how-it-works” narrative. The progress of a plant from seed to fruition lends itself to an event sequence narrative, but often the narrative is complicated by the fact that things happen concurrently. A story about how an electric bell works must recognize that the same movement of the striker that sounds the bell also breaks the circuit that drew it to the bell, thus resulting in its springing back from the bell in readiness for a new cycle. Nutrition, respiration, and circulation are all processes that can be given explanations in narrative form, with the stories gaining complexity and undergoing correction as students’ knowledge advances. Evolutionary explanations of the emergence of wings, legs, sexual reproduction, and so on, all have a story-like structure that students can grasp, elaborate, and criticize.

Oejjord (2003, p. 145) has set out criteria that a causal narrative must meet in order to be of scientific value. These are similar to the criteria of strong explanations set out above. Causal narratives that fall short on these criteria are often called “just so” narratives. I would argue that “just so” stories have value in education as improvable ideas—ideas that acquire more explanatory strength as they encompass more facts and as problems with them are solved. Constructing causal narratives is something children and novices can do that engages them with important scientific concepts and problems and that gives them realistic practice in a mode of thinking (abduction) that is a vital mechanism of creative work with ideas (Paavola & Hakkarainen, 2005b). The point most relevant to the issue of epistemology’s relation to pedagogy is that, while the epistemology of science projects a landscape that contains heights beyond normal student capabilities, it also contains relatively gentle slopes that ordinary students can negotiate.

The knowledge creation/ knowledge building alternative

Project-based science is aimed at engaging learners more directly in the kinds of practices that characterize “doing science” in the real world (Marx, Blumenfeld, Krajcik, & Soloway, 1997). However, project-based learning covers a wide range of approaches. At one extreme we have the traditional school project, in which students produce documents on selected topics that resemble encyclopedia articles in content and often in structure as well. Despite use of Web resources and computer-based representation tools, the basic pedagogy has not changed significantly in a hundred years. Then there are projects organized around some fictitious or real-life theme (e.g., planning a trip to Mars or analyzing pollution in a nearby stream). Although these can have genuine merit, they fall short of engaging students in the core enterprise of theory building—working to make systematic sense of the physical, biological, and social world. There is an alternative that engages students directly in this core sense-making enterprise. It is Knowledge Building (Scardamalia & Bereiter, 2003, 2006, 2014), defined as the production and continual improvement of ideas of value to a community. (I capitalize the term when referring to the cited program but use lower case when discussing concepts and processes.) Paavola and Hakkarainen (2005a) have shown that there is a close affinity between knowledge building as carried out in schools and knowledge creation as carried out in research laboratories and innovative companies. Scardamalia and I have argued (Bereiter & Scardamalia, 2014) that they are indeed the same concept, with differences in application due to context. One of these differences, of course, is that in an education context knowledge building/knowledge creation is expected to produce learning—and its success is ultimately judged by learning outcomes rather than by the quality of the ideas collectively produced.

Knowledge Building is not limited to any particular subject. It may include artistic and engineering types of knowledge creation. The creation of literary works, invention of mechanical and electrical things that work, and various sorts of creative computer program development may all constitute knowledge creation in so far as something new is added to the cognitive resources available to the knowledge-building community. However, the most distinctive and challenging kind of knowledge building in educational contexts is the creation and improvement of explanatory theories or theory-like “conceptual artifacts” (Bereiter, 2002). New education

standards are calling for teaching “big ideas,” and most modern programs honor this call in some fashion; but Knowledge Building makes the ideas themselves objects of inquiry.

The driver of students’ idea-centered work is knowledge-building dialogue, which may involve argumentation but only as part of a collaborative effort to solve problems of explanation (Scardamalia & Bereiter, 2006). Students do experiments and other kinds of empirical work, often self-initiated, but always with the purpose of advancing the knowledge-building dialogue. The vast information resources of the Worldwide Web become an integral part of the process, which makes Knowledge Building a distinctly 21st-century approach. It is not expected that students will invent or discover through experiments big ideas like universal gravitation, natural selection, and photosynthesis, but neither is it left to the teacher to convey these ideas through didactic instruction. Instead, it is expected that students will encounter and assimilate them through their use of Web resources in efforts to improve their own explanations of natural phenomena. Whereas guided discovery, argumentation, and project-based learning may engage students in some of the elements of scientific practice, Knowledge Building incorporates all of these into the fundamental scientific effort to build coherent theoretical explanations of the phenomenal world.

Conclusion

Separating pedagogy from the epistemology of disciplines makes sense in many practical respects, but only up to a point. Creative instructional designers will do more than schedule topics according to logical or empirical dependencies. They will try to identify learnable concepts that are legitimate simplifications or reduced versions of target concepts (White, 1993). But doing this requires getting deeply into the structure and essence of the concepts, a sort of thing philosophers of science do. Kid-level concepts need to lead in the right direction, toward eventual grasp of the target concepts and not toward roadblocks and misconceptions.

In this paper I have elaborated on other ways in which the epistemologies of the sciences and science education intersect. With regard to students understanding the nature of science, Kirschner and I and many science educators agree that the conventional version of “the scientific method” is fundamentally flawed. But its basic flaw is epistemological, with its pedagogical drawbacks following as a consequence. I reviewed several competing conceptions of what distinguishes science from non-science, an epistemological issue that I would not presume to judge. But from a pedagogical point of view I think there is no doubt that a programmatic conception is preferable to a no-difference or a methodological conception, no matter how well-reasoned these may be. This is not epistemology directing pedagogy, but it is a pedagogical choice that could not have been formulated without the work of Lakatos and other philosophers of science.

Understanding what constitutes real science is one thing, but doing real science is another, and one that arguably deserves priority. Naïve empiricism provides a simple answer to engaging students in real science: Get them practicing “the scientific method” by generating hypotheses (guesses), turning them into predictions, and testing the predictions. (For an indication of how ubiquitous this practice is, do an image search on “science fair.”) But if students are to do the kind of science that significantly advances understanding of how the world works, they need to build coherent explanations and work to make them better. Their theorizing should lead them to experimentation, whereas hands-on experimentation leading to theorizing, although it does happen in real science, is more a pipe-dream of learning-by-doing enthusiasts than something that can be expected to motivate inquiry in the classroom. Building and improving explanations means working *within* a discipline, not just carrying out some of the visible practices of researchers and applying received ideas to what one observes. Providing support for knowledge creation—what in educational contexts we call Knowledge Building (Bereiter & Scardamalia, 2014)—requires attending seriously to and supporting the knowledge processes through which research programs advance knowledge frontiers. In designing this kind of education, pedagogy cannot be separated from epistemology, as Kirschner would seem to advocate. Instead, the epistemology of the disciplines needs to be extended to encompass the peculiar conditions of disciplinary knowledge production in educational settings.

References

- Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bereiter, C. (2012). Theory building and education for understanding. In M. A. Peters, P. Ghiraldelli, B. Zarnic, and A. Gibbons (Eds.), *Encyclopaedia of educational philosophy and theory*. Web document: http://www.ffst.hr/ENCYCLOPAEDIA/doku.php?id=theory_building_and_education_for_understanding
- Bereiter, C., & Scardamalia, M. (2010). Can children really create knowledge? *Canadian Journal of Learning and Technology*, 36(1). Web: <http://www.cjlt.ca/index.php/cjlt/article/view/585>.
- Bereiter, C., & Scardamalia, M. (2014). Knowledge building and knowledge creation: One concept, two hills to

- climb. In S. C. Tan, H. J. So, & J. Yeo (Eds.), *Knowledge creation in education*, pp. 35-52. Singapore: Springer Science + Business Media
- Bruner, J. (1960). *The process of education*. Cambridge: Harvard University Press.
- Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.
- Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, 28 (3), 235-251. 43.
- Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching*, 35, 623–654.
- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic reasoning in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86, 175-218.
- Chuy, M., Scardamalia, M., Bereiter, C., Prinsen, F., Resendes, M., Messina, R., Hunsburger, W., Teplovs, C., & Chow, A. (2010). Understanding the nature of science and scientific progress: A theory-building approach. *Canadian Journal of Learning and Technology*, 36(1). Published online at <http://www.cjlt.ca/index.php/cjlt/article/view/580>
- Gardner, P. L. (1975). Science and the structure of knowledge. In P. L. Gardner (ed.), *The structure of science education*. Hawthorn, Australia: Longman Australia.
- Kirschner, P. (1992). Epistemology, practical work and academic skills in science education. *Science and Education*, 1, 273–299.
- Kirschner, P. A. (2009). Epistemology or pedagogy, that is the question. In S. Tobias & T. M. Duffy. *Constructivist instruction: Success or failure?* (pp. 144-157). New York: Routledge.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimally guided instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching.
- Lakatos, I. (1970). Criticism and the growth of knowledge. In I. Lakatos, & A. Musgrave (Eds.), *Falsification and the methodology of scientific research programmes*. Cambridge: Cambridge University Press.
- Marx, R. W., Blumenfeld, P. C., Krajcik, J. S., & Soloway, E. (1997). Enacting project-based science. *Elementary School Journal*, 97, 341-358.
- Moss, J. & Beatty, R. (2006). Knowledge building in mathematics: Supporting collaborative learning in pattern problems. *International Journal of Computer-Supported Collaborative Learning 1* (4), 441-465.
- Oeijord, N. K. (2003). *Why Gould was wrong*. Bloomington, IN: iUniverse.
- Paavola, S. & Hakkarainen, K. (2005a). The knowledge creation metaphor – an emergent epistemological approach to learning. *Science & Education*, 14(6), 535-557.
- Paavola, S. & Hakkarainen, K. (2005b). Three abductive solutions to the Meno paradox - with instinct, inference, and distributed cognition. *Studies in Philosophy and Education* 24(3-4), 235-253.
- Resnik, D. B. (2000). "A pragmatic approach to the demarcation problem". *Studies In History and Philosophy of Science Part A* 31 (2): 249–267. doi:10.1016/S0039-3681(00)00004-2.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.) *Liberal education in a knowledge society* (pp. 67-98). Chicago: Open Court.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge building. In J. W. Guthrie (Ed.), *Encyclopedia of education* (2nd Ed.) (pp. 1370-1373). New York: Macmillan Reference.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 97-118). New York: Cambridge University Press.
- Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (2nd ed.). New York: Cambridge University Press.
- Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into World 3. In K. McGilley (Eds.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 201-228). Cambridge, MA: MIT Press.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MA: MIT Press.
- Windschitl, M. (2004). Folk theories of “inquiry”: How preservice teachers reproduce the discourse and practices of an atheoretical scientific method. *Journal of Research in Science Teaching*, 41(5), 481-512.
- White, B. (1993). Causal models and intermediate abstractions: A missing link for successful science education? In R. Glaser (Ed.), *Advances in instructional psychology* (vol. 4), pp. 177-252. Hillsdale, NJ: Lawrence Erlbaum Associates.

Embodied Cognition in Observational Amateur Astronomy

Flávio S. Azevedo, The University of Texas at Austin, flavio@austin.utexas.edu
Michele J. Mann, The University of Texas at Austin, mjmann@utexas.edu

Abstract: We investigate the embodied cognitive practices of amateur astronomers, as they engage in observational routines in the field. In particular, we take an “interactionist” perspective to the problem and address the question: How is the body used as a resource for practical, conceptual, and communicative purposes in the process of planning, searching for, and sighting celestial objects? The basis for our analysis is a set of ethnographic video and field note records collected over several years of fieldwork. We scrutinize the moment-by-moment interactions among practitioners for episodes in which different modalities of embodied action are recruited and coordinated to produce and communicate complex meanings in the practice. As a whole, therefore, we seek to empower learners—as the ICLS 2016 calls for—by bringing insight into foundational aspects of knowing and learning.

Keywords: embodied cognition, amateur astronomy, hobbies, out-of-school learning

Introduction

Responding to ICLS 2016’s call for research “... to provide [renewed] insight into how people learn,” in this paper we investigate the embodied cognitive practices of amateur astronomers, as they engage in observational routines in the field. By and large, the human body has been relegated to a secondary role in accounts of reasoning and learning (Hall & Nemirovsky, 2012; Stevens & Hall, 1998). By contributing to approaches that work to re-inscribe the body into cognitive processes (e.g., Goodwin, 2011, 2013; Nemirovsky, Rasmussen, Sweeney, & Wawro, 2012), our goal is to add to basic theorizing on knowing and learning in STEM practices and to fine-grained analyses of the situated, context-dependent nature of these processes. We thus seek to empower learners—as the Conference calls for—by bringing insight into foundational aspects of learning.

We pursue this inquiry from a perspective known as “interactionist” (Jordan & Henderson, 1995). As such, we seek to document how the body is used as a resource for action and reasoning, *in situ*, on a moment-to-moment basis, in interaction with other bodies, material infrastructures (e.g., tools and instruments) and other parameters of practice, such as communities’ norms and values, divisions of labor, histories of participation, and so on (e.g., Hall & Stevens, 1995; Stevens, 2000). Specifically, we ask the question: How is the body used as a resource for practical, conceptual, and communicative purposes in the process of planning, searching for, and confirming sight of a celestial object?

As we will see, this question becomes particularly interesting and complex when we consider the conditions of practice of observational amateur astronomy and the specifics of the “task” of finding any single celestial target. For good seeing, astronomers practice under the darkest possible conditions and only red-screened light sources (e.g., flashlights and computer screens) were allowed on site. Because the practice is densely populated with tools and various artifacts—including telescopes, tripods, binoculars, computers, books and various star charts and notebooks, tables and chairs, among others—and observations are often carried out in small groups, these conditions created some important practical challenges, such as coordinating joint activities and communicating information on celestial scenes to peers and visitors.

Finding a celestial target clearly involves conceptual complexity as well. The process typically unfolds as follows. First, the astronomer picks a target, frequently one missing from a pre-made, themed list of observational targets (or similar). Then, he/she may read some about the object, say, in search of information regarding its shape, general look, location, and relationship to nearby celestial bodies. To do so, the astronomer may resort to a number of sources, from handwritten notes (taken at home), to printed books and other reference materials, to computers (and other electronic, handheld equipment), and others. Following that, he/she will chart a “star hopping” path through visible objects in the vicinity of, and towards the target, and finally attempt to sight it through the scope. In all, this makes for an iterative and relatively extended pursuit in which practitioners continuously assemble to one another representations and descriptions of a target scene, and coordinate their understandings of it. Bodies take center stage in these processes, as we will see.

Theoretical framework

The interactionist perspective that we adopt insists on a socio-cultural, activity-systemic perspective in which knowing and learning are actively produced in the transactions among elements of such a system—including the

organization of the activity (or practice), the distribution of roles and attributions within the activity, the material and historical conditions under which participants work, etc. (e.g., Gibson, 1979; Jordan & Henderson, 1995; Saxe, 1991; Hutchins 1995; Cole, 1996; Engeström, 1999). For our inquiry into embodied cognition, this means that we see “human thinking and learning [as] intimately tied not only to the body, but to a body that interacts with others and is active in social and cultural settings already rich with mediating artifacts that afford particular kinds of joint activity” (Hall & Nemirovsky, 2012, pp. 213-214). We take it that multiple modalities of embodied action are engaged in these processes—including gestures, talk, touch, prosody, gaze, body posture and orientation, and tool use (Goodwin, 2011; Hall, 1996; Nemirovsky et al., 2012; Stevens, 2012)—and our work is aimed at documenting how these modalities are manifested in observational amateur astronomy practice.

Further, as Goodwin (2011) illustrates, multiple modalities of embodied action are coordinated in activity, each mutually elaborating on one another as means of articulating and expressing complex meaning. In addition, because the oft-collaborative character of the task of finding a celestial object, coordinations and alignments between participants’ bodies should also be expected (Stevens & Hall, 1998; Hall, 1996). These should be further intertwined with the immediate conditions of practice, as we will see in the analysis.

Finally, we note that—as in other techno-scientific practices (Latour, 1987; Lynch & Woolgar, 1990)—observational amateur astronomy is heavily mediated by technical representations (e.g., star charts, maps, and pictures), tools (telescopes, binoculars, and others), and various artifacts. In action, these serve to structure and extend the active body and are critical for an account of how people create meaning within the collective, naturally occurring exchanges in the field.

Methods

The data we analyze here are ethnographic video records and field notes (Hammersley & Atkinson, 1995) produced during two distinct research studies. In the first study, which took place between 2002 and 2003, the first author (FA) documented the individual and collective practices of amateur astronomers with the goal of explaining structures and processes of extended, interest-based participation (Azevedo, 2013). In the second—which spanned the full year of 2014—we returned to the field with the explicit goal of documenting the embodied practices that permeates amateur astronomy field practice and which the first study had hinted at. In all, we collected 8 hours and 55 minutes of video in study one and 6 hours and 11 minutes in study two, and several pages of field notes, theoretical and analytical memos across both studies.

Studies one and two were carried out on different settings as well. Study one in fact included two settings—Mt Hillview and Lake Countryside, both located in Northern California—each of which was frequented by different communities of astronomers. (All names are pseudonyms.) Study two took place within a single community of astronomers that met at the High Meadows observation site in the Texas hill country. In all cases, astronomy clubs and communities held a variety of events and we focused on the most common and recurrent ones, namely the public “star parties” (i.e., outreach efforts) and members-only nights held at that site.

Of particular importance, based on observations in study one, in study two we sought to address the challenges of data collection under the dark conditions characteristic of astronomy field practice. This was especially crucial in light of the goal of documenting, at a fine grain of detail, the interactions between participants and the bodily practices involved, the use of star charts, telescopes and other tools, and how these were coordinated and elaborated upon throughout the “task” of seeing an object.

To tackle this problem, we used an infrared (IR) camera fitted with an infrared lens. Because the camera shed light undetectable to the human eye, we were able to illuminate and record participants’ interactions and long-term work without disturbing the original conditions under which they carried out their practice. Had we used a regular camera and illuminated the scene with red-screened flashlights, we would have circumvented the technical problem, but fundamentally altered the constraints and affordances that practitioners regularly encounter and thus the very character of the embodied practices we sought to study.

Analysis and results

We scrutinized the full set of ethnographic records for moments in which participants’ (astronomers and visitors) bodies seemed to take an active role in the ongoing observational activity. We then bracketed these moments for further analysis and watched the corresponding videotape segments several times, all along transcribing parts that required clarification or elaboration (for whatever reason). With episodes of embodied cognitive practices catalogued and described, we classified them into categories that reflect either a modality of embodied action and reasoning or their functional role in the activity. We organize the presentation of results around these categories.

Averted vision

Averted vision is a technique that is crucial to observing deep sky, and therefore faint, celestial objects. The technique works by exploiting the basic structure of the human eye. Briefly, the eye has two distinct types of receptor cells—i.e., cones and rods. Cones are found mostly in the central area of the retina and their function is to make out details and colors in a scene. However, cones are not very sensitive to light. Rods, on the other hand, populate the periphery of the eye and they are very sensitive to light, though they cannot make out object details or colors.

When applied to amateur astronomy practice, the technique posits one should look at a faint object at an angle—rather than directly, straight on—thereby exposing the most sensitive part of the eye to the scene (say, in a telescope or binoculars). The exact angle of off-center looking varies across individuals, but usually it falls between 8 and 16 degrees. Both specialized books and amateur astronomers report that it takes time to learn the technique, but the gains in “seeing objects” are recognizably high (Dickinson & Dyer, 2008).

Because averted vision requires training a body part itself to perform in unusual ways, the technique perhaps best exemplifies how deeply embodied cognitive processes can be in a STEM practice, and astronomy in particular. The technique is perhaps also a prototypical example of learning-by-doing across STEM disciplines; no matter how much one might hear explanations of averted vision, mastering the technique takes extended individual practice.

The body as medium of communication

The unique conditions of practice that characterize observational amateur astronomy create practical demands that seem quite specific to the hobby. In particular, the severe constraints imposed on the visual channel create challenges for effective communicative practices, especially given the potentially complex nature of finding a specific object among many plausible candidates. Apparently to compensate for this “loss,” amateur astronomers rely on various alternative ways of communicating information relevant to the ongoing task and the body is frequently a protagonist in these processes.

Particularly noteworthy, we have observed that bodies may be used literally as inscription surfaces to communicate and explicate details of celestial scenes (say, as seen through a telescope’s eyepiece). As an example, in an episode during the Mt Hillview event of 9/7/2002, Sally was searching for the M103 open star cluster in the constellation of Cassiopeia. Upon finding it, Sally offered to share the view with FA. FA then sat next to the telescope and Sally stood to his side. She then rested her left hand on his left shoulder and proceeded to use her right indicator finger to inscribe the telescope scene on FA’s back! To do so, she first traced a circle, starting halfway up FA’s back and moving clockwise down and up. Such a circle was to represent the circular field of view offered by the scope’s eyepiece. Sally then moved on to mark star hopping points on FA’s back, always within the limits of the circle and always discursively elaborating on her actions, checking on FA’s understanding as she went along.

In a similar manner, in our most recent study in the High Meadows site (5/3/2014), we documented an event in which a senior astronomer inscribed the Big Dipper asterism on the palm of a visitor’s (a young girl) hand. The goal was to explain to the girl how to find the star Polaris by following an imaginary straight line extending from two of Big Dipper’s stars—a commonly known strategy.

In all, these events show how seeing a celestial object is literally an interactional achievement in that it requires touching bodies as a way to convey technical and observational information. Simultaneously, the episodes illustrate how modalities of embodied action are coordinated in activity, one elaborating on the other so as to achieve complex meaning that neither modality could individually convey (Goodwin, 2000, 2011).

Gesturing

Gestures (Kita, 2000; Roth, 2000) permeate all aspects of observational amateur astronomy, from planning an observation to carrying it out and explaining one’s “seeing” to others. We organize our narrative around the functional uses of gestures throughout the collaborative problem solving process of seeing an object, folding into it an account of how gestures appeared in conjunction with other modalities of embodied action.

Highlighting

Highlighting refers to a set of practices that “divide a domain of scrutiny into a figure and a ground, so that events relevant to the activity of the moment stand out” (Goodwin, 1994, pp. 610). In an archeological excavation, for example, an experienced researcher may highlight to an assistant the conceptually relevant marks on a patch of dirt by gesturing with a dowel to selectively single out aspects of the dirt and elaborating verbally on what constitutes foreground and background in a “scene” (Goodwin, 2000, 2003). In amateur astronomy, highlighting takes on various forms and they are particularly ubiquitous when astronomers communicate to one another specific

information relevant to a star hopping sequence.

To exemplify, consider an episode in which two astronomers (Guy and Bob) set out to observe the Sombrero galaxy. Guy was the more experienced of the two and he took the lead in charting a path to their common target. After some studying of a few charts, Guy then proceeded to explain to Bob his star hopping sequence. Switching his gaze back and forth between the night sky and the star chart he held with his left hand, Guy traced a “C” pattern connecting objects in the vicinity of the target (Figure 1), all along narratively elaborating on aspects of the pattern to which he was pointing. By doing so, he highlighted to Bob (and others) an emergent, relevant pattern among many other possibilities and this pattern could be used to anchor the path to the target.

As a whole, highlighting practices of amateur astronomers almost always picked either emergent features of sets of celestial objects (as above) or well known, landmark object or configurations (e.g., the Big Dipper). As a rule, the gestures making up these practices were performed over various media, including the sky, star charts, computer screens, picture books, and others.



Figure 1. Four snapshots showing Guy’s tracing of a “C” shape to highlight a set of adjacent objects. For about one minute, he repeated the gesture 3 times, tracing the “C” back and forth (7/26/2014, High Meadows).

Measuring and way finding

Again as part of communicative and coordination needs, participants use gestures to create relational, on-the-fly measures that can aid in locating objects in a star hopping sequence. To illustrate with a single case, consider an episode in which Jack and Brian had just finished confirming sight of a particular celestial target. Jack passed the telescope seat to Brian and proceeded to describe what he (Jack) had seen. The following interaction ensued.

- Jack: ((to Brian)) Okay, if you notice you can... when you look at the Telrad ((scope viewfinder)) you can use your averted ((vision)) a little bit and move... it is a little above... see the three stars ((points with left indicator finger to the sky and gestures a line through three adjacent stars)) and the one that is directly above it ((gestures up and down on a line)), it is about half again above it... when I went that far that’s what I got.
- Brian: [((inaudible))]
- FA: [Half what? It’s about half what?]
- Jack: ((to FA)) Uhh, when you go three stars ((points with left indicator finger to third star in the sequence)) and then there’s a star above it ((“holds” two points with left indicator and thumb, Figure 2.a)). You go half that distance up ((points with right indicator to a region above the points fixed by the left hand, Figure 2.b))... or half again.
- FA: Okay.

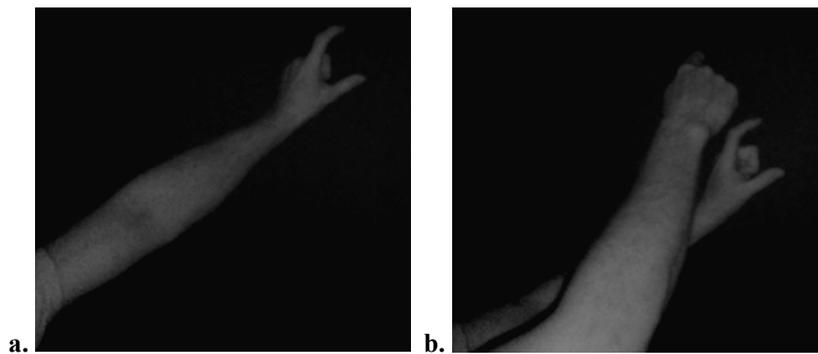


Figure 2. (a) Jack “holds” two points (stars) that were key in his star hopping procedure—his left thumb pointing to one and left indicator pointing to another. The distance between the two fingers then emerges as a measuring stick in the star hopping sequence; (b) Jack points with his right indicator to the direction of the hopping sequence, stating that target is to be found half a measuring stick distance added to last star in the hopping sequence (10/26/2014, High Meadows).

Beginning in turn 1, Jack went briefly through the star hopping procedure that they had previously worked out, in the process using an improvised measure of distance to locate the position of the target celestial object. Such a distance was a gestural yardstick (Figure 2.a) that had just emerged as an artifact of the pointing to stars in the hopping sequence. As Jack elaborated in turn 3, *half* of that measure was now to be added atop of the last star in the hopping sequence—thereby flexibly reusing the emergent yardstick to communicate object location. As with highlighting practices, gestures used in measuring always appeared in coordination with other modalities of embodiment (tool use and speech) and it too occurred always across various media that formed the infrastructure to the practice.

Conclusions and implications

We have begun documenting the embodied cognitive practices of observational amateur astronomy with the goal of contributing to our general understanding of knowing and learning in STEM practices. The human body has been conspicuously absent from accounts of knowing and learning and our work here is meant to help ameliorate this gap.

In all, we have found many results that align with the extant literature, including tool use and gestures as essential forms of embodiment for knowing and communicating in action (Goodwin, 2011). We have also uncovered some novel forms of embodied activity, such as the use of averted vision (as a way of improving one’s seeing of faint celestial objects) or inscribing the body with astronomical scenes, which we showed to reflect both the core disciplinary aspects of astronomy but also the conditions of observational amateur astronomy practice. As we extend our investigations within and beyond amateur astronomy, we expect to continue uncovering a variety of new forms of embodied action and reasoning in STEM disciplines and to shed further light on processes of knowing and learning.

References

- Azevedo, F. S. (2013). The tailored practice of hobbies and its implication for the design of interest-based learning environments. *Journal of the Learning Sciences*, 22(3), 462-510.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Cambridge University Press.
- Dickinson, T., & Dyer, A. (2008). *The Backyard Astronomer’s Guide*. Buffalo, NY: Firefly Books.
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, and R. L. Punamaki (Eds.), *Perspectives on activity theory* (pp. 19-38). Cambridge, MA: Cambridge University Press.
- Gibson, J. J. (1979). *An ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3), 606-633.
- Goodwin, C. (2000). Action and embodiment within situated human activity. *Journal of Pragmatics*, 32, 1489-1522.
- Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet*, (pp. 217-241). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Goodwin, C. (2011). Building action in public environments with diverse semiotic resources. *Versus*, 112-113, 169-182.
- Hall, R. (1996). Representation as shared activity: Situated cognition and Dewey's cartography of experience. *Journal of the Learning Sciences*, 5(3), 209-238.
- Hall, R., & Nemirovsky, R. (2012). Introduction to the special issue: Modalities of body engagement in mathematical activity and learning. *Journal of the Learning Sciences*, 21(2), 207-215.
- Hall, R., & Stevens, R. (1995). Making space: A comparison of mathematical work in school and professional design practices. In S. L. Star (Ed.), *The cultures of computing*, (pp. 118-145). Oxford, UK: Blackwell Publishers.
- Hammersley, M. and Atkinson, P. (1995). *Ethnography principles in practice*. London: Routledge.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practices. *The Journal of the Learning Sciences*, 4(10), 39-103.
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and Gesture*, (pp. 162-185). Cambridge, MA: Cambridge University Press.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.
- Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Lynch, M., & Woolgar, S. (1990). *Representation in Scientific Practice*. Cambridge, MA: MIT Press.
- Nemirovsky, R., Rasmussen, C., Sweeney, G., & Wawro, M. (2012). When the classroom floor becomes the complex plane: Addition and multiplication as ways of bodily navigation. *Journal of the Learning Sciences*, 21(2), 287-323.
- Roth, W.-M. (2000). Gestures: Their role in teaching and learning. *Review of Educational Research*, 71, 365-392.
- Saxe, G. B. (1991). *Culture and cognitive development: Studies in mathematical understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stevens, R. R. (2012). The missing bodies of mathematical thinking and learning have been found. *Journal of the Learning Sciences*, 21(2), 337-346.
- Stevens, R. R. (2000). Divisions of labor in school and in the workplace: Comparing computer and paper-supported activities across setting. *Journal of the Learning Sciences*, 9(4), 373-401.
- Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in techno-science. In M. Lampert and M. L. Blunk (Eds.), *Talking mathematics in school: Studies of teaching and learning* (pp. 107-149). New York, NY: Cambridge University Press.

Acknowledgments

We thank the astronomy clubs and amateur astronomers who participated in this study for their time and friendship. We also thank three anonymous reviewers for their comments and suggestions.

Designing a Data-Centered Approach to Inquiry Practices With Virtual Models of Density

Jonathan M. Vitale, Jacqueline Madhok, and Marcia C. Linn,
jonvitale@berkeley.edu, jjmadhok@berkeley.edu, mclinn@berkeley.edu
University of California, Berkeley

Abstract: New standards advocate for instruction that combines disciplinary knowledge and science and engineering practices. In this study we explore the design of an 8th grade science curriculum featuring interactive virtual models and guided graph construction exercises to support learning about density and inquiry practices. We apply observations about students' difficulties in the 1st iteration to the redesign of the curriculum. In particular, we observed that students' often misinterpreted the virtual models or used them inappropriately to confirm prior ideas. Similarly, even when guided to construct appropriate graphs, students were unable to conceptually link the abstract representation to concrete conceptual knowledge. On the other hand, by centering student thinking on the data in the 2nd iteration and using the models to illustrate abstract concepts conveyed in graphs, students were more likely to perform appropriate investigations and develop coherent, integrated explanations.

Introduction

In this paper we explore the use of graphs and virtual models to facilitate middle school students' understanding of density. Both models and graphs are tools used in authentic science practice because they facilitate inquiry and reveal complex patterns. Virtual models have the potential to facilitate systematic testing of ideas with accessible, concrete representations (Smith, Snir, & Grosslight, 1992) when students apply appropriate scientific practices (Perkins & Grotzer, 2005). Graphs can supplement virtual models by highlighting general principles and illustrating complex relationships (Wilkerson-Jerde & Wilensky, 2014). To benefit, students need to translate between the abstract features of graphs and the underlying scientific context (Shah & Hoeffner, 2002). We investigate ways to guide students through interactive graph activities that link the concrete features of a model with abstract features of a graph to support conceptual development of the concept of density.

Student ideas about density

Density is an important, but challenging concept for children to grasp. Young children rarely differentiate mass and volume, thus precluding a notion of density (Kohn, 1993; Piaget, 1972; Smith, Carey, & Wiser, 1985). Older children are able to differentiate mass and volume, but do not immediately comprehend how the relationship between the two determines an object's buoyancy or is related to the object's material composition (Snir, 1991). Rather, students generate a variety of ideas that apply to specific situations. Some attribute sinking behavior exclusively to the mass or weight of an object, others consider mass when size is equal (Howe, 2002; Linn & Pulos, 1983). To address students' varied ideas, researchers have tried using demonstrations or models that directly contradict intuitive beliefs. For example, a teacher might demonstrate that a large, heavy board of wood floats, while a small, light pebble sinks. These demonstrations rarely result in long-term impacts because they do not address all of students varied ideas, and because students may lack the ability to generalize from limited cases (Perkins & Grotzer, 2005).

Although virtual models allow for rapid testing of multiple conjectures, including the specific ideas held by the student, they may still pose challenges for generalization. In many cases, virtual models generate similar learning outcomes as their physical counterparts (Klahr, Triona, & Williams, 2007). Yet, virtual models may add value if they offer representations that help students attend to underlying relationships and mechanisms. For example, Smith, Snir, and Grosslight (1992) introduced a "grid-and-dots" representation to demonstrate the underlying mechanism of density (i.e., more dots within a grid box represents higher density). Likewise, graphs can capture general relationships: Rowell and Dawson (1977) presented graphs to depict density as a constant ratio between mass and volume. Specifically, 9th grade students were prompted to measure and calculate the mass and volume of objects, graph this data, and draw lines of best-fit. In this activity, identical materials are graphed along a line with a constant slope. Thus, in this case, the graph representation served two functions: as a tool for inquiry and a means to illustrate an important concept. Yet, despite the potential for this activity, nearly three-quarters of the students did not recognize that a single, constant-slope line should pass through graphs of identical materials. Whether the graphing representation itself or its deployment in instruction inhibited learning gains is unclear.

Knowledge integration and graphs

Regardless of the representations of density, the way student activities are contextualized within a curriculum can have a substantial impact on learning. For example, in a related study on water displacement Linn and Eylon (2000) found that assigning students to reflection activities had a greater influence on learning than particular features of the model or representations (i.e., level of control). Research suggests that beyond specific representations, a curriculum is successful when it facilitates knowledge integration – i.e., a proactive process by the learner to resolve disparate ideas and construct coherent explanations (Linn & Eylon, 2011). Curriculum designers can support knowledge integration by prompting students to explain their initial ideas, observe new phenomena to generate new ideas, distinguish discrepant ideas, and synthesize their understanding through reflection (Linn & Eylon, 2011). Virtual models and graphs, particularly with adaptive guidance, can facilitate this process by highlighting important discrepancies between ideas.

Following this framework, interactive computer models and graphing activities can provide alternative perspectives on the same concepts. Students, therefore have an opportunity to test novel, potentially discrepant ideas. Automated assessment and guidance can then be used to challenge student ideas (Vitale, Lai, & Linn, 2015). In this study we explore the initial development and redesign of a curriculum created to introduce virtual models and graphing activities with automated guidance to support knowledge integration.

Research overview

The design of the density curriculum was informed by principles of knowledge integration (Linn & Eylon, 2011) and dynamic visualization (McElhaney, Chang, Chiu, & Linn, 2014). To resolve the many open questions concerning how best to take advantage of graphing and modeling activities an exploratory research approach was applied. Design-based research (DBR) is an emerging methodological perspective that facilitates initial theory building while concurrently addressing instructional goals (Barab & Squire, 2004). Although specific DBR approaches vary widely, we utilized the following set of assumptions to guide this research:

- While researchers may have initial conjectures, surprises should be expected (Sandoval, 2014).
- Teacher, researcher, and peer interaction play an important role in learning and should be explored rather than minimized (Collins, Joseph, & Bielaczyc, 2004).
- Due to the surprising nature of DBR, classroom practices and instructional material design should be flexible to emergent circumstances (Collins et al., 2004)
- Multiple data sources are necessary to analyze complex situations (Barab & Squire, 2004).

Guided by these assumptions we implemented the density curriculum in two design iterations to investigate the following research question: how can virtual models and graph activities be designed to mutually support inquiry and learning? We used observational notes and quantitative data to evaluate the effectiveness of our first design and apply what we learned to a re-design. We then implemented this redesign in two classrooms from the same school in the following year and compare the learning performances between design iterations.

Iteration 1

We designed iteration 1 to establish a baseline for performance on instructional and assessment items and to discover patterns of successful students. We used these results to refine the curriculum.

Student and school sample

Iteration 1 was conducted with two 8th grade teachers and their 205 students. Ms. S (173 students) had over 20 years of science teaching experience. Ms. P. (32 students) was a second year teacher. The school serves a diverse community (42% reduced lunch, 13% ELL) in a suburban area of the western United States. Students performed the pretest and posttest individually. For the curriculum, although students were expected to provide their own responses, they were advised to collaborate with an adjacent student.

Materials

All computer materials, including pretest and posttest were constructed in the Web-based Inquiry Science Environment (WISE). The main curriculum unit, “Sink or Float”, can be previewed at this link: <http://wise.berkeley.edu/previewproject.html?projectId=11723>.

Pretest and posttest

We constructed a test with 7 multiple choice and 4 open response items (subset coded by two raters, κ 's > .8). Several items prompted students to explain whether specified objects would float, while other items prompted students to apply this knowledge to the following graph of mass vs. volume:

Mass vs. volume graph interpretation. The graph in Figure 1 displays the mass and volume of 6 points representing objects placed in water. Items 3 and 5 target the concept of density as an intensive property of material. Items 6 and 7 target the relationship between mass-volume ratio and flotation.

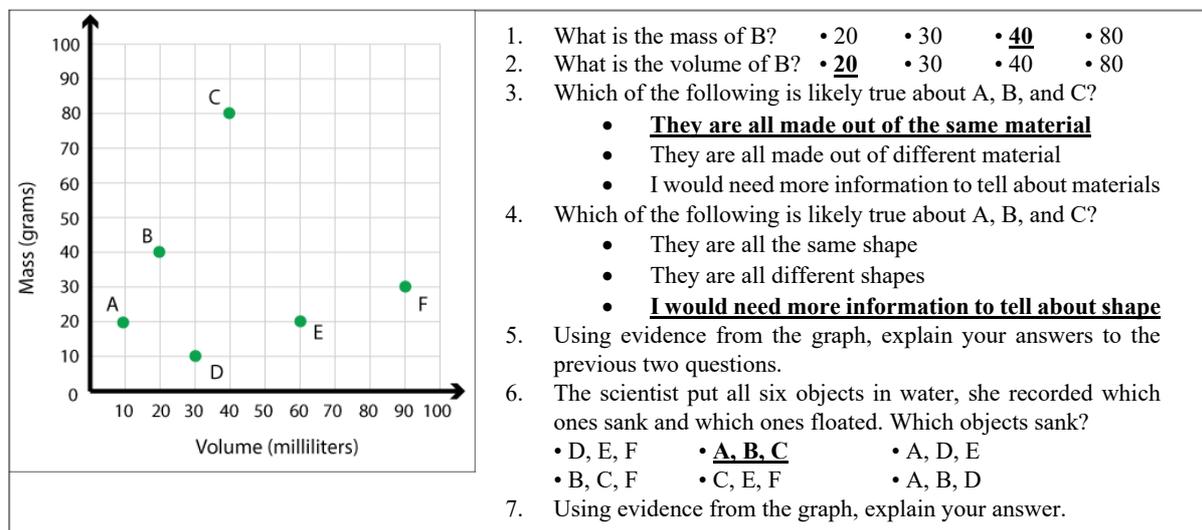


Figure 1. Mass vs. volume graph pre-post items. Correct multiple choice answer are bolded and underlined.

Curriculum unit

“Sink or Float” was developed to facilitate understanding of density and buoyancy with a series of open investigations. In addition to exploratory models and graphs each activity began with relevant prediction questions and finished with reflection prompts to engage students in knowledge integration.

Virtual laboratory. Developed with a popular open source physics engine, the virtual laboratory provided students with an open and realistic environment to test ideas (Figure 2). Students were prompted to investigate claims made by fictional characters (e.g., Aida: “It’s the mass that matters, not the volume.”) by constructing blocks from diverse materials, measuring their mass and volume, and testing if they sank or floated. Data (material composition, mass, volume, and sink/float) was automatically displayed on a following “table-explanation” html page, which prompted students to explain whether the tabular data supported the fictional claim. Later in the unit (following “materials simulation”, Figure 4) this data was re-displayed with density and students were prompted to find a relationship between density and other object features (material composition).

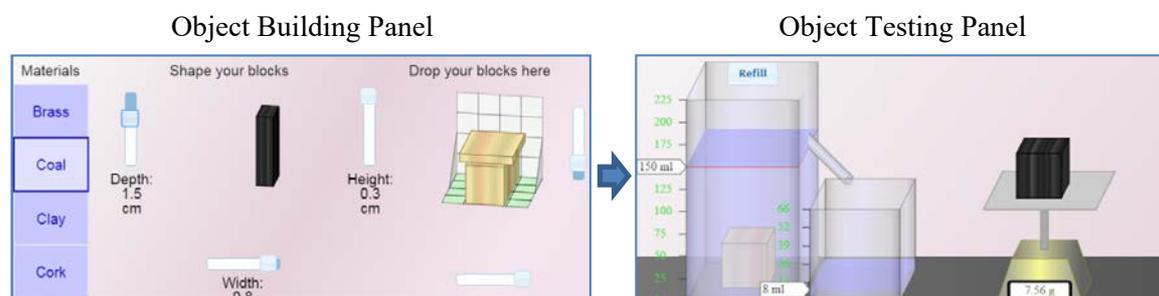


Figure 2. Virtual laboratory. In these panels (displayed vertically in curriculum), students construct objects out of various materials, and then test the materials with virtual measurement tools.

Mass vs. volume graph construction. Following the virtual laboratory, students were prompted to plot points representing objects that would sink, objects that would float, and then draw a line to divide all potential points (Figure 3). Automated guidance was provided to alert students to the number of incorrect points or problems with their dividing line (e.g., not starting from the origin). Students had unlimited opportunities to revise the graph. Graphs were scored with a knowledge integration rubric (Vitale et al., 2015) that distinguished between 5 levels

of complexity: irrelevant (KI = 1), non-normative (e.g., points divided by horizontal line, KI = 2), partial (points distributed in small, but correct area, KI = 3), full (points distributed widely across correct area, divider not correct, KI = 4), and complex (points correct, divider correct, KI = 5).

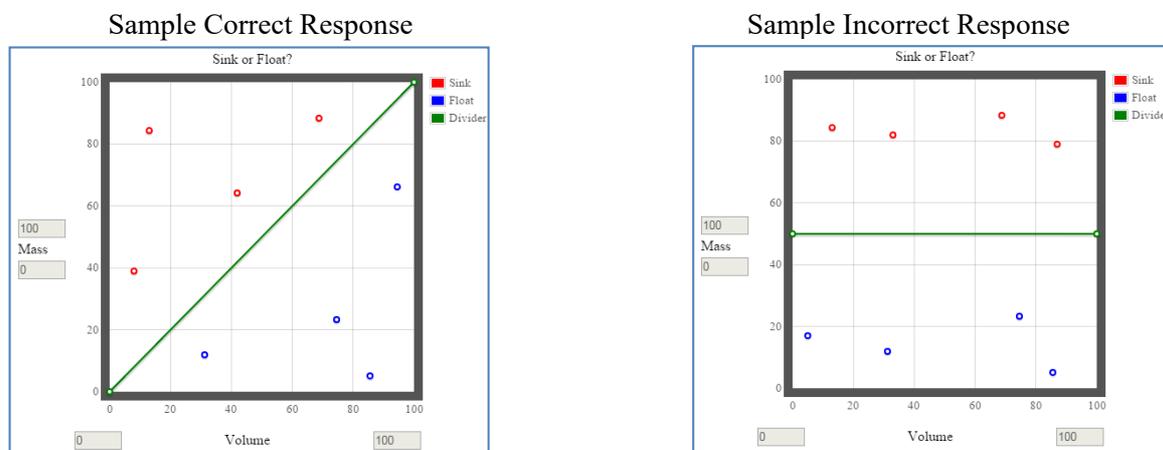


Figure 3. Students plot points to represent objects that sink (red), float (blue), a line to divide points.

Materials simulation. In this activity students explored relationship between mass, volume, density, and buoyancy of objects made from one of eight materials. Figure 4 displays four questions posed to students as they engaged the simulation (although we focus on 2 – 4, because 1 could be answered as “mass” or “density”).

1. What changes when you switch materials?
 - **Mass** • Volume • **Density**
2. What stays the same as you change the shape of an object?
 - Mass • Volume • **Density**
3. What happens why you double the volume of an object (and keep the material the same)? [M is mass, D is density]
 - M doubles, D doubles • M halves, D halves
 - M stays the same, D doubles • M stays the same, D halves
 - **M doubles, D stays the same** • M halves, D stays the same
 - M doubles, D halves • M halves, D doubles
 - M stays the same, D stays the same
4. Which of the following determines how fast an object will sink or float?
 - Mass • Volume • **Density** • Width

Figure 4. Students test properties of objects made from 8 materials, and then respond to questions.

Analysis

Pretest and posttest

Overall, scores from pretest to posttest rose significantly with a medium effect size [$t(203) = 7.2, p < .001, d = 0.5$]; however, focusing on the graph items (Figure 4), the effect size was substantially lower [$t(200) = 3.0, p < .01, d = 0.2$], suggesting that the curriculum was less effective at promoting learning on the graphing applications than for more general applications (e.g., predicting whether a wooden table would float).

Curriculum unit

Virtual laboratory. For the “Test Aida’s Idea” (mass determines sinking) step, 163 of 226 students performed valid work (i.e., constructing at least one object). Of these 163 students only 27% constructed objects that could falsify Aida’s claim (i.e., at least one sinking object with mass less than or equal to a floating object).

Likewise, on the following table-explanation step, 66% of students (incorrectly) agreed with Aida’s claim or gave a mass-based explanation for sinking, while only 19% (correctly) disagreed with Aida or gave a density/material-based explanation for sinking. Of those 44 students who constructed objects that would falsify

Aida's claim less than half (39%) explicitly disagreed with Aida. Even of those that did, only two made explicit connections between data in the table and Aida's error.

Mass vs. volume graph construction. On their final graph 46% of students produced a graph that displayed an irrelevant or non-normative understanding about the relationship between mass and volume (KI = 1, 2), while 39% of students produced graphs with correct, well-distributed points (KI = 4, 5).

Materials simulation. In this step 47%, 46%, and 57% of students observed that density changes as materials change, remains the same as volume doubles, and determines speed of sinking, respectively.

Case study

To better illustrate the difficulties and affordances of the activities we present the case of a student who was clearly engaged and made gains from pretest to posttest (> 8 gain points), but nonetheless struggled with the initial simulation and graphing activities. We used log data to follow this student's trajectory.

Student "Jen". Jen initially agreed with Aida's claim (mass determines sinking) and constructed identical 8 ml cubes from each of the four materials. By controlling for volume, Jen implemented a special case in which Aida's claim is true, thus reinforcing her prior belief. As such, on graph construction Jen plotted sink and float points separated by mass, dispersed across the x-axis (like Figure 3, right). After receiving guidance, Jen returned to the simulation and the table-explanation step, although she did not construct any new objects. She then returned to the graph step and was able, presumably with guidance, to move points to the correct region of the graph. However, Jen still retained a horizontal dividing line and provided explanations on subsequent reflection exercises that demonstrated confusion between mass and density. On the materials simulation Jen correctly identified the role of density in the three target questions. In following steps Jen clearly linked density to materials.

Discussion

Both quantitative and qualitative results indicate that the virtual laboratory did not effectively challenge students' prior conceptions. It is notable that Jen constructed objects of identical volume, but on the graph plotted points across a range of volumes. In this case a misapplied scientific practice (i.e., controlling variables) limited Jen as she used the model, whereas no such prior knowledge about exploring continuous graphs impeded construction of more diverse artifacts. Yet, even for those students who generated appropriately diverse objects in the virtual lab, less than half applied this knowledge to contradict Aida's claim. Conversely, while the graph may have provided a stronger alert of an error, the graph construction proved insufficient for generating new ideas. Jen's return to the virtual laboratory after receiving graph guidance suggests that, even though it did not lead to new ideas, the concrete model provides a more accessible representation of concepts.

On the other hand, the "materials simulation" clearly helped Jen understand the link between materials and density. This simulation closely integrated the visual model with numerical data, facilitating rapid testing of ideas. This suggests that reframing activities with virtual models to focus students on data representations may enhance inquiry practices and learning.

Iteration 2

To implement a data-centered approach to the application of virtual models we performed two major changes to the curriculum. For the virtual laboratory we randomly assigned either a table or graph in both the prompt (for critique) and the model output to determine which representation is more effective. For graph construction, to help students link the models and graphs, we displayed virtual model that depicted objects represented in the graph. By comparing the outcomes and behaviors of students in this iteration to those of iteration 1 we test whether greater model-graph integration facilitates more exploratory behavior and stronger learning outcomes.

Student and school sample

Iteration 2 was conducted within the same school as iteration 1 in the following year with a new cohort of 239 8th grade students. This study included Ms. S from iteration 1 (140 students) and a first year teacher (Ms. L, 99 students). Unlike iteration 1 students performed the curriculum unit in assigned pairs.

Materials

The pretest and posttest remained the same as iteration 1. Changes in the curriculum unit are detailed below and can be previewed at this link: <http://wise.berkeley.edu/previewproject.html?projectId=14827>.

Curriculum unit changes

Although several minor changes were made, we focus here on changes to steps explored in iteration 1. By merging the affordances of the “materials simulation” to the virtual laboratory and the graph construction exercise we hoped to make this simulation redundant, and thus removed it from the curriculum.

Virtual laboratory. Students were randomly assigned to a model that displayed either a graph (Figure 5) or table displayed below the laboratory during testing and in a pop-up dialog immediately following a completed measurement. Additionally to facilitate greater attention to the data and more diverse construction, students were presented with either a graph or table (matching their output) in the step prompt depicting the non-normative under investigation. For example, “Aida’s graph” shows points separated by mass (Figure 5).

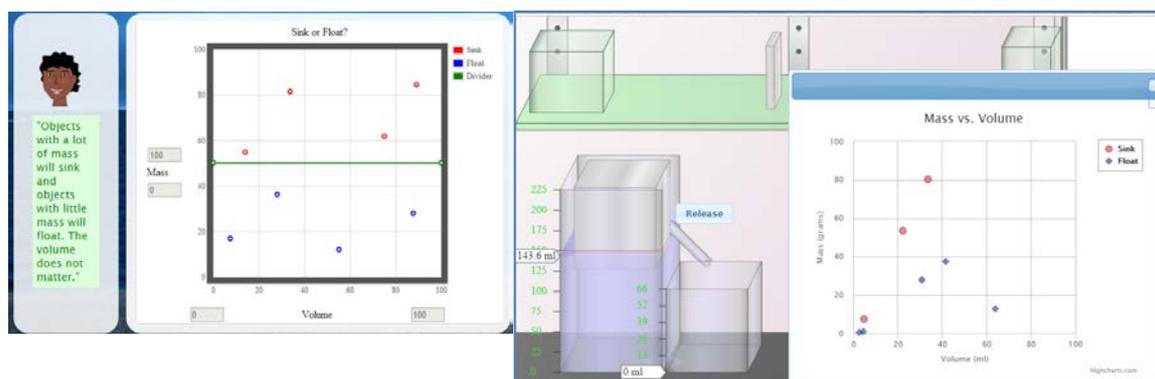


Figure 5. Virtual laboratory with graph. Students presented with a graphical interpretation of non-normative idea and prompted to investigate. As students conduct measurements a corresponding graph is displayed.

Mass vs. volume graph construction. To help students link data to a concrete representation, we displayed a simple model illustrating the properties of points selected on the graph with a block whose darkness and size corresponded to the density and volume, respectively. Additionally, students could drag points to observe continuous changes in features of the model. Post submission text guidance was similar to iteration 1 (e.g., stating number of incorrect points) but prompted students to observe the model.

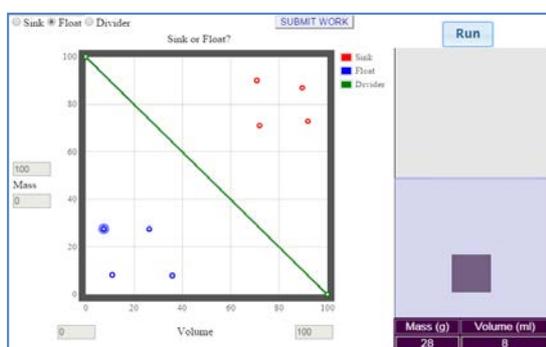


Figure 6. Graph sink and float with linked model displaying properties of selected points on graph. In this case, a point at (8, 28), incorrectly labeled as “Float” (blue), is shown to sink.

Analysis

Pretest and posttest

Students made large gains from pretest to posttest [$t(238) = 12.3, p < .001, d = 0.8$]. Compared to iteration 1, students had similar pretest scores [$t(441) = 0.7, p > .1$], but achieved significantly higher gains in this iteration [$t(441) = 3.4, p < .001, d = 0.3$]. We found no particular differences between students who received graphs or tables to represent their data; however, this needs further exploration to make any strong conclusions.

Curriculum unit

Virtual laboratory. For “Test Aida’s Idea” 65 of 120 workgroups constructed objects that could falsify Aida’s claim. In contrast to the 27% from Iteration 1, this represents a significantly higher proportion [$\chi^2(1) = 20.4, p <$

.001]. Additionally, less than half (48 of 117) of the students agreed with Aida’s claim after completing this activity, representing a significantly lower proportion than iteration 1 [$\chi^2(1) = 18.0, p < .001$].

Mass vs. volume graph construction. Unlike iteration 1, new logging software allowed for a more fine-grained analysis of student work. Figure 7 displays score distributions for initial and final graphs produced in iteration 2 and the final graphs for iteration 1 for comparison. Final scores for iteration 2 were shifted significantly higher (using Wilcoxon for non-normal distributions) than iteration 2 initial scores [$V = 716, p < .001$]. Additionally, the final iteration 2 scores were shifted significantly higher than the iteration 1 final scores [$W = 8336, p < .001$], demonstrating a greater value for the revised, model-enhanced guidance.

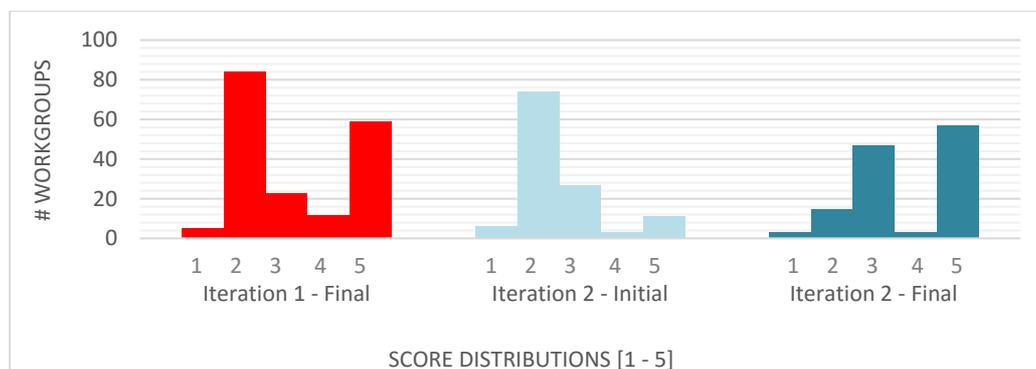


Figure 7. Graph construction score distributions, by iteration. Each distribution shows frequency of each score from 1 to 5, with modes at 2, 2, and 5, from left-to-right.

Classroom observations. In addition to the curriculum changes noted above, Ms. S incorporated new discussion activities into her instruction. One such activity was a discussion of “claims, evidence, and reasoning”, where students were asked to discuss each with peers. Ms. S then presented students with an example response from the WISE unit that displayed unclear use of evidence to justify a claim and asked students to rewrite this response and share their new reasoning. As such 52% of Ms. S’s students recognized that Aida’s claim was false, while only 33% of Ms. L’s students arrived at this conclusion [$\chi^2(1) = 3.1, p = .08$].

General discussion

Quantitative analyses demonstrate significant improvement between iterations that reflect increases in the coherence of student ideas. Determining the factors leading to improvement is limited by the quasi-experimental nature of the study. However, the similarity of pretest scores suggests that the two groups of students from the same school had similar prior knowledge and abilities. Of course, it is unclear whether the improvement was due to changes in the curriculum or to improved implementation strategies. We suspect a combination of both.

Regarding the curriculum, the links between their investigations and the points on the graph clearly helped students to better coordinate tests of materials. This resulted in more instances of “falsification” for the second iteration than for the first. The teacher-led discussion of claims, evidence, and reasoning likely played a role by defining what is expected as strong evidence to justify or falsify a claim.

For the graph construction, the large shift from initial to final graphs demonstrates the value of the linked model. Observation of students revealed a number of new strategies in the revised version of the curriculum, including the production of numerous data points to test diverse possibilities and fill the space of the graph. Some students took advantage of the real-time link by dragging points across the graph space and observing simultaneous changes in the model, thus adding a dynamic component to their use of the graph. Future work will investigate these and additional graph interpretation strategies.

Conclusions

Computer models are an important tool for scientists and engineers, but can be challenging for middle school students to use appropriately. Designing models to meet the needs of students requires analysis of their prior knowledge and likely strategies. Students often need new strategies as well as guidance to take advantage of these types of models and conduct effective investigations (McElhaney et al. 2014; Perkins & Grotzer, 2005). In the first version of the curriculum students primarily used investigations to confirm their beliefs. In this case the realism and familiarity of the virtual tools may have limited students’ creative expression. The revision of the curriculum shifted the data representations to a more central role for idea development and used the models in a

supplementary role to illustrate abstract concepts. The gains between iterations suggest that this configuration led to greater diversity of ideas and stronger links between evidence and theory.

More generally, this study suggests that data representations, specifically graphs, represent an important space for students to apply, elaborate on, and distinguish between their ideas. The graph used in this study allowed students to explore changes in volume and mass systematically. Activities requiring the production and interpretation of graphs, when carefully designed, compel students to identify general principles and think about the relationship between critical variables (Shah & Hoeffner, 2002). Thus, graphs provide a framework for thinking about models productively.

In addition to illustrating specific disciplinary concepts, graphing is an authentic tool that is applicable across disciplines (NGSS Lead States, 2013). This provides teachers with an opportunity to address science practices explicitly in their instruction using concrete examples from student work. As the discussion of “claim, evidence, and reasoning” given in iteration 2 suggests, this can be an effective way to both encourage more substantive investigation in a specific activity and link the activity to more general epistemological themes.

References

- Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *The Journal of the Learning Sciences*, 13(1), 1–14.
- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, 13(1), 15–42.
- Howe, A. (2002). *Engaging children in science*. Upper Saddle River, NJ: Merrill Prentice-Hall.
- Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school students. *Journal of Research in Science Teaching*, 44(1), 183–203.
- Kohn, A. (1993). Preschoolers' reasoning about density: Will it float? *Child Development*, 64(6), 1637–1650.
- Linn, M. C., & Eylon, B. (2000). Knowledge integration and displaced volume. *Journal of Science Education and Technology*, 9(4), 287–310.
- Linn, M. C., & Eylon, B.-S. (2011). *Science learning and instruction: Taking advantage of technology to promote knowledge integration*. New York: Routledge.
- Linn, M. C., & Pulos, S. (1983). Male-female differences in predicting displaced volume: Strategy usage, aptitude relationships, and experience influences. *Journal of Educational Psychology*, 75(1), 86–96.
- McElhaney, K. W., Chang, H.-Y., Chiu, J. L., & Linn, M. C. (2014). Evidence for effective uses of dynamic visualisations in science curriculum materials. *Studies in Science Education*, 51(1), 49–85.
- Perkins, D. N., & Grotzer, T. a. (2005). Dimensions of causal understanding: the role of complex causal models in students' understanding of science. *Studies in Science Education*, 41(1), 117–165.
- Piaget, J. (1972). *The child's conception of physical causality*. Totowa, NJ: Littlefield, Adams.
- Rowell, J. A., & Dawson, C. J. (1977). Teaching about floating and sinking: An attempt to link cognitive psychology with classroom practice. *Science Education*, 61(2), 245–253.
- Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, 23(August 2014), 18–36.
- Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14(1), 47–69.
- Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concepts of size, weight, and density. *Cognition*, 21, 177–237.
- Smith, C., Snir, J., & Grosslight, L. (1992). Models to facilitate using conceptual change: The case of weight-density differentiation. *Cognition and Instruction*, 9(3), 221–283.
- Snir, J. (1991). Sink or float—what do the experts think?: The historical development of explanations for floatation. *Science Education*, 75(5), 595–609.
- States, N. L. (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academy Press.
- Vitale, J. M., Lai, K., & Linn, M. C. (2015). Taking advantage of automated assessment of student-constructed graphs in science. *Journal of Research in Science Teaching, Advance on*. doi:10.1002/tea.21241
- Wilkerson-Jerde, M. H., & Wilensky, U. J. (2014). Patterns, probabilities, and people: Making sense of quantitative change in complex systems. *Journal of the Learning Sciences*, (May 2015), 1–48.

Fostering University Freshmen's Mathematical Argumentation Skills With Collaboration Scripts

Freydis Vogel, Technical University of Munich, freydis.vogel@tum.de
Ingo Kollar, University of Augsburg, ingo.kollar@phil.uni-augsburg.de
Stefan Ufer, Ludwig-Maximilians-Universität Munich, ufer@math.lmu.de
Kristina Reiss, Technical University of Munich, kristina.reiss@tum.de
Frank Fischer, Ludwig-Maximilians-Universität Munich, frank.fischer@psy.lmu.de

Abstract: Students often have problems formulating and using arguments in mathematical contexts. Therefore, we investigated to what extent a collaboration script helps students overcome their problems and acquire mathematical argumentation skills. In two previous studies, we showed that collaboration scripts can have positive effects on learning cross-domain argumentation skills in the mathematical context. Yet, the effectiveness of the script depended on individual prerequisites such as final high school grade (GPA) and self-regulation skills. In this study, $N = 96$ participants learned in one of three script conditions. We found that a high-structured domain-general collaboration script for argumentation was more effective for acquiring domain-specific mathematical argumentation skills than a low-structured or an adaptable one. Furthermore, only in the condition with the low-structured script, learners' self-regulated learning skills played an important role for the learning outcomes.

Keywords: mathematics, argumentation, collaboration scripts

Mathematical argumentation

Argumentation is a common mode for knowledge-generating dialogues in social science, politics, or everyday negotiation. Against this background, investigating the potentials of argumentation for knowledge-generation in a mathematical context might not be that intuitive. Yet, argumentation also plays an important role in mathematics, both in informal mathematical reasoning (Thurston, 1994) and even more so when it comes to formal proof, which is considered the ultimate form of mathematical argument. Finding promising conjectures and constructing mathematical proofs typically requires the use of heuristic strategies (rather than algorithms) and frequent comparisons between the promises and problems of different solution strategies (Reiss et al., 2008). Moreover, the construction of an acceptable proof requires the repeated production and evaluation of single arguments as well as of the structure of the overall argumentation. This process of proof construction, which is "usually held informally between mathematicians to develop, discuss or communicate mathematical problems and results" (Douek, 1999, p. 129; see also Thurston, 1994), can be considered a substantial argumentation process, requiring justifications of the strategic decisions as well as the mathematical inferences made. Thus, during the process of finding a mathematical conjecture or proof, argumentation skills are required at different points either to construct sound arguments, e.g. with the components introduced by Toulmin's (1958) model of argument (Pease, Smaill, Colton, & Lee, 2009), or to be able to engage in an argumentative dialogue.

According to an experts' process model proposed by Boero (1999), this process comprises phases of exploration as well as phases of consolidation of arguments into a deductive chain. The explorative phases are executed in a rather tentative way that includes revisions of steps made before. The discussion with others during those phases seems to be a promising way to frequently evaluate and strengthen the chosen approach to solve the proof task, since the others could bring new ideas and perspectives to refine ideas and arguments. Other phases include the selection and organization of coherent arguments which ideally lead to the formal proof requested by the initial task (for elaborated descriptions of the process of proof see Boero, 1999; Reiss, et al., 2008), which require evaluation of the arguments according to standards of the discipline. Also this evaluation process can benefit from monitoring by collaboration partners.

Apart from supporting problem solving processes, the engagement in an argumentative dialogue has more benefits to offer. The process of constructing and applying arguments in a social discursive collaborative learning process is ascribed a high potential for deeper elaboration of the learning content and thus a key moderator for the enhancement of learning processes ("arguing to learn"; Andriessen, Baker, & Suthers, 2003). Even though the formal rules for accepting proofs are quite specific to mathematics, general guidelines for the construction of arguments in a collaborative discussion as well as for a sustainable social discursive process of argumentation between dialogue partners are considered to be more or less applicable to other domains as well. For instance, Toulmin's (1958) argument model is widely used for the evaluation of single arguments in multiple domains (van

Eemeren & Grootendorst, 2004). Also the dialectical format of argumentation - simplified as the cycle of 'argument', 'counter-argument', and 'synthesis' - might function as a common ground for social discursive activities where two or more dialogue partner are engaged in an argumentative discourse (Leitão, 2000). Studies have detected that students often struggle when trying to construct valid mathematical argumentations (e.g. Heinze, Reiss, & Rudolph, 2005). Students are also often not able to engage in a deeper elaborative dialogue while learning collaboratively (Cohen, 1994). As students have problems effectively using argumentation in mathematical learning processes (Hillbert et al., 2008), it is important to find ways to appropriately scaffold students' argumentation in the context of mathematical proof tasks.

Scaffolding mathematical argumentation

One way to help students to acquire knowledge and skills related to mathematical argumentation is by offering collaboration scripts for argumentation. Collaboration scripts aim at guiding students through meaningful and beneficial collaborative learning processes by distributing roles and activities among the students (King, 2007; Kollar, Fischer, & Hesse, 2006). Collaboration scripts have shown to be effective in both face-to-face and computer-supported collaborative learning (e.g. O'Donnell & Dansereau, 1992; Rummel & Spada, 2005, Weinberger, Fischer, & Mandl, 2003). For the design of collaboration scripts, an explicit reference to argumentation frameworks proved to be helpful. For example, Weinberger, Stegmann and Fischer (2010) developed a script that guided students through an argumentative discourse by offering explicit support on crucial points in the argumentation process. For instance, when students were supposed to construct arguments or counterarguments, they were asked to address single elements of Toulmin's (1958) argument model (e.g. claim, data, qualifier). In other computer-supported collaboration scripts, students' discourses were guided by distributing the roles of discussants (e.g. De Wever, Schellens Van Keer, & Valcke, 2008) or by prompting them to perform single activities of a discursive argumentation cycle (e.g. Hron, Hesse, Cress, & Giovis, 2000).

In recent studies, the collaboration script approach has been transferred to the mathematical domain as well. In an own study, an argumentation-related collaboration script showed positive effects on students' cross-domain argumentation skills (Kollar et al., 2014), both in the learning process and a subsequent individual post test. However, positive effects on domain-specific mathematical argumentation skills could not be found. Additionally, in this study the final high school qualification grade (GPA) was significantly related to the effectiveness of the collaboration script on domain-general argumentation skills. The better the GPA grade, the more the students benefited from the collaboration script (Kollar et al., 2014). Yet, in other domains, collaboration scripts were found to exert positive effects on domain-specific learning (e.g. Weinberger, et al., 2010), albeit not consistently. As the GPA was such an important factor for the effectiveness of the collaboration script in the study by Kollar et al. it can be hypothesized that the used collaboration script did not offer enough hints and structure, especially for learners with lower GPA.

A further criticism that is often raised against the use of collaboration scripts to support collaborative learning efforts is that collaboration scripts may structure learners' activities in a too rigid way. This might lead to "overscripting", i.e. that the collaboration script inhibits learning by constraining learners in using their own, possibly effective learning strategies (Dillenbourg, 2002). One way to avoid overscripting is to offer the learners adaptive collaboration scripts. These collaboration scripts would automatically be adapted to a learner's current needs (e.g. Diziol, Walker, Rummel, & Koedinger, 2010). Unfortunately, such adaptive learning scenarios are enormously complex and expensive to develop. Therefore, a more feasible way to avoid overscripting might be to create adaptable collaboration scripts that give the learners themselves the possibility to adapt the script to their own needs. Yet, the learners might need a certain amount of self-regulation skills to benefit from such adaptable learning scenarios (Vogel et al., 2015).

Goals of the study

The goal of this study was to compare the effects of a high-structured collaboration script, a low-structured collaboration script and an adaptable collaboration script in the context of mathematical proof tasks on students' acquisition of mathematical argumentation skills. Furthermore, we investigated to what extent the learner's GPA and their self-regulated learning skills would be related to learning in each of the three collaboration scripts. In line with prior research, we expected the high-structured collaboration script to have a more positive impact on learning mathematical argumentation skills than the low-structured collaboration script. Furthermore, we also expected students in the adaptable script condition to outperform students who received the low-structured collaboration script, as the adaptable script offers each learner the amount of structure s/he (thinks s/he) needs. Therefore, we expected that the self-regulated learning skills would be positively related to learning especially in the condition with the adaptable script, but also in the low-structured condition since both offer learners the

opportunity to self-regulate their learning much more strongly than the high-structured script. We expected the GPA to be positively related to learning in the low-structured script condition.

Research questions

- RQ1: To what extent does a low-structured collaboration script, a high structured collaboration script and an adaptable collaboration script affect learner's acquisition of mathematical argumentation skills?
- RQ2: To what extent are the learners' self-regulated learning skills related to learning mathematical argumentation skills when learning with either a low-structured, a high-structured or an adaptable collaboration script?
- RQ3: To what extent is the learners' GPA related to learning mathematical argumentation skills when learning with either a low-structured, a high-structured or an adaptable collaboration script?

Methods

Participants and design

The study was implemented in a two weeks preparation course that was conducted at three different German universities in fall 2011. Participants were $N = 96$ ($n = 58$ female; $n = 38$ male) mathematics students and mathematics teacher students at the beginning of their university studies. Their mean age was $M = 19.53$ years. The students were randomly assigned to the three treatment conditions (1) low-structured argumentation script ($n = 29$), (2) high-structured argumentation script ($n = 34$) and (3) adaptable argumentation script ($n = 33$).

In each condition, we grouped students in dyads of high as well as low achievers, based on their final high school qualification grade. For each of the three learning tasks, we randomly formed new dyads within the high and low achievers groups nested in the experimental conditions to minimize the impact on knowledge acquisition one specific learning partner might exert on a given learner.

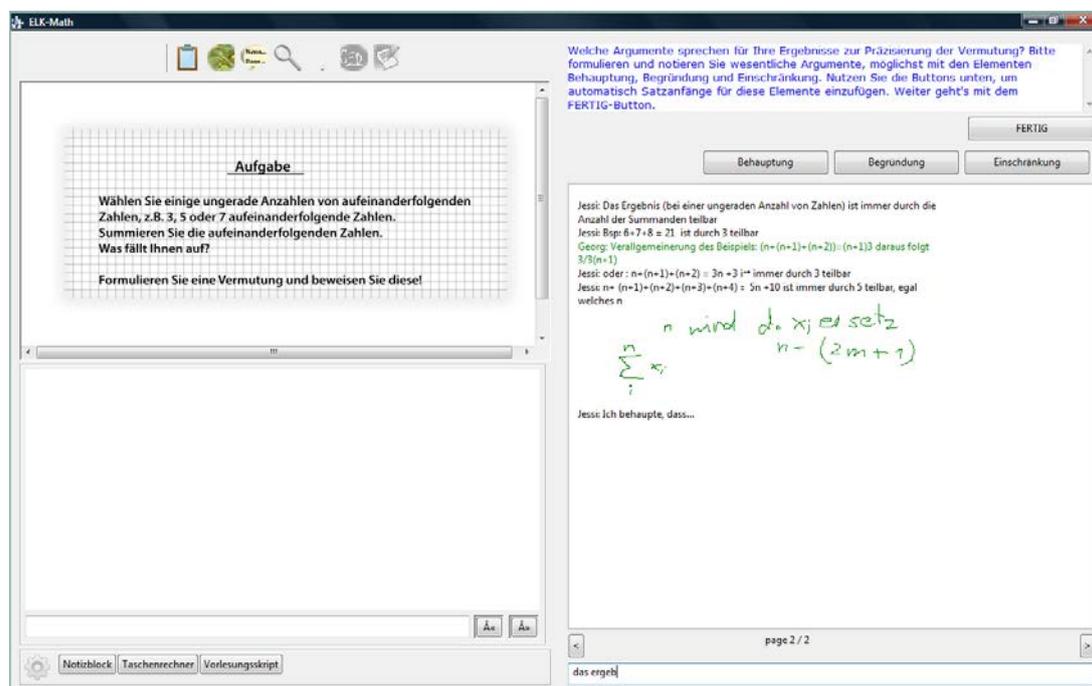


Figure 1. Screenshot of the computer-supported learning environment with the proof problem on the left side, script prompts on the upper right side and the communication area below the script prompts on the right side.

Setting and learning environment

Students learned collaboratively in dyads in a computer-supported learning environment (see Figure 1) on three different problem solving tasks in the context of mathematical proof. Additionally to the proof tasks, the learning environment offered six driving questions derived from Boero's steps of mathematical proof (Nadolski, Kirschner,

& van Merriënboer, 2006). The learners were seated vis-à-vis each other and each learner was equipped with one laptop, a graphic tablet and a mouse. The two learners of each learning dyad were connected via a computer-supported learning environment that displayed the current mathematical proof task and the script prompts in the respective treatment conditions. The computer-supported learning environment provided the learning task itself and a shared communication area where the two learning partners could exchange text messages as well as drawings. They were also asked to type the most important results of their discussion into the communication area even though they were also allowed to discuss orally.

Conditions of the learning environment and learners' pre-requisites

Type of collaboration script. Three types of collaboration scripts were compared in this study: a low-structured script, a high-structured script and an adaptable script. The students were asked to collaboratively discuss their ideas about the proof tasks in all treatment conditions. In the condition with the low-structured script, prompts repeatedly sequenced the discussion into the phases (1) argument, (2) counterargument and (3) synthesis. During each phase, prompts distributed the roles of 'talker' and 'listener' to the learning partners. The roles were switched for every new argument cycle. In the condition with the high-structured script, students were additionally prompted to formulate sound arguments (i.e. to formulate claims, data and rebuttals based on Toulmin's, 1958 argumentation model) in each of these three steps. In the condition with the adaptable argumentation script, learners were allowed to choose between the high structured and the low structured argumentation script at six points in time within each treatment session (see Figure 2). Before they adapted the script, the learners had to discuss which script they wanted to take in order to come to a joint solution.

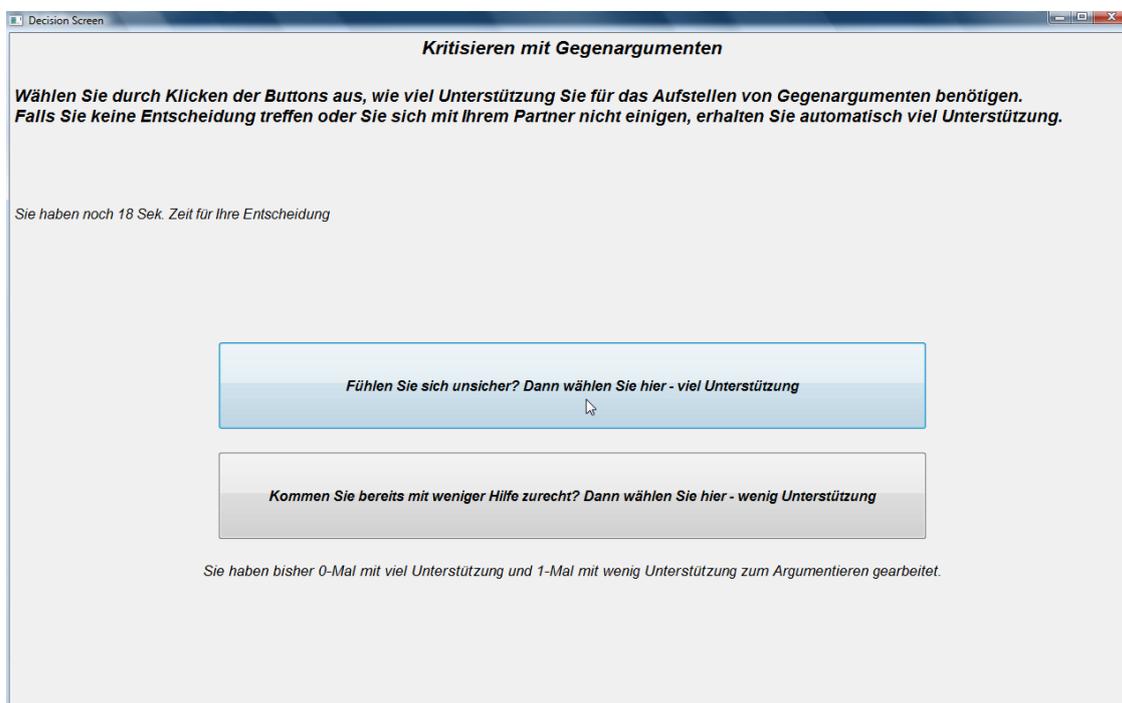


Figure 2. Screenshot of the decision screen with two buttons to adapt the script either into a high-structured or into a low-structured script.

GPA. For the GPA as moderator variable the students were asked for their final high school qualification grade. The grades ranged from 1.00 to 3.50 with an average grade of 1.94.

Self-regulation skills. Self-regulation skills were measured with an 18-items questionnaire (adapted from Fisher, King, & Tague, 2001) in which students rated the extent to which they typically apply certain self-regulation strategies (e.g.; "I prefer to plan my own learning"; "I am systematic in my learning"). The resulting scale proved reliable (Cronbach's $\alpha = .77$).

Instruments and outcome measures

To analyze students' individual learning outcomes, we conducted parallel pre- and post-tests one day before and after the treatment phases, comprising 17 open items each (cf. Kollar et al., 2014). Five items on schematic argumentation with elementary rules from number theory (e.g., "Show that for all natural numbers, a and b , the following statement is true: If 7 divides $a+3b$ then 7 divides $2a+13b$."), which required transformations of the algebraic expression and application of rules from the courses' number theory lectures. Proof skills in elementary number theory were examined by six items (e.g., "Prove the following statement: The sum of five consecutive numbers is divisible by five."), and six items tested performance in open-ended argumentation problems (e.g., "Prove or refute the following statement for natural numbers a and b : If you multiply the sum of a and b with the difference of a and b , you will always obtain an even number."). Two trained, independent raters coded all items. Inter-rater reliability was good (Mean of $ICC_{unjust} = .79$). Where discrepancies remained, raters discussed them until they reached a consensus. Reliability was good for both tests (Cronbach's alpha: $\alpha = .82$ for the pre-test, $\alpha = .80$ for the post-test). For the statistical analyses all scores were merged and scaled to values between zero (nothing correct at all) and one (everything correct).

Results

First, we checked if the three script conditions differed in the pre-test mathematical argumentation skills of the participants. An ANOVA did not yield significant differences ($F(2,93) = 2.10, n.s.$). To answer the first research question, an ANCOVA with pre-test mathematical argumentation skills as a covariate was conducted.

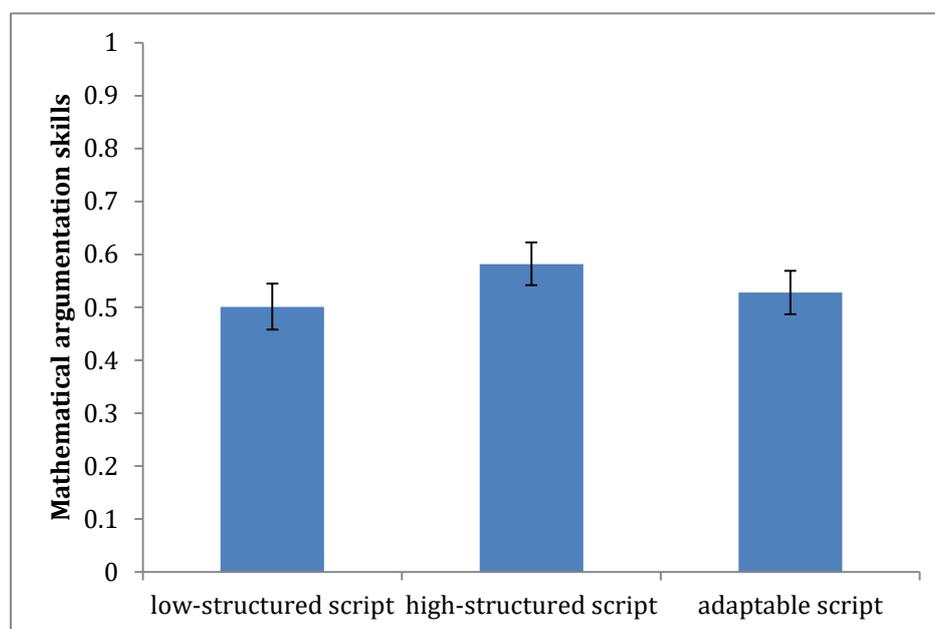


Figure 2. Learners' mathematical argumentation skills by script condition.

The descriptive statistics showed that regarding the outcome score, learners in the high-structured script condition outperformed the learners in the adaptable script condition. Learners in the low-structured script condition reached the lowest outcome score for mathematical argumentation skills. The ANCOVA revealed a significant effect of the script conditions on students' acquisition of mathematical argumentation skills ($F(2,92) = 3.91, p = .02$, part. $\eta^2 = .08$). Post-hoc comparisons between the three levels of the independent variable showed that it was the condition with the high-structured script ($M = 0.58, SE = 0.02$) that reached significantly higher results in mathematical argumentation skills than the condition with the low-structured script ($M = 0.50, SE = 0.02$). The adaptable script condition ($M = 0.53, SE = 0.02$) was not significantly different from the high- and low-structured script (see Figure 2).

To answer research questions RQ2 and RQ3, we conducted regression analyses in each of the three conditions with pre-test mathematical argumentation skills, self-regulation skills and GPA as predictors and post-

test mathematical argumentation skills as dependent variable. Results of the regression analyses showed that only in the low-structured script condition, self-regulation skills were significantly and positively related to the acquisition of mathematical argumentation skills. GPA was no significant predictor at all for post-test mathematical argumentation skills (see Table 1). Also, in all conditions the pre-test mathematical argumentation skills served as significant predictor for the post-test argumentation skills.

Table 1: Regression models for learning mathematical argumentation skills within each of the three conditions.

	<i>B</i>	<i>SE B</i>	β
low-structured script¹			
Pre-test mathematical argumentation skills	0.708	0.116	.738***
Self-regulation skills	0.109	0.051	.264*
GPA	0.049	0.038	.202
high-structured script²			
Pre-test mathematical argumentation skills	0.633	0.109	.742***
Self-regulation skills	0.002	0.056	.005
GPA	0.033	0.041	.100
adaptable script³			
Pre-test mathematical argumentation skills	0.609	0.130	.686***
Self-regulation skills	0.047	0.054	.110
GPA	0.019	0.038	.079

Notes: ¹R² = .662***, ²R² = .591***, ³R² = .597***; *p < .05, **p < .01, ***p < .001

Discussion and conclusions

The results on mathematical argumentation skills are in line with our hypothesis reflecting the results of studies on collaboration scripts for argumentation within other domains (e.g. Weinberger et al., 2010). Thus, learning with a high structured collaboration script helped students extend their domain-specific skills in mathematical argumentation. Since it was the high-structured script that led to better skill acquisition compared to the low-structured script, the findings of this study are not favouring a strategy of minimal scripting when trying to avoid possible overscripting (Dillenbourg, 2002). Furthermore, the difference between the high structured and the low structured script lies in the additional prompting for the use of Toulmin's (1958) argument structure. The use of the structure itself might have helped student to elaborate more deeply on the mathematical content and through this they may have acquired better mathematical argumentation skills (King, 2007). It is plausible that increasingly making the script more complex would eventually lead to an overscripting (Dillenbourg, 2002). The more important question than *how much* script support is necessary, might be *which script support* is actually helpful. Thus, a systematic variation of different script components based on different frameworks (Toulmin, 1958, Leita, 2002) might be a promising research direction for the future.

As Vogel et al. (2014) point out, collaboration scripts rarely lead to domain-specific learning gain, if they are not combined with adequate domain-specific instructional support. In a past study by Kollar et al. (2014), applying a collaboration script did not support the acquisition of mathematical argumentation skills substantially better than unstructured collaboration, when combined with problem solving or heuristic worked examples. The guided problem solving support using driving questions (Nadolski et al., 2006) based on an experts' model of the proof process (Boero, 1999), as applied in this study, represents a compromise between the two approaches used by Kollar et al. (2014). The results indicate that this approach might offer better opportunities for students to actually make use of the collaboration script support.

The support with an adaptable collaboration script for argumentation was not significantly different from leaning with a high-structured or a low-structured script regarding the acquisition of mathematical argumentation skills. Maybe the learners were overwhelmed with the task to adapt the collaboration script to their own needs. Further studies should investigate which kind of support would be necessary for students for being able to adapt the collaboration script in a fruitful way. Also, computer support might be used to adapt the script automatically to the responses students give (e.g. Diziol, Walker, Rummel, & Koedinger, 2010).

Furthermore, learners in the condition with the low-structured collaboration script reached higher scores in the final mathematical argumentation skills test when they had good self-regulation skills. This implies that especially in the low-structured script condition, the learners with better self-regulation skills might have been able to adjust their argumentation on a higher level than it was supported by the script and thus engage in more elaborated learning dialogue, finally achieving better learning outcomes (King, 2007).

The same effect could not be found in the other two conditions. This is especially surprising for the adaptable script condition. In this condition the learners were explicitly asked to adapt the learning environment to their own needs. We expected that an adequate adaption of the learning environment would require a certain amount of self-regulation skills. Yet, there was no difference for students in the adaptable script condition with higher or lower self-regulation skill in their acquisition of mathematical argumentation skills. One reason for this could be that all learners, no matter if they were good self-regulators or not, always adapted the script into one direction without reflecting if this script might fit to their needs. Again, this is a further argument to find better instructions to help learners within the process of adapting their learning environments.

What can be seen as success for the design of the collaboration scripts is the fact that in none of the three conditions the GPA was an important factor for learning mathematical argumentation skills. The collaboration script seems to “balance” to some extent the differences learners might have which might be expressed by the GPA the learners have.

Of course, this study is not without limitations. What can be seen as drawback of this study is that it has no real control condition in the sense of a waiting group or a group learning without any support. Yet, this study was deliberately designed to compare different kinds of collaboration scripts which should lead to better insights about how collaboration scripts should be designed in order to support domain-specific learning.

In conclusion, more structure seems to be better than less when we aim at designing collaboration scripts to positively affect domain-specific learning outcomes. Mainly this might be true because less structure is accompanied by higher demands on self-regulation skills which are quite unevenly distributed among secondary school students as well as freshmen at university.

References

- Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. In J. Andriessen, M. Baker, & D. Suthers (Hrsg.), *Arguing to learn: Confronting cognitions in Computer-Supported Collaborative Learning environments* (pp. 1-25). Dordrecht: Kluwer.
- Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. Preuve: International Newsletter on the Teaching and Learning of Mathematical Proof, (July/August 1999). Retrieved from <http://www.lettredelapreuve.it/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html>
- Cohen, G. C. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1-35.
- De Wever, B., Schellens, T., Van Keer, H., & Valcke, M. (2008). Structuring asynchronous discussion groups by introducing roles: Do students act in line with assigned roles? *Small Group Research*, 39(6), 770-794. doi:10.1177/1046496408323227
- Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL* (pp. 61–91). Heerlen: Oben University Nederland.
- Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R., (2010). Using intelligent tutor technology to implement adaptive support for student collaboration. *Educational Psychology Review*, 22(1), 89-102. doi:10.1007/s10648-009-9116-9
- Douek, N. (1999). Some Remarks about Argumentation and Mathematical Proof and their Educational Implications. In Proceedings of the First Conference of the European Society of Research in Mathematics Education (Vol. I, pp. 125–139). Osnabrück: Forschungsinstitut für Mathematikdidaktik.
- Fisher, M., King, J., & Tague, G. (2001). Development of a self-directed learning readiness scale for nursing education. *Nurse Education Today*, 21(7), 516–525. doi:10.1054/nedt.2001.0589
- Heinze, A., Reiss, K., & Rudolph, F. (2005). Mathematics achievement and interest from a differential perspective. *Zentralblatt für Didaktik der Mathematik*, 37(3), 212-220. doi:10.1007/s11858-005-0011-7
- Hilbert, T. S., Renkl, A., Kessler, S., & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction*, 18(1), 54–65. doi:10.1016/j.learninstruc.2006.10.008
- King, A. (2007). Scripting collaborative learning processes: A cognitive perspective. In F. Fischer, I. Kollar, H. Mandl & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives* (pp. 13-37). New York: Springer.
- Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts - a conceptual analysis. *Educational Psychology Review*, 18(2), 159-185. doi: 10.1007/s10648-006-9007-2

- Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning and Instruction, 17*(6), 708-721. doi: 10.1016/j.learninstruc.2007.09.021
- Kollar, I., Ufer, S., Reichersdorfer, E., Vogel, F., Fischer, F., & Reiss, K. (2014). Effects of collaboration scripts and heuristic worked examples on the acquisition of mathematical argumentation skills of teacher students with different levels of prior achievement. *Learning and Instruction, 32*, 22–36. doi: 10.1016/j.learninstruc.2014.01.003
- Leitão, S. (2000). The potential of argument in knowledge building. *Human Development, 43*, 332–360.
- Nadolski, R. J., Kirschner, P. A. & van Merriënboer, J. J. G. (2006). Process support in learning tasks for acquiring complex cognitive skills in the domain of law. *Learning and Instruction, 16*(3), 266-278.
- O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz, & N. Miller (Eds.). *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 120-141). New York: Cambridge University Press.
- Pease, A., Smaill, A., Colton, S., & Lee, J. (2009). Bridging the gap between argumentation theory and the philosophy of mathematics. *Foundation of Science, 14* (1-2), 111-135. doi:10.1007/s10699-008-9150-y
- Reiss, K. M., Heinze, A., Renkl, A., & Groß, C. (2008). Reasoning and proof in geometry: effects of a learning environment based on heuristic worked-out examples. *ZDM The International Journal on Mathematics Education, 40*(3), 455–467. doi:10.1007/s11858-008-0105-0
- Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14*(2), 201-241.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum
- Thurston, W. P. (1994). On Proof and Progress in Mathematics. *Bulletin of the American Mathematical Society, 30*(2), 161–177. doi:10.1090/S0273-0979-1994-00502-6
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
- Van Eemeren, F. H. & Grootendorst, R. (2004). *A Systematic Theory of Argumentation: The pragma-dialectical approach*. Cambridge: Cambridge University Press.
- Vogel, F., Kollar, I., Ufer, S., Reichersdorfer, E., Reiss, K., & Fischer, F., (2015). Fostering argumentation skills in mathematics with adaptable collaboration scripts: Only viable for good self-regulators? In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the material conditions of learning. The Computer-Supported Collaborative Learning Conference (CSCL) 2015 – Volume II* (pp. 576-580). International Society of the Learning Sciences: University of Gothenburg.
- Vogel, F., Kollar, I., Wecker, C., & Fischer, F. (2014). The role of content support and transactivity for effects of computer-supported collaboration scripts on domain-specific learning: A meta-analysis. In: F. Xhafa, L. Barolli, F. Palmieri, M. Koeppen, & V. Loia. (Eds.) *Proceedings of the 2014 International Conference on Intelligent Networking and Collaborative Systems* (pp. 677-682). Los Alamitos, CA: IEEE Computer Society. doi:10.1109/INCoS.2014.82
- Weinberger, A., Fischer, F., & Mandl, H. (2003). Collaborative knowledge construction in computer-mediated communication: Effects of cooperation scripts on acquisition of application-oriented knowledge. *Zeitschrift für Psychologie, 211*(2), 86-97.
- Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior, 26*, 506-515. doi: 10.1016/j.chb.2009.08.007

Coordinating Collaborative Chat in Massive Open Online Courses

Gaurav Singh Tomar, Sreecharan Sankaranarayanan, Xu Wang, and Carolyn P. Rosé
gtomar@cs.cmu.edu, sreechas@cs.cmu.edu, xuwang@cs.cmu.edu, cprose@cs.cmu.edu
Carnegie Mellon University

Abstract: An earlier study of a collaborative chat intervention in a Massive Open Online Course (MOOC) identified negative effects on attrition stemming from a requirement for students to be matched with exactly one partner prior to beginning the activity. That study raised questions about how to orchestrate a collaborative chat intervention in a MOOC context in order to provide the benefit of synchronous social engagement without the coordination difficulties. In this paper we present a careful analysis of an intervention designed to overcome coordination difficulties by welcoming students into the chat on a rolling basis as they arrive rather than requiring them to be matched with a partner before beginning. The results suggest the most positive impact when experiencing a chat with exactly one partner rather than more or less. A qualitative analysis of the chat data reveals differential experiences between these configurations that suggests a potential explanation for the effect and raises questions for future research.

Introduction

The field of Computer Supported Collaborative Learning (CSCL) has a rich history extending for nearly two decades, covering a broad spectrum of research related to learning in groups, especially in computer mediated environments. In contrast, a major limitation of the current generation of Massive Open Online Courses (MOOCs) is a lack of social presence. Analyses of attrition and learning in MOOCs both point to the importance of social engagement for motivational support and overcoming difficulties with material and course procedures. Effective collaborative learning experiences are known to provide many benefits to learners in terms of cognitive, metacognitive, and social impact (Kirschner, Paas, & Kirschner, 2009; Webb & Palinscar, 1996). These experiences offer a potentially valuable resource for MOOCs, if affordances can be provided that facilitate high quality collaborative learning interactions in the absence of human facilitators that can keep up with the high enrolment in such courses.

Traditionally, MOOC platforms, including team based MOOC platforms like NovoEd, have not offered synchronous interaction opportunities or even instantaneous forms of social awareness. MOOC contexts allow for asynchronous interaction and possibly even collaboration. However, the asynchronous nature leads to some unsavory experiences. For example, sometimes participants spend time posting a thoughtful post but get only a cursory reply. Participants sometimes have to wait days or even weeks to get a response to a question or receive feedback on completed work. This lack of immediacy implies that information can be out of date by the time someone views it, or worse, that the student needing help or feedback gave up and dropped out before the response was posted. These types of issues can hinder motivation. These limitations at the interface level stem from limitations at the architecture level due to challenges in scaling immediate update protocols. Through integration protocols such as the Learning Tools Interoperability protocol (LTI) chat tools and other forms of synchronous interaction have found their way into a few MOOCs. However, just the ability to integrate a synchronous chat room into a MOOC does not solve the problem. Challenges remain regarding coordinating the times of the discussions as well as supporting the functioning of an ongoing discussion.

The proposed solution to this problem is to support synchronous interaction within the MOOC context. Students may prefer synchronous activities because it offers them a greater experience of active involvement and social connection. Coordinating synchronous social engagement in a MOOC is challenging, however, and the challenges that arise can be frustrating for students. An earlier study identified positive and negative effects on attrition of a collaborative chat intervention that required students to be matched with exactly one partner prior to beginning the activity (Ferschke et al., 2015b). Negative effects occurred when students had to return multiple times in an attempt to be matched with a partner. That study raised questions about how to provide the benefit of synchronous social engagement without the coordination difficulties. In this paper we present a careful analysis of an intervention designed to overcome coordination difficulties by welcoming students into the chat on a rolling basis as they arrive rather than requiring them to be matched with a partner before beginning. This design raises questions about how the impact of the experience differs depending on conditions within the less controlled social environment within the chat.

In the remainder of the paper, we describe the motivation for the study from the Computer Supported Collaborative Learning (CSCL) literature. We then describe the intervention we evaluate in this paper. Next we present both quantitative and qualitative analyses of our intervention. Finally, we conclude with discussion and future directions.

Theoretical foundation

The fact that conversational interaction may provide an opportunity for (collaborative) learning has been underscored by many learning theories ranging from cognitive to socio-cultural perspectives (Kirschner, Paas, & Kirschner, 2009; Webb & Palinscar, 1996). However, it is well known that without support, many instances of collaborative learning fail. In light of this fact, the field of collaborative learning has produced a wide variety of forms of scaffolding often referred to as scripts. Collaboration scripts may operate at the macro level, providing task structuring and role assignment. Or it may operate at the micro level, structuring the nature of the flow of contributions to the discourse.

Our goal is to find the best ways to provide learners with a space to interactively hone their understanding of concepts related to a specific domain so that they have the chance to display their own reasoning, experience how others display their reasoning, challenge and be challenged by others. In order for the interactions to provide a meaningful learning experience for students, it was essential for these activities to be well integrated into the instructional design, and not treated as an afterthought or an appendage. This leads to more authentic, and ecologically valid learning experiences for students (Tudor, 2000; Van Lier, 2004). In addition to this embedding, interactions are mainly meaningful for learning when they are structured and scaffolded in an appropriate manner.

In this paper, we build on a paradigm for dynamic support for group learning that has proven effective for improving interaction and learning in a series of online group learning studies conducted in classroom settings. In particular we refer to using tutorial dialogue agent technology to provide interactive support within a synchronous collaborative chat environment (Kumar et al., 2007; Chaudhuri et al., 2008; Chaudhuri et al., 2009; Kumar et al., 2010; Ai et al., 2010; Kumar & Rosé, 2011). Introduction of such technology in a classroom setting has consistently led to significant improvements in student learning (Adamson et al., 2014), and even positive impacts on the classroom environment outside of the collaborative activities (Clarke et al., 2013). While it would seem to be desirable to import such technology into a MOOC setting to provide a learning experience that is both more instructionally valuable and socially supportive, such an introduction comes with its technical, methodological, and theoretical challenges.

In our setup, the goal is for the agent to aid the students in integrating their respective understandings of the concepts they previously encountered individually in the video lectures and other assignments that precede the chat activity. On the other hand, it also provides an opportunity to develop their collaborative skills (O'Donnell, 1999).

Collaboration platform

The chat tool system deployed in the MOOC study reported in this paper is Bazaar (Adamson et al., 2014), which is used to provide interactive support within online synchronous collaborative chat activities. Reflection activities were authored by the instructor team to provide students the opportunity to reflect substantively about the material they learned in each unit. Each chat activity included several specific reflection topics to provide a consistent macro level structure to the chats for each unit. Very little micro-level support was offered in addition to the macro-level structuring.

The most similar previously published study regarding a collaborative chat intervention in a MOOC was published by Ferschke and colleagues (Ferschke et al., 2015a). In that study, in order to facilitate the formation of ad-hoc study groups for the chat activity, Ferschke and colleagues (Ferschke et al., 2015b) made use of a simple setup referred to as a Lobby. Students entered the Lobby with a simple, clearly labeled button integrated with the edX platform. Upon entering the Lobby, students were asked to enter a username that would be displayed in the chat. Once registered in the Lobby, the student waited to be matched with another participant. If the student was successfully matched with another learner who arrived at the Lobby within a couple of minutes to interact with, he and his partner were then presented with a link to click on to enter a chat room created for them in real time. Otherwise they were requested to come back later. Some students needed to make up to 15 attempts in order to be successfully matched for a chat. Thus, many students were frustrated. A follow up analysis (Ferschke et al., 2015b) reports a negative impact on commitment to the course for students who experienced this frustration. An important lesson learned from this study was that whereas providing the opportunity for synchronous chat was positive for students for whom it was possible to be matched for a chat easily, this positive effect was balanced with a negative effect in the case where the lack of critical mass despite the total enrollment of 20,000 students from their MOOC was not sufficient to enable a quick match.

In order to address the difficulties reported in the earlier study (Ferschke et al., 2015b), we did not employ an intermediate Lobby interface in our design. Instead we exposed a single, continuous chat room that multiple students could join at any time on a rolling basis, rather than beginning all at the same time, and without having to be matched with a particular peer student. The conversation continues as long as there is at least one person in the room. The architecture is able to facilitate a chat with only one student present and keeps the conversation going for any number of participants. That way, students can enter the chat on their own schedule and join other students who are engaged in the activity. The chat room automatically resets once all participants have left the chat.

In order to provide a very light form of micro-level script-based support for the collaboration, the computer agent facilitator in our design kept track of the students in the room and prompts that had been given to the students. For each student, it tracked the events of their joining and leaving the chat room. For topic prompts, it tracked the times they were given and the students who were present in the chat room at that time. This tracking was done to make sure each student could engage with all question prompts of the reflection exercise and no student saw the same question prompt twice unless the activity script restarted and students decided to repeat it.

Another design consideration was to address a difficulty identified in the earlier study, which was that the students did not always stay on topic and sometimes engaged in goalless interactions or neglected to interact at all. One possible explanation is that these students may have misunderstood the question. To address this issue in the current design, the agent kept track of any student and group inactivity that went on for more than two minutes. It also kept track of how many utterances each student contributed. It aggregated students' messages and computed a similarity score between student messages and the topic prompt to test whether the students were on topic. This was also done over a two-minute window. If the group was dormant or did not say anything that is relevant to the topic in that two-minute window then the agent generated a poke message, which rephrased the question again by providing timely hints that make it more explicit what is expected from a contribution. A contentful group discussion on a question prompt was observed to last for about 10 minutes on average.

Although a single, continuous chat room provides the capacity to solve the problems with synchronous collaboration in a MOOC observed in the earlier study, it may also entail additional coordination challenges. For instance, a student may not be motivated to join in the middle of the discussion or might feel lost due to a lack of a frame of reference for where to start upon joining an in progress discussion. To address this issue, we introduced a prompt for the agent to ask the group to summarize the state of the conversation for incoming students in order to help ramp up new participants and encourage current participants to voice their understanding of discussed concepts. Alternatively, sometimes an agent-generated summary is given, mentions a list of topics already discussed as well as a summary of the current topic. To ensure the agent does not interrupt the discussion with repeated summarization requests, these requests were issued only if at least two topic prompts had been discussed.

Results

The most positive result of the earlier Ferschke et al. study (Ferschke et al. 2015b) was the observed reduction in attrition associated with participation in a collaborative chat as measured through a survival analysis. For our work, it is important to determine not only whether we have accomplished the specific changes to student experiences that motivated our design work. We also want to ensure that we maintain or even enhance the positive impact observed with the earlier approach. In order to measure the impact of participation in a collaborative chat on attrition, we adopt a survival analysis approach. Survival analysis is a statistical modeling technique used to model the effect of one or more indicator variables at a time point on the probability of an event occurring on the next time point. In our case, we are modeling the effect of participation in a collaborative chat on probability that a student ceases to participate actively in the course on the next time point.

Methodology

Survival models are a form of proportional odds logistic regression, and they are known to provide less biased estimates than simpler techniques (e.g., standard least squares linear regression) that do not take into account the potentially truncated nature of time-to-event data (e.g., users who had not yet ceased their participation at the time of the analysis but might at some point subsequently). In a survival model, a prediction about the probability of an event occurring is made at each time point based on the presence of some set of predictors. The estimated weights on the predictors are referred to as hazard ratios. The hazard ratio of a predictor indicates how the relative likelihood of the failure (in our case, student dropout) occurring increases or decreases with an increase or decrease in the associated predictor in the case of a continuous variable, or presence vs. absence of the factor in the case of a binary variable. A hazard ratio of 1 means the factor has no effect.

If the hazard ratio is a fraction, then the factor decreases the probability of the event. For example, if the hazard ratio was a number of value .4, it would mean that for every standard deviation greater than the average of

the continuous predictor variable (or where a data point has value 1 for a binary predictor variable), the event is 60% less likely to occur (i.e., $1 - n$). If the hazard ratio is instead greater than 1, that would mean that the factor has a positive effect on the probability of the event. In particular, if the hazard ratio is 1.25, then for every standard deviation greater than the average of the continuous predictor variable (or where the data point has value 1 for a binary predictor variable), the event is 25% more likely to occur (i.e., $n - 1$).

Survival analyses are correlational analyses, and as such they do not provide causal evidence for an effect. However, lack of an effect in a survival analysis would suggest that the data fail to provide causal evidence as well. A positive effect in a survival analysis would suggest that it makes sense as a next step to manipulate the associated factor so that causal evidence for a positive effect could be measured.

Specifying the model

In our survival model we include control variables, independent variables, and a dependent variable. Our primary interest is how the independent variables related to participation in collaborative chats make predictions about the dependent variable, which indicates course dropout. However, control variables are essential in accounting for variances in the participants that may influence attrition. For example, some students are more active in the course in general, indicating a priori greater commitment to the course. If we do not account for this in our model, we cannot say whether it's the intervention we are introducing that is leading to the observed difference in the dependent variable, or it's because other confounding variables such as the difference priori commitment level of students' drives their participation in the intervention that leads to the effect.

Unit of analysis. In order to assess the impact of measured factors at each time point during a student's trajectory through the course, it is necessary to decide what the unit of analysis is. In other words, it is necessary to determine what the time interval to use in the survival model is. Even the most active participants in the course did not participate every day. However, very active participants returned to the course more than once within a week. Based on preliminary exploration into the dataset, we chose our unit of analysis to be a 7-day period of time.

Student population. We conducted our analysis in an edX MOOC called Big Data in Education (BDEMOOC), which was taught by Ryan Baker from Teacher's College, Columbia University. The MOOC was launched in the fall of 2015.

We included in our analysis all students who had at least clicked once during the course to enter a chat. There were 401 such students in our dataset. The collaborative chat activities were positioned as enrichment activities after the individual work for the week was completed. In order to enter a chat, students clicked on a button in the courseware page. We expect that students who attempted to participate in a chat were more active on average than students who did not since the chat activities were positioned at the end of the unit, so they would be mainly consumed by students who had completed the other assignments of the course. However, students who did not complete the other activities were not prevented from participating. The baseline data points in our model are the ones when students did not click to enter a chat during the corresponding time interval.

Control variables. An important indicator of a priori commitment to active engagement was the number of clicks on course videos. We included a standardized count of video clicks during the time period between two consecutive data points as a control variable in our analysis.

Independent variables. As mentioned, the baseline for our comparison is data points that did not include any attempt to enter into a chat. There were 21,968 of these. If a student clicked to enter a chat, four possible things could happen. 1) Some number of students experienced technical difficulties due to their Internet configuration, which didn't allow them to connect to the chat server port from their network. There were 145 such data points, which we label as Malfunction data points. 2) Another possibility is that the student entered the chat, but no other student joined them there. There were 349 such data points. We label these as alone data points. 3) There were 145 data points where students entered a chat and there was exactly one partner student. We labeled these as Pair data points. 4) Finally there were 15 data points where there were more than two students present. We refer to these as Group data points. We included binary independent variables in our survival model for Malfunction, Alone, Pair, and Group.

In order to assign these binary variables, it was necessary to compute the maximum number of peers that the student interacted with in a particular chat session. We calculated this number by counting the number of peers present each time the student performs an event (join, leave, post a message). From the set computed during a chat session we take the maximum. This number reflects the largest group size the student can possibly have actively interacted with during the session.

Dependent variable. We referred to the dependent variable as Drop. This was a binary variable that was 1 for the last time interval of a student's active participation in the MOOC and 0 otherwise.

Survival model results

The main results of the analysis (summarized in Table 1) are consistent with those of the earlier study (Ferschke et al., 2015b) in that we see a consistent trend suggestive of positive impact when students participated in a chat. The most positive impact is in the case where students chat with exactly one partner student. There the hazard ratio is .6 indicating a 40% increase in probability that the student will still be active at the next time point. However, the effect is only marginal ($p = .06$). The trend for data points where students chat alone or with more than one partner student is also positive, however the effect is not significant. There is a significant negative impact of experiencing a malfunction due to network connectivity issues. However, even this is less strong than that reported by (Ferschke et al., 2015b) for students who were required to click multiple times in order to be matched for a chat. Ferschke et al. (2015b) report a hazard ratio of 2.33 for a standardized count of match attempts. Overall, the results suggest that welcoming students into the chat on a rolling basis rather than requiring them to be matched with a partner student in order to enter is a more advantageous strategy since students who experience a pair chat benefit, and those who chat alone or with more than one student are not harmed.

Table 1. Survival table with estimates that measure the impact of control variable (Video Clicks), and independent variables (Malfunction, Alone, Pair, and Group) on probability of survival

Independent Variable	Hazard Ratio	p-value
Standardized Video Clicks	.97	$p = \text{n.s.}$
Malfunction	1.7	$p < .001$
Alone	.89	$p = \text{n.s.}$
Pair	.6	$p = .06$
Group	.8	$p = \text{n.s.}$

Taking a closer look: Post hoc analyses of results

The results of the survival model suggest that students who participated in the BDEMOOC and chatted with exactly one partner student experienced more benefit. In our Post Hoc analysis we attempted to better understand what was different in their experience. We find evidence that students experienced a richer interaction when they participated in the chat with a single partner student. One strong indicator of this richness is in comparing amount of time spent in the chat based on how many students were participating. We see a significant difference in the average time spent in the chat between students who had a partner in the chat and student who didn't ($F(2,563) = 3.1, p < .05$). Chats with more than two students were not significantly longer or shorter in the average time of the session than either. Informally, we also see less frustration and more exchanged reasoning in the chats containing exactly two students.

Table 2. Comparison of average time of chat sessions in relation to the number of students participating.

Number of students in the chat room	Mean time spent	Standard deviation	Median time spent
1	389.44	755.67	0.0
2	501.58	874.23	1920.0
3	582.86	586.31	1560.0

In order to further investigate what is the difference between dialogues in chat rooms that contain only one person, two people, and more than two people, we manually coded frustration in the dialogues and compared the difference between different conditions. Since frustration was a factor that influenced dropout in the Ferschke et al. (2015b) study, we believed this would offer a useful lens on the chat processes at work in our study.

Two coders coded 100 dialogue turns generated by learners in the collaborative chats as expressing frustration or not in order to check the inter rater reliability of frustration coding. The coders achieved a kappa of 0.784. The two coders then continued to code the rest of the dataset, which contains 2917 dialogue turns in total. The coders then computed the average number of contributions coded as expressing frustration in each condition. The trend in proportion of contributions expressing frustration across sets was consistent with expectations. The proportion of contributions expressing frustration was lowest in the chats with exactly two participants. The rooms with single student generate the highest concentration of messages expressing frustration (5.78%), and the rooms with more than two students rank second (3.59%), while the rooms with pairs generate the lowest concentration of messages expressing frustration (2.86%). In testing the significance of the difference between the three conditions, we found that only the difference in the proportion of frustration messages expressed between the condition of single students and that of pairs is marginally significant (Z -score = 1.7662, p = 0.077).

We then took a closer look at what happened in the chat room in different conditions. When there is only one person in the chat room, the student tends to check whether others are present, and sometimes expresses boredom and disappointment at being alone. Here are examples from several different chats with individual students:

Student 1: shrugs shoulders

Student 2: Is it just me and VirtualRyan?

Student 3: Am I missing anything? I feel silly just chatting into space! :P

Another student was excited in the beginning, and he talked to himself for a while, nobody responded to him, and then he left.

Student 4: Is anyone here?

Anyone wants to discuss learning curves and other visualization methods?

This was the best week so far. The discussion about the various curves gave insight into how these techniques can be use to understand the impact of curriculum or how to help students directly.

I'll try again another time.

leave

Above we reported that students in pairs remained in the chat room longer than individual students. Informally we observed a higher concentration of expressed reasoning in the chats with pairs. Here are some examples of conversation when exactly two people are in the room. The examples below show that when there are two people in the chat room, they frequently exchanged perspectives, and offered positive comments on each other's ideas. We underlined the parts that they are exchanging, referring to, building upon and commenting on each other's ideas.

Example 1:

P: The dataset I looked at included more than time-related variables, so I would think yes. There are grades received at different stages, reason why a student took a class, etc.

D: Hmm, I agree with the baseline idea that the decision tree works better at categorical data.

D: Good. Thanks for the inspiration.

Example 2:

T: Well, to continue our theme, I would use data mining to predict MOOC dropout in China

E: That sounds interesting. What type of predictors were you thinking you may use?

T: I am going to use the "assignment accomplishment" predictor.

T: The reason I chose the predictor is based on my own experience. So how about yours?

E: I am working to develop new online programs and have access to all of the data in our LMS so thought I may be able to help tell students what they should be doing to increase their chance of success.

- E:* It may also help us check use of the LMS and flag students who may have persistence issues....just a thought. I don't have experience in data mining before.
- T:* And I would like to add a second feature: some complaints in the discussion area, to predict the dropout, as well.
- E:* Good thoughts! What all is included in the assignment accomplishment predictor?
- T:* Thanks. but just a thought, not very detail. I believe the more assignment accomplished, the more successful the subject may reach the endline.

Discussion

In this paper, we presented an analysis of results from a collaborative chat intervention that suggests positive impact of the intervention when students experienced the chat with exactly one partner. A post hoc analysis suggests that students had a richer interaction when they were in the chat with exactly one partner. One possible explanation is that the support offered to groups was not sufficient to scaffold the interaction for more than two students. However, the number of chats with more than two students was small. And our analysis is merely correlational. Thus, we are unable to make strong claims about the reason for the pattern we found with this study.

One limitation of the current study is that some students experienced a negative impact from technical difficulties, namely the port connection issues discussed earlier. Though we carefully isolated the data points directly impacted by this issue in our analysis, we cannot eliminate the possibility that the negative impact of this issue on the students who directly experienced it may have affected others through hearing about the issues in the discussion forums. This issue has now been resolved in our infrastructure.

More substantively, despite the positive impact of collaborative chats when the conditions were ideal (e.g., where exactly one partner student was present), we identified some challenges with the intervention. For example, we observed that many students who were chatting with agent in the absence of peers felt that agent was not a good listener, felt bored, and sometimes left room in the middle of the discussion. They expressed a desire to be able to discuss the questions with the agent instead of just giving their reflections on the question prompt being asked. Another issue was that some dynamics of the agent were tuned to the situation where multiple human students were present and thus were not appropriate in the case where only one student was present. Thus, a more dynamic timing strategy for the task structuring may have been received better.

In our follow up work, in order to alleviate the boredom felt by individual students in the chat room, we are developing a design that supports a more substantive exchange between a student and the agent in the absence of other human partners in the interaction. Specifically, we will make use of tutorial dialogue agent technology, such as Knowledge Construction Dialogues (KCDs) (Rosé & VanLehn, 2005; Kumar & Rosé, 2011), which may provide more intensive engagement. Thus, if a single student is present in the room, the facilitation agent may lead the student step by step to construct substantive explanations related to the issues targeted by the collaborative chat exercise. Adaptive time of task structuring will also be employed.

Conclusions

This paper addresses questions on how to effectively orchestrate synchronous discussion activities in MOOCs despite practical coordination challenges. It presents an evaluation of a specific collaborative chat intervention integrated within an edX MOOC. While an earlier study offered results associating participation in a chat session with reductions in attrition over time (Ferschke et al., 2015b), that study was not able to separate the effect of the chat activity from the effect of synchronous interaction with a peer. In this paper, we present a correlational analysis suggesting differential effects of a chat experience depending upon the number of peers present. The positive impact associated with interaction with a single partner student is consistent with the results presented in the earlier study, though slightly weaker, and thus provide more confidence in the finding that interaction with a peer is associated with at least temporary elevation in commitment to the course. The study also suggests further directions to improve the intervention for future MOOCs.

References

- Adamson, D., Dyke, G., Jang, H., & Rosé, C. P. (2014). Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education*, 24(1), 92-124.
- Ai, H., Kumar, R., Nguyen, D., Nagasunder, A., & Rosé, C. P. (2010, January). Exploring the effectiveness of social capabilities and goal alignment in computer supported collaborative learning. In *Intelligent Tutoring Systems* (pp. 134-143). Springer Berlin Heidelberg.

- Chaudhuri, S., Kumar, R., Joshi, M., Terrell, E., Higgs, F., Alevan, V., & Rosé, C. P. (2008, January). It's not easy being green: Supporting collaborative "green design" learning. In *Intelligent Tutoring Systems* (pp. 807-809). Springer Berlin Heidelberg.
- Chaudhuri, S., Kumar, R., Howley, I. K., & Rosé, C. P. (2009, July). Engaging Collaborative Learners with Helping Agents. In *AIED* (pp. 365-372).
- Clarke, S., Chen, G., Stainton, K., Katz, S., Greeno, J., Resnick, L., Howley, H., Adamson, D., Rosé, C.P.: The impact of CSCL beyond the online environment. In: *Proceedings of Computer Supported Collaborative Learning* (2013)
- Ferschke, O., Howley, I., Tomar, G., Yang, D., & Rosé, C. P. (2015a). Fostering Discussion across Communication Media in Massive Open Online Courses. In *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning. Gothenburgh, Sweden*.
- Ferschke, O., Yang, D., Tomar, G., & Rosé, C. P. (2015b, June). Positive Impact of Collaborative Chat Participation in an edX MOOC. In *Artificial Intelligence in Education* (pp. 115-124). Springer International Publishing.
- Greybeard, A. P. & Wizwoman, B. Z. (2009). *Designing our way to new learning: Two decades of innovation and improvisation in the classroom*. New York, NY: Scholarly Publishers, Inc.
- Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, 21(1), 31-42.
- Kumar, R., Rosé, C. P., Wang, Y. C., Joshi, M., & Robinson, A. (2007). Tutorial dialogue as adaptive collaborative learning support. *Frontiers in artificial intelligence and applications*, 158, 383.
- Kumar, R., Ai, H., Beuth, J. L., & Rosé, C. P. (2010, January). Socially capable conversational tutors can be effective in collaborative learning situations. In *Intelligent tutoring systems* (pp. 156-164). Springer Berlin Heidelberg.
- Kumar, R., & Rose, C. P. (2011). Architecture for building conversational agents that support collaborative learning. *Learning Technologies, IEEE Transactions on*, 4(1), 21-34.
- O'Donnell, A. M. (1999). Structuring dyadic interaction through scripted cooperation. *Cognitive perspectives on peer learning*, 179-196.
- Webb, N. M., & Palincsar, A. S. (1996). *Group processes in the classroom*. Prentice Hall International.
- Tudor, I. (2001). *The dynamics of the language classroom*. Ernst Klett Sprachen.
- Van Lier, Leo. *The ecology and semiotics of language learning: A sociocultural perspective*. Vol. 3. Springer Science & Business Media, 2004.

Acknowledgement

This work was funded in part by grants from Google and the Gates foundation and NSF grant ACI-1443068.

Community Knowledge, Collective Responsibility: The Emergence of Rotating Leadership in Three Knowledge Building Communities

Leanne Ma, OISE/University of Toronto, leanne.ma@mail.utoronto.ca
Yoshiaki Matsuzawa, Aoyama Gakuin University, matsuzawa@si.aoyama.ac.jp
Bodong Chen, University of Minnesota, chenbd@umn.edu
Marlene Scardamalia, OISE/University of Toronto, marlene.scardamalia@utoronto.ca

Abstract: Developing cultural capacity for innovation is an educational imperative. The challenge in schools is to foster a culture of sustained, creative work with ideas, as in out-of-school Collaborative Innovation Networks (COINs) and cyberteam that self-organize to create knowledge. In this study, we examined the online knowledge work of three Knowledge Building classes, where young students assumed collective responsibility for creating and improving their community knowledge. We adopted the COIN concept of rotating leadership to visualize collective responsibility for knowledge advancement. Using a mixed methods approach, we conducted social and temporal network analyses, then content analyses of student notes to further assess cases of student leadership. Overall, we found relatively decentralized student networks, with most students leading the group at different points in time; when leading, students were connecting unique ideas to the larger class discussion. We discuss our findings within the context of designing embedded, transformative assessment for knowledge building communities.

Keywords: Knowledge Building; Collaborative Innovation Network; collective responsibility; rotating leadership; self-organization; knowledge creation; innovation

Introduction

To prepare students for a world of constant change (Takeuchi & Nonaka, 1986) within innovation-driven societies (OECD, 2010), many 21st century education movements (e.g., Johnson, 2009; Ananiadou & Claro, 2009) aim to equip students with skills in creativity, critical thinking, collaboration, and communication. Apart from changing curricular goals, however, most classroom activities continue to be teacher-directed, with students being assessed individually against traditional benchmarks (Griffin, McGraw, & Care, 2012). Such fixed structures and processes prevent innovation, adaptability, and new competencies from emerging in schools (Sawyer, 2015; Scardamalia et al., 2012). One way to address educational challenges of the 21st century is by exploring innovative organizational configurations for knowledge creation in schools (Bereiter, 2002; Sawyer, 2006; Philip, 2011).

Knowledge Building, which is synonymous with knowledge creation (Scardamalia & Bereiter, 2014; Bereiter & Scardamalia, 2014), represents a longstanding effort at building cultural capacity for innovation. It aims to transform education into a knowledge-creating enterprise by creating a culture of innovation in classrooms and engaging students directly in sustained, creative work with ideas, so that students are engaged continuously in creating knowledge of value to the community (Scardamalia & Bereiter, 2003). From a young age, students are empowered to take ownership of their learning through participation in high-level decision processes, so that all share collective responsibility for advancing the community goals (Scardamalia, 2002).

In contrast to many activity-centered and procedure-based constructivist pedagogies, Knowledge Building uses an idea-centered, principle-based design approach (see Scardamalia, 2002 for overview of 12 principles). For example, the principles of *idea diversity*, *improvable ideas*, and *rise above* prioritize students' ideas at the center of class interactions and highlight the iterative nature of idea generation, refinement, and invention in knowledge creation processes that enhance the breadth and depth of group understanding and achievement. Principles of *epistemic agency*, *democratizing knowledge*, and *community knowledge, collective responsibility* create contexts for empowering students to take charge of knowledge creation processes at the highest level, including defining problems, setting goals, contributing to and monitoring goal progress, synthesizing advances at new and unexpected levels, and more generally, giving priority to knowledge creating interactions as cultural practices (Bereiter & Scardamalia, 2014). In a successful Knowledge Building classroom, the teacher and students embody all twelve principles as they work collaboratively toward creating community knowledge. The teacher supports student agency and autonomy by facilitating improvisation in their own practice through opportunistic collaborative engagement (Zhang et al., 2009), wherein the teacher encourages students to self-organize into small groups based on emergent goals of the community. Similarly, Collaborative Innovation Networks (COINs; Gloor, 2006), which create knowledge and drive innovations around the world, function based

on principles of autonomy and self-organization. Whereas highly productive teams have stable leaders, COINs operate in a decentralized fashion with various emergent leaders rotating leadership over the course of a project (Gloor et. al., 2003; Kidane & Gloor, 2007). Members share collective responsibility for their knowledge work through a high degree of connectivity, interactivity, and sharing (Gloor, 2006). The idea-centered, principle-based design approach to Knowledge Building pedagogy emphasizes the role of self-organization in knowledge creation and innovation (Hong & Sullivan, 2009; Scardamalia & Bereiter, 2014), which is consistent with how COINs operate.

Because the Knowledge Building process is emergent and non-linear, assessment designs must complement the 12 Knowledge Building principles in a way that supports and sustains collective progress, as indicated by the principle of *concurrent, embedded, and transformative assessment*. Previous work aimed at developing principle-based indicators of collective responsibility for community knowledge advancement used descriptive statistics (e.g., van Aalst et. al., 2012) to assess online reading and writing behaviours and social network analyses (e.g., Philip, 2010) to assess collaboration patterns within the community. Additionally, semantic measures (e.g., Hong et. al., 2015) and lexical measures (e.g., Sun, Zhang, & Scardamalia, 2010) were developed to assess the diffusion of ideas within the community and the growth of community knowledge over time. While quantitative measures are useful for capturing patterns of social interactions and community connectedness, qualitative analyses are useful to assess the quality and coherence of ideas, concepts, and theories shared between students. Thus, in designing Knowledge Building assessments, it is important to integrate social network analyses with content analyses in order to understand the complex process of collaborative meaning making (van Aalst, 2012). We further add that the ideal assessment would integrate social, semantic, and temporal aspects of Knowledge Building.

The current study is exploratory in nature, with the goal of developing a new method to assess collective responsibility for knowledge advancement in three Knowledge Building classes. We hypothesized that if young students are really taking on collective responsibility and self-organizing to create community knowledge, we would find rotating leadership as an emergent phenomenon of these classes, as in COINs. In adopting the stance that a Knowledge Building class is a complex, self-organizing system with multiple components interacting at multiple levels, we created two sets of research questions which corresponded to two levels of analyses:

1. At the group level, what does rotating leadership look like in the Knowledge Building class? How many students emerge as leaders over the course of the inquiry?
2. At the individual level, what is happening when a student is leading? How are they contributing to the group discourse?

Methods

In recognition of the complex and dynamic nature of the Knowledge Building process, we adopted a complementary mixed methods design (Greene, Caracelli, & Graham, 1989) in order to harness the strengths of quantitative and qualitative analyses and develop a more holistic understanding of the phenomenon under study: collective responsibility for knowledge advancement. We used quantitative methods (i.e., social and temporal network analyses of online interactions) to address the first sets of questions, and we used qualitative methods (i.e., content analyses of online discourse) to address the second sets of question.

Data sources

Our samples consisted of Knowledge Building classes at the Dr. Eric Jackman Institute of Child Study participating in Beyond Best Practice and Ways of Contributing research projects (2002-2012) located in Toronto, Canada. In order to test our hypothesis, we selected successful Knowledge Building classes and conducted secondary analysis to validate a new method of assessment. We selected classes in which 1) students engaged in sustained inquiry for a minimum of three months, 2) students documented their knowledge advances through extensive writing online, and 3) Knowledge Forum support was integrated into daily classroom practices. Knowledge Forum (Scardamalia, 2004) is an online community space optimized to support knowledge creation processes. Students contribute ideas as notes in conceptual spaces called views; connect ideas with “build-on” notes; and generate explanations/syntheses with “rise-above” notes. Student work in Knowledge Forum involves continuous reading, writing, and revising of notes to advance the community knowledge. In summary, Case 1: Grade 4 light consisted of 22 students engaged in inquiry about light for three months. Students wrote 380 notes across 8 views: Light, How Light Travels, Colors of Light, Light and Materials, Natural and Artificial Light, Shadows, Images in our Eyes and in Films, and All We See Is Light. Case 2: Grade 1 water consisted of 21 students engaged in inquiry about water for three months. Students wrote 391 notes across 3 views: All about Clouds, Where did water come from?, and Evaporation. Case 3: Grade 4 rocks consisted of 23 students engaged

in inquiry about rocks for four months. Students wrote 269 notes across 3 views: Rocks and Minerals, Volcano/Lava, and The Big Bang and the Universe. See Table 1 for summary of three cases.

Data analyses

The first stage of analysis involved quantitative methods. Student notes in Knowledge Forum were spellchecked and exported into KBDeX (Knowledge Building Discourse Explorer; Oshima, Oshima, & Matsuzawa, 2012) in order to facilitate social and temporal network analyses based on a list of content-related words extracted from the Ontario Curriculum of Science and Technology, which served as benchmark concepts in the community knowledge. KBDeX produces network visualizations of the learners network, notes network, and words network based on the co-occurrence of key words in each network, with the thickness of edges representing the strength of connections. For example, the learners network shows idea sharing across learners, the notes network shows idea overlap across notes, and the word network shows how ideas are connected in the community knowledge as the group discussion unfolds. KBDeX also produces temporal visualizations of network metrics, such as betweenness centrality, which indicates the extent to which a member influences other members of the group (Gloor et. al., 2003). Whereas a betweenness centrality value of 1 means that a member is highly influential, a value of 0 means that a member is equally influential as other members. In using KBDeX, we were able to explore the three networks seamlessly in order to examine stages of the inquiry when a specific learner had high betweenness centrality, as indicated by the emergent peaks in the temporal visualization (see Figure 1). For example, when we selected a specific learner in the learners network, their corresponding notes were highlighted in red in the discourse network, and their corresponding keywords were highlighted in red in the word network (see Figures 2a, 2b, 2c respectively). The second stage of analysis involved qualitative methods. After identifying leaders with a betweenness centrality value of 0.20 or higher, we used KBDeX to explore the three networks in order to understand the discursive context that supported their emergence as the leader: 1) which students shared the same ideas as this specific leader, 2) which notes contained the same ideas as this specific leader's notes, and 3) which ideas this specific leader connected together. More specifically, the note network allowed us to conduct content analyses on notes immediately connected to that specific leader; we interpreted notes isolated from this network as student efforts directed toward diverse and unique ideas in the community knowledge. By tracing and comparing changes in the network shapes across different times, we were able to identify pivotal points in the group discourse that led to greater connectedness of ideas (i.e., explanatory coherence), which we interpreted as knowledge advancement. Below, we present one leader (i.e., influential student) from each Knowledge Building case.

Findings

Table 1 shows descriptive measures to assess Knowledge Building in all three cases. Network measures for reading and writing behaviours, such as average weighted degree (AWD), transitivity, average path length (APL), and average betweenness centrality (ABC), indicate that students were working productively in Knowledge Forum for each case. Overall, reading networks are denser than build-on networks, suggesting that students read more of their peers' notes than they did build on them. In the section below we present our findings for each Knowledge Building case from general to specific: rotating leadership at the group level, followed by a detailed account of what happened when a particular student was leading. For each case, we present our results in the following order: temporal analyses, network analyses, content analyses.

Table 1. Descriptive measures of student activities in Knowledge Forum for Cases 1, 2, and 3

Sample	Students	Notes	Network	Density	AWD	Transitivity	APL	ABC
Case 1: Grade 4 light	22	380	Reading Building-on	1.03 0.17	43.36 7.27	1.00 0.42	1.02 2.24	20.02 23.57
Case 2: Grade 1 water	21	391	Reading Building-on	0.76 0.26	31.82 10.82	0.91 0.57	1.29 2.16	19.04 28.25
Case 3: Grade 4 rocks	23	269	Reading Building-on	0.96 0.25	40.18 10.54	0.99 0.55	1.05 2.21	22.90 26.49

Case 1: Grade 4 light

Figure 2 shows temporal analysis of betweenness centrality for Case 1. The Y axis of the chart shows the betweenness centrality value, and the X axis shows the turn in discussion over time. Each coloured line represents a student, resulting in the display of 22 lines in the chart. The oscillation of coloured lines depicts the phenomenon

of rotating leadership, which means that the leading student (i.e., the student with the highest betweenness centrality) changed frequently. Of the 22 students, 20 students took a leading position, suggesting that many students were influential at different times. The legend in the top right of Figure 2 indicates the colour of student 11, who was leading between turns 166 to 178, and peaked at turn 175 (betweenness centrality = 0.20).

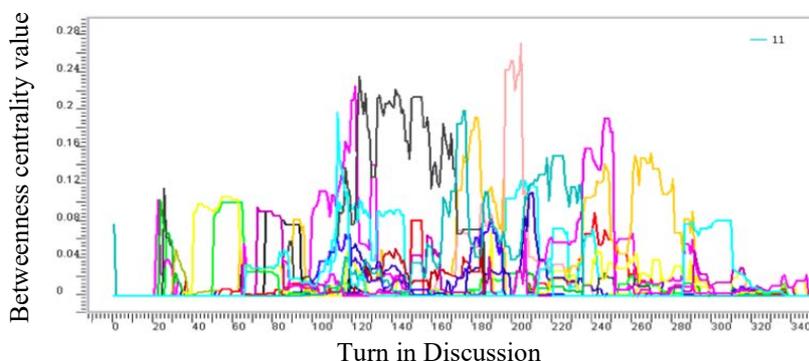


Figure 1. Temporal visualization of betweenness centrality of Case 1: Grade 4 light.

Figure 2 shows network analyses in KBDeX when student 11 was leading at turn 175. The student network in Figure 2a shows that student 11 connected students 24, 8, and 14 to the larger group network. The note network in Figure 2b shows that note 176 written by student 11, linked notes 179, 157, and 139 to the larger cluster of notes. The word network in Figure 2c shows that student 11 connected the concepts of “sunlight”, “solar energy”, and “flashlight” to the main discussion of lenses and magnifying glasses.

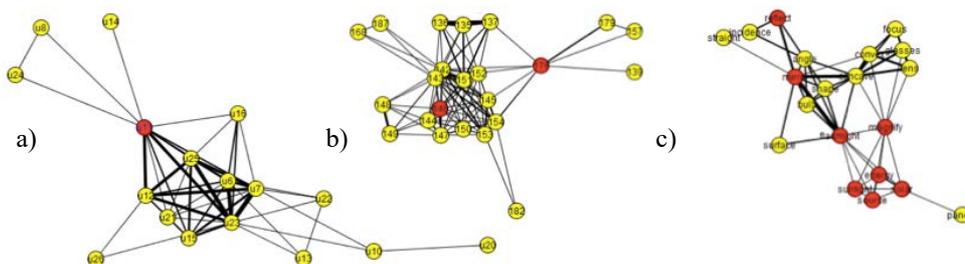


Figure 2. Network visualizations at turn 175: a) student network, b) note network, and c) word network.

Below is an excerpt of the discussion where student 11 played an important role in connecting ideas from their peers’ notes in the Natural and Artificial Light view. The problem of understanding, as indicated by student 24, is light as an energy source. Student 24 raised a question about how solar energy works, while student 8 shared their theory about solar panels, and student 14 described their experiment about light mills. Student 11 added their theory about how the source of light would relate to its strength of energy. Student 11 hypothesized that sunlight would be a stronger source than a flashlight and added that a magnify glass could be used to adjust the strength of energy. Student 11’s note connects student 24, 8, and 14’s notes to the larger discussion about lenses and magnifying glasses, where student 16 explains how different types of lenses and glasses adjust the strength of light.

- Student 24: solar power is everywhere or it should be everywhere because it’s just as good as regular energy i don’t know how it works but it is just as good.
- Student 8: [My theory]: is that the light from the power plant is sunlight that comes from solar panels.
- Student 14: [We] did an experiment that was a light mill. we went outside to see if the mill would spin by using sunlight. we don’t know which way the mill turned?
- Student 11:** [My theory]: is that the sunlight can produce stronger solar energy. under the sun. some people do it by using a magnify glass. maybe the glass focuses the light on the ant which burns it. i think if you tried to burn an ant under a flashlight the source of the light would not be strong enough.

Student 16: lens have two sides one is called convex and one is called concave. When you put light on a magnify glass it will make the light get smaller and hotter and it will close in. it also depends on the different types of glasses on the different types of shapes.

Case 2: Grade 1 water

Figure 3 shows temporal analysis of betweenness centrality for Case 2. Of the 21 students, 12 students took a leading position, suggesting that half the students were influential at different times. The legend in the top right of Figure 3 indicates the colour of student 195251, who was leading between turns 147 to 209, and peaked at turn 161 (betweenness centrality = 0.22).

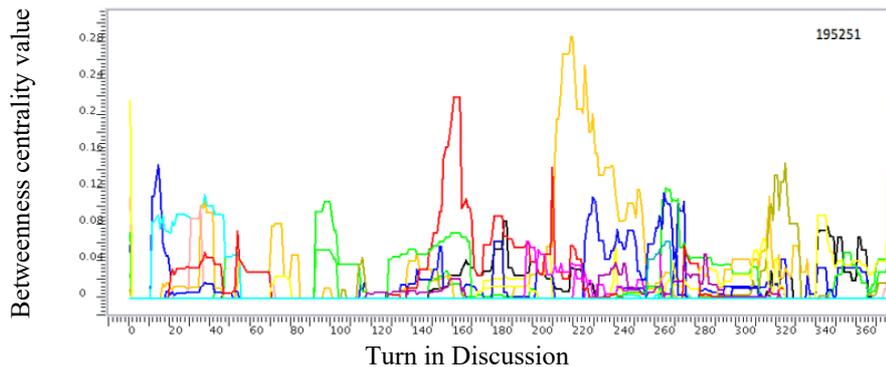


Figure 3. Temporal analysis of betweenness centrality of Case 2: Grade 1 water.

Figure 4 shows network analyses in KBDex when student 195251 was leading at turn 209. The student network in Figure 4a shows that student 195251 connected students 1044599, 1852393, and 1162041 to the larger group network. The note network in Figure 2b shows that note 183 written by student 195251, linked notes 181, 177, and 211 to the larger cluster of notes. The word network in Figure 2c shows that student 195251 connected the concepts of “space”, “atmosphere”, and “air” to the main discussion of cloud and evaporation.

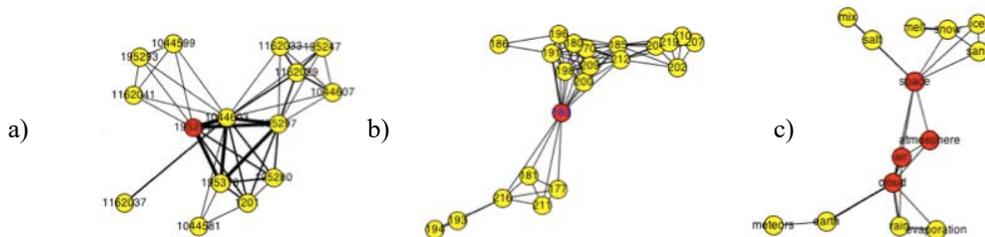


Figure 4. Network visualizations at turn 209: a) student network, b) note network, and c) word network.

Below is an excerpt of the discussion where student 195251 played an important role in connecting ideas from their peers’ notes in the All About Water view. The problem of understanding, as indicated by student 1162041, is whether or not there is water in space. Student 1162041 raised a question about where water goes in space, while student 185293 wondered about water from space mixing with salt on earth and student 1044599 shared their theory about water coming from space. Student 195251 added their theory about how the atmosphere stops clouds (and by extension, water) from entering space. Student 195251 hypothesized that since there is no air in space, there are no clouds (and by extension, water) in space. Student 195251’s note connects student 1162041, 1852393, and 1044599’s notes to the larger discussion about clouds, water, and evaporation, where student 195313 observed that you cannot see water in a cloud.

Student 1162041: [I wonder]: if water is in space where does it go

Student 185293: [I wonder]: if water came from space and mix with salt on the ground?

Student 1044599: [My theory]: it could have came from space

Student 195251: [My theory]: there is no air in space and cloud are air and i think that the atmosphere will stop it so the air can’t go

- Student 379:** [A better theory]: is that it is a place where atoms collided and formed planets. then life form came and it evolved to the life forms you see today. that is how humans came to be.
- Student 361: [I need to understand]: how did life first come to earth? (after the big bang) i thought that the earth was made out of fire until it cooled down by a icy meteorite.

Conclusions and implications

This study aimed to adopt the COIN concept of rotating leadership as a new method for visualizing collective responsibility for knowledge advancement. Our findings demonstrate that rotating leadership, as indicated by oscillating patterns of betweenness centrality, is an emergent phenomenon in Knowledge Building classes, where students work collaboratively to create community knowledge while receiving appropriate pedagogical and technological supports. In Case 1: grade 4 light, 20 of 22 students were leaders. In Case 2: grade 1 water, 12 of 21 students were leaders. In Case 3: grade 4 rocks, 11 of 23 students were leaders. When students were leading, they were generating new theories by connecting their peers' unique ideas with the larger class discussion, thus demonstrating principles of *epistemic agency*, *idea diversity*, *improvable ideas*, *rise above*, and *community knowledge*, *collective responsibility*. It is interesting to note that in all three cases, in addition to having leaders who rotate frequently, multiple leaders appeared to emerge simultaneously. Further analyses are needed to understand leadership dynamics with multiple, "shared" leaders, seemingly engaged in symmetric, mutually supportive advances. COIN theory (Gloor, 2006) suggests that having a strong core with several overlapping leaders is conducive to group creativity and productivity, thus the "sharing" of rotating leadership may also be beneficial for the Knowledge Building community as a whole.

Future work on Knowledge Building assessments should aim to verify our findings across various learning contexts (e.g., learner's age, subject matter, school culture). KBDex represents a powerful analytic tool for providing formative feedback during Knowledge Building (e.g., Matsuzawa et al., 2012, Resendes et al., 2015). We believe that rotating leadership has the potential to serve as a descriptive measure for teachers and students to assess and monitor group progress as their creative work proceeds, and ultimately help sustain innovative interactions in the 21st century classroom.

References

- Ananiadou, K. and Claro, M. (2009), "21st Century Skills and Competences for New Millennium Learners in OECD Countries", OECD Education Working Papers, No. 41, OECD Publishing.
- Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah: Lawrence Erlbaum Associates.
- Bereiter, C. & Scardamalia, M. (2014). Knowledge building and knowledge creation: Once concept, two hills to climb. In Tan, S. C. & Tan, Y. H. (2014). Perspectives of knowledge creation and implications for education.
- Gloor, P. A. (2006). *Swarm creativity: Competitive advantage through collaborative innovation networks*. Oxford: University Press.
- Gloor, P. A., Laubacher, R., Dynes, S. B., & Zhao, Y. (2003, November). Visualization of communication patterns in collaborative innovation networks-analysis of some w3c working groups. In *Proceedings of the twelfth international conference on Information and knowledge management* (pp. 56-60). ACM.
- Griffin, P., McGaw, B., & Care, E. (2012). *Assessment and teaching of 21st century skills* (p. 36). Dordrecht: Springer.
- Hong, H. Y., & Sullivan, F. R. (2009). Towards an idea-centered, principle-based design approach to support learning as knowledge creation. *Educational Technology Research and Development*, 57(5), 613-627.
- Hong, H. Y., Scardamalia, M., Messina, R., & Teo, C. L. (2015). Fostering sustained idea improvement with principle-based knowledge building analytic tools. *Computers & Education*, 89, 91-102.
- Johnson, P. (2009). The 21st Century Skills Movement. *Educational Leadership*, 67(1), 11.
- Kidane, Y. H., & Gloor, P. A. (2007). Correlating temporal communication patterns of the Eclipse open source community with performance and creativity. *Computational and mathematical organization theory*, 13(1), 17-27.
- Matsuzawa, Y., Oshima, J., Oshima, R., & Sakai, S. (2012). Learners' use of SNA-based discourse analysis as a self-assessment tool for collaboration. *International Journal of Organisational Design and Engineering*, 2(4), 362-379.

- Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: A social network analysis application for knowledge building discourse. *Educational Technology Research and Development*, 60(5), 903-921.
- Philip, D. N. (2010). Social network analysis to examine interaction patterns in knowledge building communities. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie*, 36(1).
- Philip, D. N. (2011). Knowledge building: Reinventing education for the Knowledge Age. *International Education Studies*, 4(4), p118.
- Resendes, M., Scardamalia, M., Bereiter, C., Chen, B., & Halewood, C. (2015). Group-level formative feedback and metadiscourse. *International Journal of Computer-Supported Collaborative Learning*, 10(3), 309-336.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith, & C. Bereiter (Eds.), *Liberal education in a knowledge society* (pp. 67-98). Berkeley, CA: Distributed by Publishers Group West.
- Scardamalia, M. (2004). CSILE/Knowledge Forum®. In *Education and technology: An encyclopedia*. pp. 183-192. Santa Barbara: ABC-CLIO.
- Scardamalia, M., & Bereiter, C. (2003). Beyond brainstorming: Sustained creative work with ideas. *Education Canada-Toronto*, 43(4), 4-8.
- Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 397-417). New York: Cambridge University Press.
- Scardamalia, M., Bransford, J., Kozma, B., & Quellmalz, E. (2012). New assessments and environments for knowledge building. In *Assessment and teaching of 21st century skills* (pp. 231-300). Springer Netherlands.
- Sawyer, K. (2006). Educating for innovation. *Thinking skills and creativity*, 1(1), 41-48.
- Sawyer, K. (2015). Organizational innovation and improvisational processes. In Garud, R., Simpson, B., Langley, A., & Tsoukas, H. (Eds.). *The Emergence of Novelty in Organizations* (pp. 180-215). Oxford University Press.
- Takeuchi, H., & Nonaka, I. (1986). The new new product development game. *Harvard business review*, 64(1), 137-146.
- van Aalst, J. (2012). Commentary: Combining content analysis and social network analysis. *Open and Interdisciplinary Journal of Technology, Culture and Education*, 7(2), 84-94.
- van Aalst, Chan, C. K. K., ... & Wan, W. S. (2012). The knowledge connections analyser. In J. van Aalst et al. (Eds.), *Proc. 10th Int. Conf. Learning Sciences* (pp. 361-365). Sydney, Australia: ISLS
- Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge-building communities. *The Journal of the Learning Sciences*, 18(1), 7-44.

Stories as Prototypes for Interdisciplinary Learning

Vanessa Svihla, University of New Mexico, vsvihla@unm.edu
Richard Reeve, Queen's University, reever@queensu.ca

Abstract: Although recent research has emphasized the importance of understanding science practices and how people teach and learn them, much of this is set within mono-disciplinary classrooms, providing limited vision for how people might learn from one another in the context of an interdisciplinary lab. We present data from participant observation collected over a three-year period in an interdisciplinary scientific research laboratory as a lens into how people can learn science from one another. We depict how the participants used stories as a means to prototype their ideas over time, and how their approach to treating these stories as tentative and contingent supported them to make progress and learn together.

Major issues addressed and significance of the work

Stories are appealing because they provide narrative coherence that supports comprehension and recall (Thorndyke, 1977). They are an important part of social learning (Brown, Collins, & Duguid, 1989). Here, we explore how stories serve as prototypes to support interdisciplinary learning. If we consider the normal state of affairs in how students engage in building explanations of their understanding—particularly in traditional school science, but even in many scaffolded learning technologies, their purpose is relatively straightforward; it involves rendering data into the canonical solution. Students are seldom presented with ill-structured problems (Jonassen, 1997), and thus, are confined to basic deductive reasoning in much of their learning. Yet, even a straightforward, well-structured problem can be treated as ill-structured if it is given contingent, tentative status. Treating well-structured problems as ill-structured may actually reflect learners' experience of such problems, provided they are engaging in sense-making as they are framing the problem.

Theoretical background

Bereiter and Scardamalia (2012) argue that deep understanding “is achieved through creating and improving explanatory theories” (p. 160). They further suggest that an explanatory theory about an aspect of nature, that is a scientific theory related to a natural phenomenon, is best built from data observed across multiple cases. However, where we find existing fields of research, as well as entirely new fields of research dealing with previously unexamined new cases, we find researchers striving to make sense of the data through theory building. Bereiter and Scardamalia (2012) have referred to this as being a “theory of the case” (p. 162) and have further suggested this is the type of theorizing that is common in social studies, literature and history. However, the idea that “a good theory of a case should be able to coherently explain all of the facts of a particular case” (p. 162) is in line with data we have from scientific settings where new areas of study are being undertaken, but for which overriding theories do not yet exist. More precisely, we have found that in these research settings the efforts to pursue explanatory coherence often takes the form of creating a “story about the data.”

The key to *authentic* theorizing is that it involves “continued striving toward higher degrees of explanatory coherence” (Thagard, 2000). We consider this a form of prototyping in which each prototype is tested against the contextual and contingent knowns and forms, and in/for/of a given moment. Such prototypes are tentative and liminal. In design, prototypes serve important functions; they aid designers “to either explore a design space or narrow down options and make decisions” (Hess, 2012). Sometimes, simpler prototypes lead to better design (Yang, 2005). Prototypes serve myriad purposes, with some functioning as a simplified final design and others standing in as form alone (Ulrich & Eppinger, 1995). Prototypes serve to aid designers in learning about their design ideas, communicating their design ideas, and integrating their design ideas with other existing ideas (Ulrich & Eppinger, 1995).

We argue interdisciplinary settings can serve as important learning environments, especially for prototyping explanatory stories. Interdisciplinary learning settings may be particularly beneficial to struggling students, and may invite broader participation (Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering, & Institute of Medicine, 2004; MacMath, Roberts, Wallace, & Chi, 2010). Interdisciplinary work tends to be collaborative, with each member bringing different expertise, and even fundamentally different understandings of the problem (Weingart, Todorova, & Cronin, 2008); each member, also, therefore, brings a *lack* of expertise in some aspects of the problem. This creates an environment where it can be safe to *not* know something. We see this as a particularly strong setting for treating prototypes and explanations as tentative and contingent, and for this to lead to learning.

Methodological approach

We present a case that highlights how interdisciplinary research presented opportunities for students to be positioned as having some relevant expertise, and how finding a story in the data served as a means for them to prototype their ideas and ultimately learn and make progress on their work.

Setting and participants

The geomicrobiology lab is lead by Denise (pseudonym) who was a part of the formation of her subfield. Members of the lab study the role of microbes in caves. Denise's lab is diverse and includes undergraduate and graduate students (N=14 across three years, with approximately seven students participating at any one time). During the three-year data collection period, the lab typically included one or two Native American students, one African-American student, four to six Latino/a students, and a majority of students who were first generation college attendees. The gender balance was generally close to even. Denise recruits students who are "in the middle" academically, and who might be in danger of leaving or not considering science careers.

Data sources

Data were collected at weekly hour-long lab meetings across three years. The first year was documented through field notes, resulting in 64 pages of typed notes. The second and third years included field notes and audio recording, resulting in 145 pages of typed field notes and 28 hours of audio. These were supplemented with ethnographic interviews with participants. The credibility of findings was enhanced through commonly-used strategies (Creswell & Miller, 2000; Golafshani, 2003; LeCompte & Goetz, 1982). Triangulation and external audits were conducted in the first author's research lab. Member checking occurred by presenting findings at lab meetings, informal interviews with lab members and sharing paper drafts. As part of member checking, informal interviews were recorded, and participants followed up by sending presentations and journal articles annotated to explain their progress on finding the story in the data.

Analysis

Field notes were reviewed repeatedly for regularities in lab practices. Where needed, audio records of the lab meetings were transcribed to provide a more detailed record of the interactions. We present vignettes highlighting how stories served as prototypes and how this practice fostered participation. Segments shared reflect the larger data corpus. Audio records of selected data were transcribed. Transcripts were reviewed by the study team for moments that seemed most relevant to the study. These were more carefully transcribed to document pauses and tone of voice, to better convey the conversations. The resulting transcripts were analyzed using interaction analysis (Jordan & Henderson, 1995). In particular, we attended to how participants discussed story, and how they made sense of each other's contributions.

Findings

We present a narrative depicting the evolution of one instance of finding the story in the data. This story begins, in terms of data documenting it, in Fall 2012. We noticed that Denise, the PI, would explicitly lead the lab in pursuit of "the story in the data."

Denise: We have a problem. We need to find the story in Kacy's data. Tell them what your questions are.

Kacy: If bacteria in pools have the ability to precipitate, if it's not strictly a geologic process.

Denise: Crystals don't form pendant structures. But do they form these particular structures?

Elena: Do they touch the water?

Denise: They form UNDER water.

Because this geomicrobiology lab is interdisciplinary, there is a commitment to developing shared understanding about data and problems, which in turn supports collective work to reach shared understanding of the data. About a month later, Denise informed the lab they'd be doing more "finding the story in the data." She explained, "I want us to develop as a lab group—finding the story in the data. A colleague sent me this manuscript, and I got lost in the data. I'd like to do several lab meetings on this." Indeed, they did engage in this practice over the entire Fall semester lab meetings. At the end of the semester, Denise asked the students to reflect on the semester, to help them plan out the Spring lab goals.

Denise: What have you found most valuable about lab meetings?

Jeff: Analyzing data and having people comment.

Denise: So, finding the story in the data?

Jeff: I was grouchy, but I got a lot out of it.

Despite Jeff's "grouchy" stance when the lab collectively sought the story in his data, he acknowledged its value; with two lab meetings devoted to scrutinizing his data, he initially reacted defensively, feeling as if he ought to have already known and tried everything they suggested. By the second session, he realized it was not intended to showcase his lack of progress, but rather to support him to *make* progress and to support everyone to learn. This practice—finding the story in the data—has persisted in the lab ever since.

About two years into data collection, Denise had gotten data back from analysis and she was trying to put a talk together. She brought the data—as a presentation with many charts and graphs—to the lab. As she shared the data, she explained what she had noticed. We have annotated the transcripts to help the reader navigate technical terms.

Denise: There are a few real oddball samples. [...] I have bacteria "unassigned." That is just—they could not even get it down to a phylum, right?

Aaron: Yeah. They know it's a bacterial sequence, but nothing else.

Denise: So, the fact that this one has THAT much, that's a fair amount of possibly new orders, new phyla maybe even, right? Ok, so, that's one part of the story. So, what you're looking for as we go through this, is I want some help coming up with the story, because it's kinda slim pickings. Ok, so, what I have observed so far is there are some, there are some unusual and interesting candidate phyla (*newly identified bacteria*) [...]

Aaron: What I might do here is get rid of all the sample names and then do like a, like an FMD (*ferromanganese deposit- an iron and manganese-rich, colorful deposit found on cave walls. The lab studies these to understand the role of microbes in caves*) and bracket it, so like, all your FMDs are here. You could say 'FMDs' and then do it by sample type.

Denise: Ok, I can do that.

Aaron: 'cause that'll show at least a story of "it's chaotic even within a sample type." [...] You might actually be able to see some trends by sample type.

Denise: Ok. I did that in the report that I did with a smaller subset. [...] I can actually organize by color within the FMD (*because FMDs come in different colors*). I can do that and this—THIS is really interesting and really puzzling.

In this interaction, Aaron suggested a means to group some of the samples together. When Aaron suggested bracketing all the FMDs, he was suggesting grouping data together by a common attribute. Looking at groups that share a common attribute together can sometimes help us find patterns that are masked if the pattern is not shared by the whole group. Aaron noted this as he commented, "You might actually be able to see some trends by sample type." As Denise agreed this was a good approach, she noted how "really interesting and really puzzling" the data were. As the conversation continued, Denise began to envision how to group her data together.

Denise: So, just make this a list?

Aaron: Yeah, and say "Hey, we got some cool stuff. Come check this out." Right, that's actually, that's actually super. I didn't, uh, realize that about the data set.

Denise: No, in fact once it's been through QIIME (*software the lab uses to match the DNA they have to known sequences; thus, here they are puzzling over how few matches were found down to the genus level.*) the whole thing, the whole character of the data set changed. [...]

Tania: Or maybe what you could do is make this really small, like put it in the corner and then just have the main points that you want.

Dr. S: What if you just showed the percent of them that are at class, the percent at family, the percent...(trailing off)

Denise: Is there a way to do that? Convert those to percents?
 Aaron: Well, you could just add up, so how many, how many do we have total? It's like 20% or something like that, right? Then, what, how many are at the genus level? Like 40?
 Tania: Yeah. That wouldn't be hard.
 Denise: Oh, not how many individual sequences are in //
 Jeff: //Oh no.
 Aaron: Yeah, 'cause Dr. S is right. That'll convey the information better. You're saying only, out of all of our samples there's 20% at the genus level.
 August: Yeah, you could do like a pie chart, right?

We note that, in this part of the transcript, there was much overlap between finding and telling the story. As soon as Denise began prototyping her story, she began to consider how to communicate it through charts and figures. As a participant observer, Dr. S was able to contribute; she failed to come up with another taxonomic rank, but her idea still made sense to the group. As the conversation continued, Denise mentioned finding bacteria in the cave that were unexpected, because they are normally found only on land. As they discussed this, they came up with plausible ways for such bacteria to enter a cave (e.g., brought in by a packrat). Tania noted this with some sense of relief in her voice at finding a plausible story in the data.

Tania: That can be explained!
 Denise: Maybe. (.) But this is still, it takes us off in a totally different direction. So, I'm thinking of getting rid of that. So that's what we have. I have copies of all the graphs. I wondered if we wanted to divide up into two teams, and have people look at that. Or have you seen enough here, that you can think about what else could I make as a story. (.) So any thoughts about what would be the best way to proceed?
 Aaron: I'd kinda like to see that in color by age.
 Denise: We can't do that in the next five minutes. So what could we do now to help move the story along? ((laughter))
 Tania: Wait, what was your question?
 Denise: What could we do to help move the story along, because//
 Tania: //Oh ((laughs))
 Denise: Would it be worth—I have two sets of charts.
 Tania: Let's do it guys.

Although the explanation for the unexpected bacteria in the cave provided a potential story, it was dismissed as “taking off in a totally different direction.” It was a bit of a small story, because it could be so easily explained. Thus, they left the relative certainty of that story, and returned to puzzling over the data. We note a sense of playfulness and going boldly into the data. This was no simple task—making sense of myriad graphs and charts. They divided into two groups and spent several minutes orienting to the task, trying to make sense of what they were seeing. As they began to try to make sense of the data, Tania voiced her confusion.

Tania: This one says distance, distance, distance. But then the dots look different. I don't know what that means.
 Aaron: That's a whole lot of nothing. YEP. They all say nothing.
 Tania: mmhmm. ((laughs))
 Aaron: Yep. Still says nothing.
 James: The story is there's no story.
 Aaron: No! There is a story.
 Tania: That is the story.
 James: THAT is the story? ((Laughing))
 Tania: Now we know what?
 Aaron: Each sample was kind of it's own little niche [...]

Tania: We're supposed to look at these, and, um, find a story?
Denise: What other story can we come up with?
Tania: You can play dot the dots.
Aaron: I kinda like the story. It's chaotic.
Tania: I do. I—I like it.

Again, we note a playful quality to their search for story in the data; we see this as opening a space for members to play both with their talk and with ideas. The data representations were a bit unfamiliar to the students, and no clear patterns were visible. Yet the story began to take shape. They began to prototype a story of randomness, handing it around from person to person to consider. We also point out *how* the students in this lab work with Denise. This lab seems to be a safe space for students to speak up, to admit uncertainty. We attribute some of this to the interdisciplinary nature of the lab. In such spaces, everyone experiences being more novice at times, and everyone has a point of view that is valuable. Our purpose is not to delve too deeply into how this was fostered, but to call it out as a tremendous resource that can be leveraged when setting about to find the story in the data. As they continued, Aaron defended the chaotic story that seemed to be emerging.

Aaron: I mean, you can make the argument, depending on, like, how the rest of it comes out, it's like a stochastic process (*random process*) is driving this thing.
Denise: In other words, it's completely random what's on these walls?
Aaron: Yeah.
Denise: Because, I can tell them, you know the temp and humidity are really, pretty
Tania: So, are you trying to figure out why?
Denise: I'm trying to figure out what controls this diversity (*of the bacteria, which are highly diverse*).

What we see here is that Denise, prior to fully endorsing this story, evaluated it against what she knows. Thus, with this prototype figuratively in hand, she could evaluate its fit and coherence. In contrast to how she responded to the previous story of the bacteria that was potentially brought into the cave by a packrat, this story seems to have been more compelling. While the story they were creating may be one of randomness, their search for story was far from random; rather, it was both exploratory and explanatory. They moved between analysis and inference, representing data and relationships between variables, looking for meaningful patterns, then holding these found patterns—or in this case, found randomness—against what they knew and what they expected to find. In this case, the story of randomness is unsatisfying because of this, and warranted going back to analysis, in search of other patterns, other stories.

The story, as prototyped, at a particular moment in time

We shared our initial analysis of how the lab prototyped stories, and Denise feared “that makes us look like occasionally we're just sort of bumbling idiots. [...] We need to have a happy ending to this bumbling around.” She thus shared with us how the story had developed as she prepared her presentation, and also how it had been further revised. When she presented this story at her conference, Denise showed how many of their samples could be identified at each taxonomic level (Figure 1). In her conclusions to that presentation, she reported on the value of using “next generation sequencing to show extensive, novel diversity [...] even in physically close sites in the cave.” Based on the groups' continued work on the data—a back-and-forth process of analysis and story-creation, they further made discoveries about how cave bacteria metabolize methane. She also presented several *stories in the making*—prototypes of stories tied to data that they were still trying to make sense of.

Reflecting on this story a year later, Denise explained numerous ways their story had changed.

Denise: We are a lot further along on that data because we now know, for instance, like, the black and the brown FMDs are where we find more of the new candidate phyla (*newly identified bacteria*). [...] So we know the candidate phyla are there. We also know the FMDs group together not necessarily by color. We know that depth is not important in [that cave] anyway, and we know that: um, there are processes—the metabolic processes, especially in the brown and black FMDs are, um, methanotrophic—so eating of methane compounds. They are nitrogen cycling and

maybe a little bit of sulfur and iron cycling, right? [...] Color is important in that red crusts are very different.

Tania: Yeah.

Denise: But you—if you go back, they're also very//

Tania: //‘cause red doesn’t always mean there’s iron there.

Denise: No, and also//

Dr. S: //That’s interesting ‘cause actually later in this transcript you talk about the reason the red would be different is ‘cause there’s iron.

Denise: Well, we found out with Tania—so—totally different system—[another] cave—she did an analysis—a bulk chemistry analysis and there’s no iron in the [red] samples. [...] You can’t make an assumption, is what it told us—that red means iron is present.

Thus, we see that the story was prototyped for a particular time—a conference presentation—and then further revised and reconfigured. As they have prototyped their ideas, their stories have cross-pollinated one another. Tania’s explanation about the color red not meaning that iron is present—an insight from another study—informed Denise’s understanding of these data. This highlights the tentative nature of scientific research, showing how each story is just a prototype—one that fits the data as currently understood, using the techniques available. We see this practice as eminently transferrable to other learning settings, provided participants treat their stories as prototypes—improvable and contingent. To further understand how the lab conducts this type of work, we consider their reflections on the process.

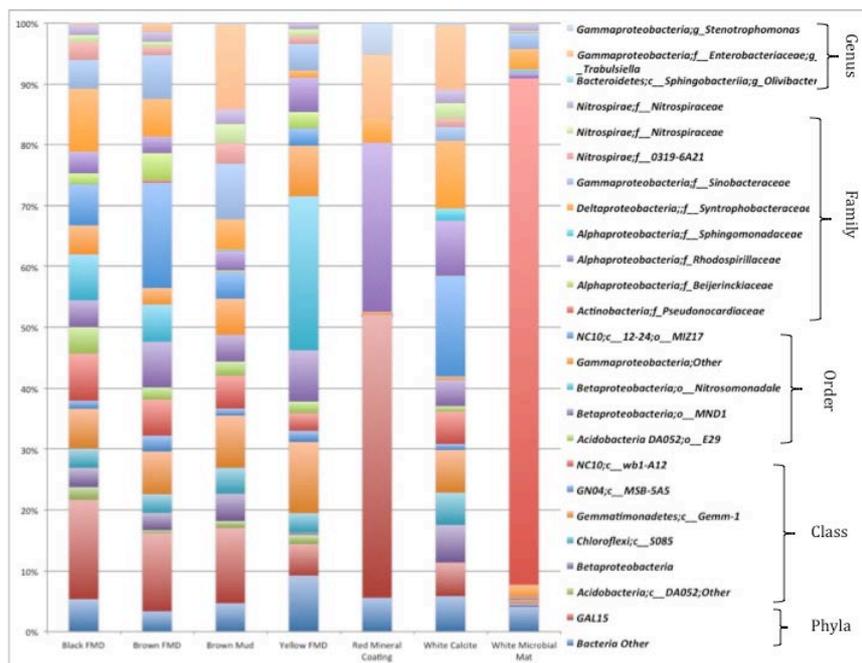


Figure 1. Image from Denise’s presentation, showing how she took up ideas from the lab discussion about how to group her data.

Reflections on finding the story

Denise described a process we documented as occurring abundantly in the lab that supported their story prototyping. Participants often brought forward representations of data, and sometimes worked hard—spending 30 to 60 minutes—to help the other members understand the representation. Many times, this understanding resulted in such significant critiques that the approach was immediately scrapped or the tentative insight dismissed. Denise described an increasing level of tolerance by two of her students for having to analyze and reanalyze data.

Denise: The other thing I'm noticing is in terms of the work that Aaron and James are doing—is the first blush of analysis is—you might as well just throw it away. Um, and then, especially when I get an email that says, “Oh, it's screwed, so I just, you know I ((laughter)) [...] So, I learn to just throw away that email and just wait a few days, and wait for the next email.

Treating such hard work as discardable and tentative is central to making progress as the group learns together. Similar to the “generative reasoning” described by Cross (2011, pp. 146-147) the lab members worked at the problems, rendering them more complex. Tania reflected on how valuable their process of sharing data with one another is, and how she approaches the practice.

Tania: And I think what also helps with, like, a group looking at data, is that they really don't understand—it's actually, I think, better. Because they see other things that if a person who's been staring at this—you know eight hours a day—it gives 'em new perspective. And what I notice I tend to do is, I tend to be like opposite. I try to think opposite of what we should be thinking or like or the way it seems to be going and like, I—I just do that. I don't know why.

Denise: That's a very helpful thing to do

Tania's insight that she takes an oppositional stance does not mean that she simply disagrees with everything. In fact, we saw in the earlier transcript that she came to agree with the story of chaos presented by Aaron. But the perspective shift she suggests is certainly one we observed her—and other lab members making. Such shifts further support the contingent nature of their prototyping; more importantly, such shifts make it safe for students to tentatively posit and/or try on ideas that may be only partially formed, or that may seem unlikely. We see this practice as supporting this lab's innovative work.

Denise: One of the things I find really helpful is we have those discussions, and then Aaron goes off and generates a bunch of new graphs or James does and when we look at those new graphs it's like tuning the focus in and out on a camera. Things come into focus more//

Evan: //that weren't there before//

Tania: //or a microscope//

Evan: //that weren't there before//

Denise: //or a microscope // just out of focus

Denise: Or he'll go off and do a—a graph in a totally different way and sometimes we find out, well, the ecological theory doesn't support that. We need to do it in a different way. And the thing that's most interesting, is the story VANISHES sometimes.

Ultimately, their process does lead to new insights and helps them contribute new understanding in their field, but it is a process with twists and turns, sometimes leading them to pursue prototypes that do not work. In such cases, they abandon them *reluctantly* yet clearly as they move on to other explanations for the data. Such abandonment shares little with how they discarded the story of the bacteria that had been brought in by a packrat, which was easily cast aside because it was not interesting. In contrast, stories that vanish are sometimes mourned; their passing leaves a void, waiting to be filled by new stories.

Conclusions and implications

In the vignettes above, we presented how one interdisciplinary research lab prototyped explanatory stories. We showed how this practice was brought in and cultivated by the PI as a new lab practice, and that lab members sometimes needed to experience it more than one time to see its value. We found that the stories were treated as shared and improvable, providing opportunities for learning. The problems the lab members wrestled with were ill-structured (Jonassen, 1997), and the stories they sought reflected their sense making as they were framing the problem. Part of their sense making involved evaluating the explanatory power (Thagard, 2000) of their stories, much like Bereiter and Scardamalia's (2012) “theory of the case” (p. 162), in which all of the particulars of a case can be coherently explained.

We particularly call out the tentative treatment of the stories, as it was this characteristic that led us to label them as prototypes used for generative reconstruction of the problem (Cross, 2011). Just as in design problem

framing, the prototypes we observed helped lab members “explore a design space or narrow down options and make decisions” (Hess, 2012); they aided them in learning about their ideas, and integrating their ideas with other existing ideas (Ulrich & Eppinger, 1995). By positioning this process of “finding the story in the data” as one of prototyping, we have emphasized the contingent and tentative nature of their progress towards meaning making. This sometimes “bumbling” process is not so different from how learners experience problems new to them, even well-structured ones. Accepting it as not only relevant to—but also central to research process could aid us in designing learning environments that are not brittle to failure and revision, but rather endorse it. We see too much of learning in school settings caught up in seeking certainty, in answering well understood problems with efficiency and accuracy; in contrast, we consider how resilient these members were when discarding a certain, yet uninteresting story, and fording back into the dragons-be-there territory of their data to generate more complex problems on which to work. Allowing students to prototype their ideas iteratively not only supports learning, it also matches practice. Treating such prototypes as contingent and tentative, using the discourse pattern of “finding the story in the data” and engaging from an interdisciplinary stance makes it safe to try on new ideas—and thus, to learn.

Our data were collected in one specific interdisciplinary lab. In seeking promising places to look at laboratory learning, we visited many labs that self-identified as interdisciplinary; we found only a few of these labs included such equitable student participation. Our purpose in this paper was not to identify the specific supports needed to create such participation, which may be requisite to engaging in the type of story prototyping we observed. However, we argue that cultivating this story-finding practice as a form of prototyping—and therefore as a tentative process—could itself help create a culture of risk-taking in learning, where it is safe to *not* know something, to push for more complex problems and ultimately, therefore, to seek deeper levels of understanding.

References

- Bereiter, C., & Scardamalia, M. (2012). Theory building and the pursuit of understanding. *Enhancing the quality of learning: Dispositions, instruction, and learning processes*, 160.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32.
- Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (2004). *Facilitating interdisciplinary research* (9780309094351). Retrieved from http://www.nap.edu/openbook.php?record_id=11153
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice*, 39(3), 124-130. doi:10.1207/s15430421tip3903_2
- Cross, N. (2011). *Design thinking: Understanding how designers think and work*: Berg.
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, 8(4), 597-607. doi:<http://nsuworks.nova.edu/tqr/vol8/iss4/6>
- Hess, T. (2012). *Investigation of prototype roles in conceptual design using case study and protocol study methods*. (MS), Clemson University.
- Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development*, 45(1), 65-94.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39-103.
- LeCompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52(1), 31-60. doi:10.3102/00346543052001031
- MacMath, S., Roberts, J., Wallace, J., & Chi, X. (2010). Research section: Curriculum integration and at-risk students: A Canadian case study examining student learning and motivation. *British Journal of Special Education*, 37(2), 87-94. doi:10.1111/j.1467-8578.2009.00454.x
- Thagard, P. (2000). *Coherence in thought and action*: Cambridge: MIT Press.
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9(1), 77-110.
- Ulrich, K., & Eppinger, S. (1995). *Product design and development*: mcgraw-hill inc. Columbus, USA.
- Weingart, L. R., Todorova, G., & Cronin, M. A. (2008). Representational gaps, team integration and team creativity: The mediating roles of conflict and coordination. *Tepper School of Business*, 528. doi:10.5465/AMBPP.2008.33662047
- Yang, M. C. (2005). A study of prototypes, design activity, and design outcome. *Design Studies*, 26(6), 649-669. Retrieved from <http://www.sciencedirect.com/science/article/B6V2K-4H2G928-1/2/b193b6bc47f3f8eb92ff7dd52a8fa45c>

In Search of Conversational Grain Size: Modeling Semantic Structure Using Moving Stanza Windows

Amanda L. Siebert-Evenstone, Golnaz Arastoopour, Wesley Collier, Zachari Swiecki,
Andrew R. Ruis, and David Williamson Shaffer
alevenstone@wisc.edu, arastoopour@wisc.edu, swiecki@wisc.edu, wcollier@wisc.edu,
arruis@wisc.edu, dws@education.wisc.edu
University of Wisconsin – Madison

Abstract: Analyses of learning based on student discourse need to account not only for the content of the utterances but also for the ways in which students make connections across turns of talk. This requires segmentation of the discourse data to define when connections are likely to be meaningful. In this paper, we present a novel approach to segmenting data for the purposes of modeling connections in discourse. Specifically, we use epistemic network analysis to model connections in student discourse using a temporal segmentation method adapted from recent work in the learning sciences. We compare the results to a purely topic-based segmentation method to examine the affordances of temporal segmentation for modeling connections in discourse.

Keywords: sliding window, epistemic network analysis, segmentation, discourse analysis

Introduction

Analyzing high-volume discourse data is a challenge in computer-supported collaborative learning (CSCL) environments because student conversations in these environments are characterized not only by what is said but by combinations of language use within social practices (Gee, 1990). This suggests that analyses of learning based on student discourse need to account not only for the content of the utterances but also for the ways in which students make connections across turns of talk. Any analysis of such connections, however, requires *segmentation* of discourse data to identify the conditions under which connections are likely to be meaningful (Hearst, 1994). In this paper, we present a novel approach to segmenting data for the purposes of modeling connections in discourse. Specifically, we use *epistemic network analysis* (Shaffer et al., 2009) to model connections in student discourse using a novel *temporal segmentation* method adapted from recent work in the learning sciences (Dyke, Kumar, Ai, & Rosé, 2012; Suthers & Desiato, 2012). We compare the results to a purely *topic-based segmentation* method to examine the affordances of temporal segmentation for modeling connections in discourse.

Theory

There are a number of theoretical perspectives in the learning sciences that describe understanding of a topic, process, domain, or practice in terms of the organization of students' understanding—that is, the way concepts, skills, and habits of mind are related to one another systematically. Chi and colleagues (1981), for example, found that experts in physics organize their understanding differently than novices. Bransford and colleagues (1999) showed that the organization of experts' content knowledge reflects their deep understanding of subject matter. diSessa (1988) suggests that while solving physics problems does require understanding basic concepts from the discipline, deep and systematic understanding comes from linking those concepts to one another within a theoretical framework. Shaffer (2012) similarly characterizes learning as the development of an *epistemic frame*: a pattern of associations among knowledge, skills, habits of mind, and other cognitive elements that characterizes communities of practice, or groups of people who share similar ways of framing, investigating, and solving complex problems.

Not surprisingly, research on discourse processing suggests that connections between concepts are made primarily on a topic-by-topic basis rather than across discourse as a whole. Gernsbacher's (1991; see also Graesser, Gernsbacher, and Goldman, 1997) theory of language processing, for example, suggests that students use the hierarchical organization of content to build understanding. Discourse is structured by topic, with concepts having clear relationships to one another within topics and few relationships across topics.

Based on this idea, *epistemic network analysis* (ENA) analyzes the structure of connections in student discourse by looking at the co-occurrence of concepts within the topics or activities that take place during learning. Building on the idea of learning as the development of an epistemic frame, ENA creates a network model of thinking as based on the co-occurrence of skills, knowledge, values, and other elements of work in some community of practice (Shaffer et al., 2009). In practical terms, ENA measures the structure of the connections

among types of talk by grouping utterances by activity into *stanzas*, or collections of related utterances. In other words, like many lines of work in CSCL and the learning sciences more generally, ENA looks at *activities* as a fundamental unit of analysis.

There are, however, two problems with such an approach. First, as Stahl, Koschmann, and Suthers (2006) argue, learning needs to be analyzed at both the group and the individual level. Stahl (2009), for example, conducted parallel qualitative analyses of the mathematics learning of a group and of the individuals in the group. But as Cress and Hesse (2013) point out, because learners work in groups, simple t-tests and ANOVAs do not do a good job of modeling the influence that groupmates have on one another. Providing a quantitative model of group discourse that accounts for the contributions of any single individual within the group discussion is thus a challenging problem.

A second problem is that aggregating connections by activity may take these connections out of context (Arvaja, Salovaara, Häkkinen, & Järvelä, 2007). While ideas are surely connected within topics or activities, those connections are most likely to occur in close temporal proximity. During discussions, students simultaneously build group and individual understanding by “saying” and replying to “what is said” (Wells, 1999). Speech typically addresses another instance of speech and anticipates a response (Bakhtin, 1986). Because “thinking and speech are, in this sense, always derivative of prior thinking and speech” (Smagorinsky, 2013, p. 23), students build on the ideas of their team members to mediate their discussion of concepts. Therefore, to measure connections in conversations, we need a method to model connection-making on shorter time scales than entire activities.

Recent work by Dyke and colleagues (2012) and Suthers and Desiato (2012) proposes using *sliding window analyses* to model temporal connections in discourse. Rather than creating summary values for all utterances in an activity, a sliding window computes a value for a smaller section of an activity—typically a small amount of time (e.g., 10 seconds) or a small number of utterances (e.g., three turns of talk; Dyke et al., 2012). The window is *sliding* in the sense that a summary value is computed for each utterance, based on the preceding lines of talk (e.g., the preceding 10 seconds or three lines of talk). This type of analysis has been used to identify shifts in topic (Rosé et al., 2008), and more generally to provide new insights on previously analyzed data (Dyke et al., 2012). Suthers and Desiato (2012) have used a sliding window approach to build a model of uptake—that is, to model connections in discourse. However, while their model showed when each actor used another actor’s contribution, this model only showed whether a connection was made, not *what* connection was made.

In what follows, we use ideas from Gee (1991) to create an ENA model of connections in discourse using a moving window approach. When analyzing discourse, Gee argues that single *lines* or utterances in talk are grouped together into sets of related lines he calls *stanzas*. His analogy is to stanzas in a poem, and this is the sense in which ENA groups turns of talk to model the co-occurrence of ideas. But Gee also suggests that stanzas themselves are grouped together into related sets that he calls *strophes*.

In this study, we use the idea of strophes and stanzas to delineate two different approaches to modeling connections using ENA, although it may be useful in other modeling approaches as well. In both cases, ENA models connections among concepts: (1) by identifying coherent topics, activities, and/or conversations in the data as strophes; and (2) by defining collections of utterances within strophes that are related to one another as stanzas. The two methods differ in the relationship between strophes and stanzas. Specifically:

1. The **Strophe Method** models connections within an entire activity or strophe: that is, all the utterances within an activity are related to one another. Or, equivalently, each strophe is composed of a single stanza.
2. The **Moving Stanza Window Method** models connections within an activity or strophe by *dividing* the strophe into multiple stanzas: that is, utterances are related to one another only within some designated stanza window. In other words, the moving stanza window method models connections only when utterances are in close temporal proximity within a strophe.

In what follows, we compare these two models by looking at data from a CSCL learning environment in which students collaboratively design solutions to engineering problems. To evaluate the strengths and limitations of these two approaches to segmentation, we created ENA models for 10 teams using both the strophe method and the moving stanza method. We focus here on the discourse of one team, and we ask:

What are the similarities and differences in how the strophe method and the moving stanza window method of segmenting data in ENA characterize the nature of connections in discourse?

Methods

The engineering virtual internship *RescuShell*

In this study, we analyzed how students who are roleplaying as engineering interns in a virtual internship interact within their teammates. In *RescuShell*, student teams conduct research and simulated experiments to develop the robotic legs for a mechanical exoskeleton for use by rescue personnel. The virtual internship is separated into 17 activities that simulate various steps in the design process, including reviewing and summarizing research reports, creating device prototypes, discussing design choices with teammates, and working to meet the needs of various internal consultants and external clients. In this study, we focused on the first eleven activities of the internship, in which students were assigned to one of five teams, each of which explores the use of a particular actuator in the design (Hydraulic, PAM, Electric, Pneumatic, or Series Elastic). Forty-four first-year engineering students participated in the virtual internship, which took approximately 15 hours to complete. From this sample, we selected one team and analyzed how these five students (4 male, 1 female) discussed the design problem in the first half of the internship.

Discourse analyses

Coding student chats

We coded each line of chat data using our Engineering Epistemic Frame Coding scheme, which identifies domain-specific epistemic frame elements (Shaffer & Arastoopour, 2014). We applied this coding scheme by using an automated key word coding process that has been validated by comparing agreement between human and computer codes with resulting Cohen's kappa scores between 0.80 and 0.98 (Chesler et al., 2015). The scheme includes 10 codes:

Data-based Justifications: Justifying decisions using data such as graphs, results, or numerical values.

Design-based Justifications: Justifying decisions using design references such as prioritization.

Client-based Justifications: Justifying decisions by referring to the client's safety, health, or comfort.

Consultant-based Justifications: Justifying decisions by stating the internal consultants' preferences.

Skill of Data: The action of using numerical values, results tables, graphs, or research papers.

Skill of Design: The action of design development, prioritizing, tradeoffs, and design decisions.

Skill of Collaboration: The action of facilitating a team meeting.

Identity of Engineer: Identifying as an engineer; possession or ownership of work.

Knowledge of Attributes: Referring to attributes: payload, recharge interval, agility, safety, or cost.

Knowledge of Inputs: Referring to inputs: actuators, ROM, materials, power sources, or sensors.

We then performed a Chronologically-Oriented Representations of Discourse and Tool-Related Activity (CORDTRA) analysis (Hmelo-Silver, Liu, & Jordan, 2009) during one activity to show the temporal pattern of the 10 codes in student discourse.

Epistemic Network Analysis

ENA models the structure of connections among engineering epistemic frame elements by quantifying the co-occurrences of codes within a stanza (Shaffer et al., 2009; Shaffer 2014). After defining the segmentation structure, ENA creates an adjacency matrix representing the co-occurrences of codes in each stanza. To construct an adjacency matrix, ENA assigns a one for each unique pair of codes that co-occur one or more times in those utterances, and a zero for each unique pair that does not appear anywhere in the stanza. ENA sums the adjacency matrices into a cumulative adjacency matrix, where each cell represents the number of stanzas (i.e., the number of adjacency matrices) in which that unique pair of codes was present. Each unit of analysis is thus represented by a cumulative adjacency matrix that summarizes the pattern of connections among codes.

ENA then converts the cumulative adjacency matrices into cumulative adjacency vectors that are projected into a high-dimensional space based on the co-occurrence of codes across segments. These cumulative adjacency vectors are normalized to control for the varying lengths of vectors by dividing each vector by its length; the resulting vector thus represents the relative frequency of co-occurrences. ENA then performs a singular value decomposition on the normalized vectors. This produces a rotation of the original high-dimensional space, such that the rotated space provides a reduced number of dimensions that capture the maximum variance in the data.

The resulting models can be visualized as networks in which the nodes in the model are the codes and the lines connecting the nodes represent the co-occurrence of two codes. Thus we can quantify and visualize the

structure of connections among engineering epistemic frame codes, making it possible to characterize student discourse during the virtual internship.

Comparison of segmentation procedures

In this study, we compared two methods of segmenting data for use in ENA: the strophe method and the moving stanza window method. For the *strophe method*, ENA creates one adjacency matrix for each activity and then sums the matrices across the 11 activities for a given team.

The *moving stanza window method* creates a *referent adjacency matrix* for each utterance, known as the *referring utterance*. The referent adjacency matrix for each utterance is constructed of two types of co-occurrences of codes: (1) co-occurrences within the referring utterance, and (2) co-occurrences of codes between the referring utterance and a specific number of previous utterances, known as the *window*. The moving window then moves to the next referring utterance and creates the next referent adjacency matrix. This process continues until the end of the strophe and then ENA sums the matrices across all utterances for that unit. No windows are made across activities (strophes), only within them. Figure 1 shows how the strophe method and the moving stanza window method create different models of connectivity.

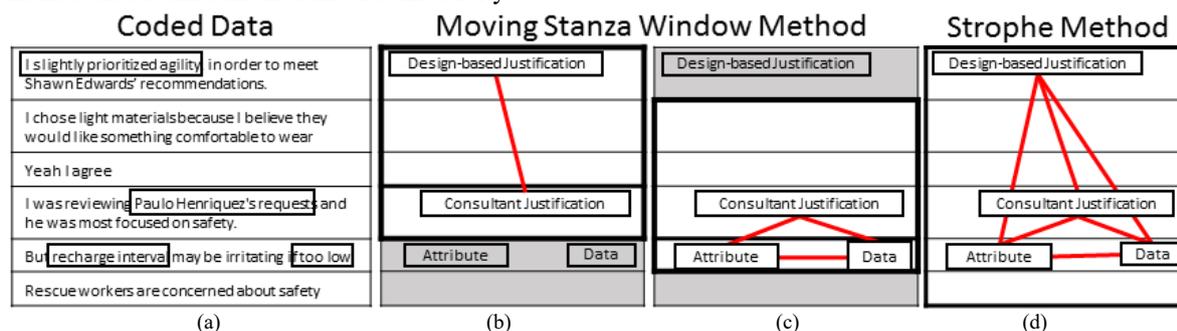


Figure 1. Example of coded data from one activity (a). The moving stanza window method analyzes connections within the referring utterance and between the referring utterance and the window (b). After analyzing a window, the moving stanza method slides to the next utterance and repeats the process of finding connections within and between the referring utterance and the window. The strophe method analyzes all connections in an activity (d).

Co-occurrences of codes within or across non-referring utterances are not included in the referent adjacency matrix, which eliminates double-counting of connections when the cumulative adjacency matrix is computed.

Comparison of network models

To analyze the different segmentation methods using ENA, we created three models: (1) a strophe model for all teams in the sample, (2) a moving stanza window model with a window size of three for all teams in the sample, and (3) a moving stanza window model with a window size of three for all students in the sample, based on a qualitative analysis of the data that suggested most explicit connections between ideas in the discourse occurred within a span of 4 or fewer lines (the referring utterance plus the preceding three turns of talk). All three of these sets were projected into the dimensional reduction for the team moving stanza model (Model 1) so the resulting networks could be compared. To analyze the differences between the two segmentation methods, we examined the discourse of one team whose models appeared to show different conclusions between methods. First, we compared Model 1 with Model 2 to understand how each segmentation type modeled team discourse, then we used Model 3 to contrast connections between individuals.

Results

For the purposes of this analysis, we looked at the conversations of one student project team. The Hydraulic team had five team members: Arden, Connor, Margaret, Jimmy, and Jordan. In what follows, we examine their collaborative design work over the first 11 activities of the virtual internship, which include background research into principles of biomechanics, as well as the design, testing, and evaluation of an initial prototype for a robotic exoskeleton.

Strophe and moving stanza window models for the hydraulic team

We used the strophe method and the moving stanza window method to model the discourse of the team. Both models (see Figure 1) show that the connections to and between the Skill of Data and the Knowledge of Inputs were prominent in the group's design discussions. This is represented by larger node sizes and thicker lines in the ENA network graph linking the nodes that correspond to those discourse elements. This is, of course, hardly

surprising, as the group’s primary goal was to choose appropriate design features (inputs) to maximize the function of their device.

However, the strophe method (Figure 2a) suggests that the Hydraulic team connected these features of design with explicit discussion of their collaboration process; in contrast, the moving stanza window method (Figure 2b) suggests that the team spent less time explicitly connecting talk about collaboration to their design work and more time talking linking the Skill of Design to other elements of the problem space, representing explicit discussion about the tradeoffs involved in the design process.

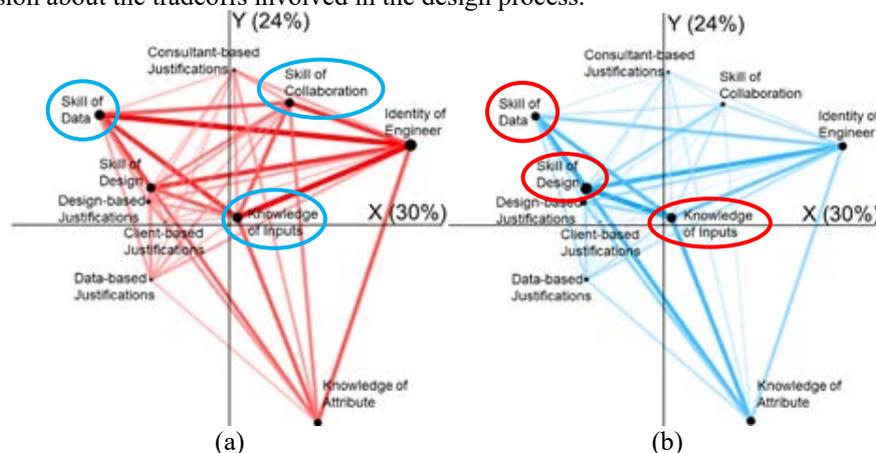


Figure 2. Network graphs of the Hydraulic team’s discourse produced using (a) the strophe method and (b) the moving stanza window method. Thicker lines denote more frequent connections between codes. Percentages indicate the amount of variance explained by each dimension, in this analysis, 54% of the total variance is accounted for in this data set.

This contrast is shown more clearly by computing the *difference* between the two network models (Figure 3). Figure 3 shows a higher number of connections in the strophe method (red lines in the figure) to the node for Skill of Collaboration, suggesting that links between the Skill of Collaboration and other elements of the epistemic frame of engineering are a prominent feature of student discourse in this model. In contrast, the moving stanza window method (blue) suggests that students made more connections to the Skill of Design.

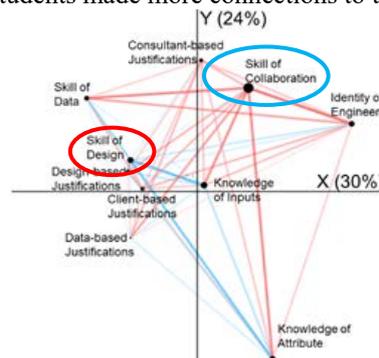


Figure 3. Subtracted network of the Hydraulic team’s discourse, in which blue connections occur more frequently with the moving stanza method and red connections occur more frequently with the strophe method.

Comparing connections within activities

To explore these differences between the two models, we examined the frequency of codes within each activity in the virtual internship. For example, when students met with their teammates to design devices, the discourse included references to the Skill of Collaboration, which was one of the key differences between the two models. To understand why there was such a substantial difference in connections to the Skill of Collaboration, we examined the CORDTRA for this activity (Figure 4).

The CORDTRA shows that students explicitly talked about collaboration only at the start and at the end of the activity. Applying the strophe method to this activity produced connections between Skill of Collaboration and codes that appeared *at any point within the activity*, even though students only talked explicitly about collaboration at the beginning and the end of the discussion.

In contrast, applying the moving stanza window produced connections between codes only if they co-occurred within close proximity, that is, within three utterances of the referring utterance. Thus, the moving stanza window model shows a less prominent role for the Skill of Collaboration.

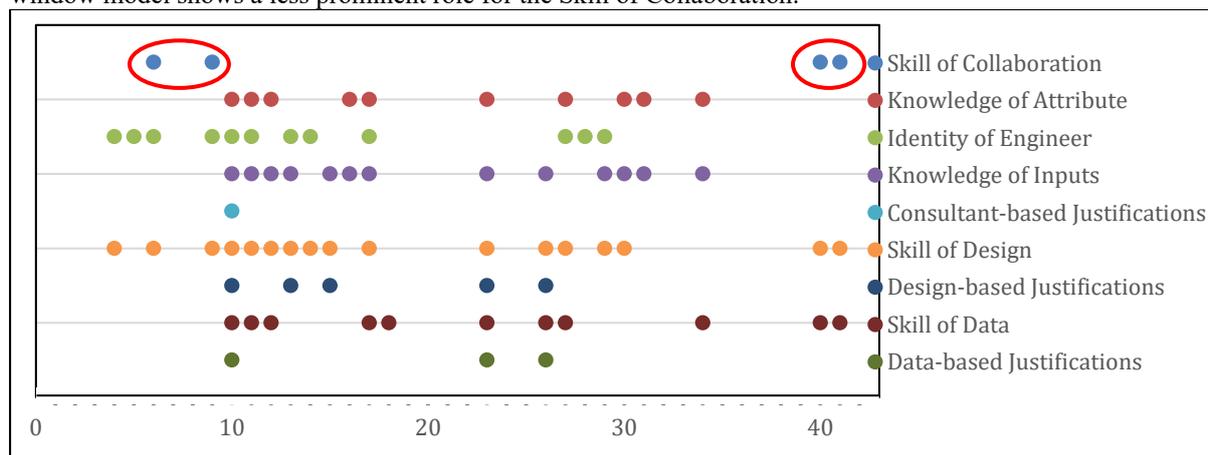


Figure 4. CORDTRA diagram of Hydraulic team discourse codes during one activity.

Contrasting connections between individuals

A second consideration in comparing the strophe method and the moving stanza window method is that the strophe method suffers from the same limitation as many extant techniques for modeling CSCL: it can model a group conversation, but it does not do a good job of modeling the participation of one individual *in the context of* a group discussion. The moving stanza window method, in contrast, can account for this important component of collaborative learning.

The reason for this difference is that the strophe method uses a *single adjacency matrix to model each activity*, and that matrix incorporates the contributions of all members of the group. There is thus no good way to disentangle the contribution of any one individual. The moving stanza window method, on the other hand, *models each utterance as an adjacency matrix*, showing the connections it contributes to the group discourse. As a result, we can use the moving stanza window method to examine the connections that each individual makes to the collaborative discussion of the group.

For example, we modeled the contributions of two students, Connor and Jimmy, to the Hydraulic team's discussion. We constructed a network model of each of two students' contributions, where each model includes only those stanza windows in which the referring utterance belonged to that individual (Figure 5). These models thus represent the unique contributions to the team discussion made by each student.

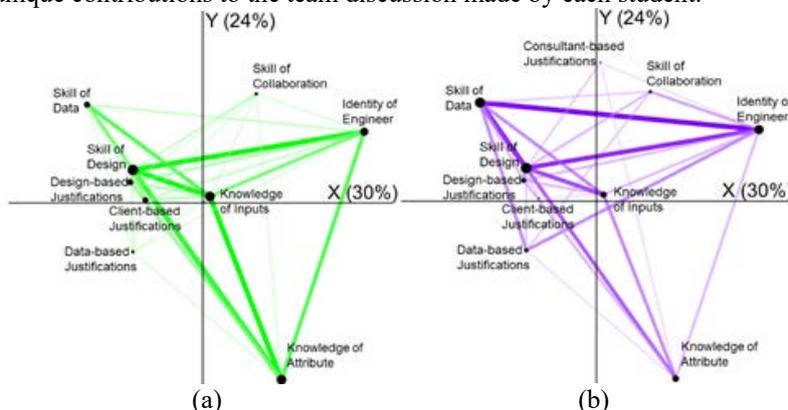


Figure 5. Moving stanza window model for Connor (a) and Jimmy (b). Thicker lines denote more frequent connections between codes.

The networks using a moving stanza window method show that across all eleven activities, Connor and Jimmy's individual contributions to the group discourse differ. Connor's network (Figure 5a) shows a higher number of connections between Knowledge elements in the design domain suggesting that he commonly contributed information about Attributes and Inputs in design discussion. On the other hand, Jimmy's network

(Figure 5b) shows a higher number of connections between Skill of Data and Skill of Design suggesting that his utterances integrated design tradeoffs into the group conversation.

Table 1 illustrates this in a short excerpt from one of the group’s discussions about interpreting experimental data. In this excerpt, Jimmy’s second comment (Line 2) makes a connection between the Skill of Data and the Skill of Design. He argues that graphs showing the results of benchmark testing (Skill of Data) help the team make an “informed decision” (Skill of Design) about their design choices. Two turns later (Line 4), Connor adds to the discussion by introducing information about specific attributes and inputs of the design: he talks about the performance parameters (payload, agility, and battery life) of some of the design choices that the team is considering (cadmium batteries and piezoelectric sensors).

Critically, this model using the moving stanza window method shows that it is Connor who builds on Jimmy’s discussion about data and design by contributing information about inputs and attributes. The moving stanza window methods separately models both Jimmy’s original contributions to the team discussion and the fact that Connor’s contribution builds on Jimmy’s utterance two lines before.

Table 1: Brief excerpt of the Hydraulic team’s discussion of findings during the graphing activity.

	Student	Chat Utterance	Code
1	Jimmy	They all had both advantages and disadvantages. There was no "obvious" best choice.	
2	Jimmy	The graphs indicated the properties of all the different options and made a comparable visual illustration to make an informed decision on which combination to use.	Skill of Data, Skill of Design
3	Jordan	The graphs detailed what aspects of power sources and control sensors are important--namely, the numerical data.	Skill of Data, Knowledge of Inputs
4	Connor	I suggested using cadmium batteries with piezoelectric sensors , together they make a strong combination of payload and agility while keeping costs in a moderate range and having strong battery life .	Knowledge of Inputs, Knowledge of Attributes

Discussion

Our results thus suggest that the strophe method and the moving stanza window method identified different types of connection-making in student discourse. In particular, the strophe method summarized the connections made by student teams based on activity, but it could not differentiate individual contributions to team discussions. The moving stanza method, in contrast, accounted for the connections that were made based on activity and temporal proximity; importantly, this method was also able to model the contributions of individual students to team conversations.

Of course, which of these models is most appropriate depends on the theory of discourse that is being modeled. For example, if we assume that talk at the beginning of an activity frames everything that follows—or similarly, if talk at the end of an activity builds on everything that preceded it—then the strophe method is more appropriate, because it models connections among all of the talk within a single activity. If, on the other hand, we want a model that is sensitive to the temporal proximity of talk, then the moving stanza window method is a better choice, as it models connections locally within an activity, such that very early turns of talk are not related to ideas that arise much later in the discussion. In addition, the moving stanza window has the benefit of also modeling the role of individual contributions to group discussions.

This study, of course, is limited in that it focused on the activities of one group of students working in one CSCL context. However, this work highlights empirically a key theoretical distinction between models of connectivity in discourse, and perhaps more importantly, it demonstrates that the moving stanza window method makes it possible to use ENA to model both group discourse and the contributions of individuals to the group within a CSCL context.

References

- Arvaja, M., Salovaara, H., Häkkinen, P., & Järvelä, S. (2007). Combining individual and group-level perspectives for studying collaborative knowledge construction in context. *Learning and Instruction, 17*(4), 448-459.
- Bakhtin, M. (1986). *Speech genres and other late essays*. Trans. Vern W. McGee. University of Texas Press.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.
- Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121-152.

- Chesler, N. C., Ruis, A. R., Collier, W., Swiecki, Z., Arastoopour, G., & Shaffer, D. W. (2015). A novel paradigm for engineering education: Virtual internships with individualized mentoring and assessment of engineering thinking. *Journal of biomechanical engineering*, 137(2), 024701.
- Cress, U., & Hesse, F. W. (2013). Quantitative methods for studying small groups. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, and A. O'Donnell (Eds.), *The international handbook of collaborative learning*, (pp. 93-111). Routledge.
- DiSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 47-70). Lawrence Erlbaum Publishers.
- Dyke, G., Kumar, R., Ai, H., & Rosé, C. P. (2012). Challenging assumptions: Using sliding window visualizations to reveal time-based irregularities in CSCL processes. In *Proceedings of the 10th International Conference of the Learning Sciences* (Vol. 1, pp. 363-370).
- Gee, J. P. (1990). *Social linguistics and literacies: Ideology in discourses*. Falmer Press.
- Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 217-263). Academic Press.
- Graesser, A. C., Gernsbacher, M. A., & Goldman, S.R. (1997). Cognition. In T. A. van Dijk (Ed.), *Discourse: A multidisciplinary introduction* (pp. 292-319). Sage.
- Hearst, M. A. (1994). Multi-paragraph segmentation of expository text. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics* (pp. 9-16). Association for Computational Linguistics.
- Hmelo-Silver, C. E., Liu, L., & Jordan, R. (2009). Visual Representation of a Multidimensional Coding Scheme for Understanding Technology-Mediated Learning about Complex Natural Systems. *Research and Practice in Technology Enhanced Learning*, 4(3), 253-280.
- Rosé, C., Wang, Y. C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, 3(3), 237-271.
- Shaffer, D. W. (2012). Models of situated action: Computer games and the problem of transfer. In C. Steinkuehler, K. Squire, S. Barab (Eds.), *Games learning, and society: Learning and meaning in the digital age*, (pp. 403-433). Cambridge University Press.
- Shaffer, D. W. (2014). *User guide for Epistemic Network Analysis web version 3.3*. Madison, WI: Games and Professional Simulations Technical Report 2014-1.
- Shaffer, D. W., & Arastoopour, G. (2014). *Guide to RSdata.csv sample ENA data set*. Madison, WI: Games and Professional Simulations Technical Report 2014-3
- Shaffer, D. W., Hatfield, D., Svarovsky, G., Nash, P., Nulty, A., Bagley, E. A., ... Mislevy, R. J. (2009). Epistemic Network Analysis: A prototype for 21st century assessment of learning. *The International Journal of Learning and Media*, 1(1), 1-21.
- Smagorinsky, P. (2012). *Vygotsky and literacy research: A methodological framework* (Vol. 2). Sense Publishers.
- Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, (pp. 409-426). Cambridge University Press.
- Stahl, G. (2009). *Studying virtual math teams*. Springer Science and Business Media.
- Suthers, D. D., & Desiato, C. (2012, January). Exposing chat features through analysis of uptake between contributions. In *45th Hawaii International Conference on System Science* (pp. 3368-3377). IEEE.
- Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology*, 79(1), 1-28.
- Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. Cambridge, England: Cambridge University Press.
- Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

Acknowledgments

This work was funded in part by the National Science Foundation (DRL-0918409, DRL-0946372, DRL-1247262, DRL-1418288, DUE-0919347, DUE-1225885, EEC-1232656, EEC-1340402, REC-0347000), the MacArthur Foundation, the Spencer Foundation, the Wisconsin Alumni Research Foundation, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.

Realizing Research-Practice Connections: Three Cases From the Learning Sciences

Susan McKenney, University of Twente, susan.mckenney@ou.nl
Jan van Aalst, University of Hong Kong, vanaalst@hku.hk
Cory Forbes, University of Nebraska-Lincoln, cory.forbes@unl.edu

Abstract: Empirical insights from the learning sciences must be translatable to contexts of application to maximize impact. Despite the importance of dissemination, however, little guidance is available to help researchers proactively bridge the research-practice divide. To address this need, we draw from theoretical perspectives on the utility of empirical research to characterize, illustrate, and discuss three cases involving the communication and translation of research. By offering framework for analysis and initial findings based on three projects in the learning sciences, this paper helps learning scientists conceptualize a range of research-practice connection types and envision how they can take shape in learning sciences projects.

Keywords: research valorization, impact, DBR, RDD, teacher community

Introduction

A robust body of knowledge now exists to describe how policymakers and educators access, value and use research (Broekkamp & van Hout-Wolters, 2007; de Vries & Pieters, 2007; Vanderlinde & van Braak, 2010); various modes through which knowledge is generated and shared (Bauer & Fischer, 2007; Lavis, et al., 2003); and what aspects of evidence-based practice and research utilization in other fields can be applied to education (Thomas & Pring, 2004). However, both the scholarly insights and effective practices have yet to become widely spread. Even though researchers are becoming increasingly required to disseminate research findings among practitioners, few models are available in the learning sciences and many researchers find it daunting. The present paper addresses this need by (a) outlining modalities of fruitful research-practice connections; and (b) analyzing how these modalities are reified in three existing projects in the learning sciences.

Theoretical underpinnings

Factors affecting the uptake and use of scientific outputs

Educational research has long been criticized for its weak link with practice. Explicit attempts to close the research-practice gap have been underway for over four decades. Informed by the work of Rogers (1969), and review of over 2600 research studies, Havelock (1971) published a landmark report on the dissemination and use of scientific outputs. Havelock identified seven general factors that could account for how scientific outputs are taken up and used: linkage, structure, openness, capacity, reward, proximity and synergy. *Linkage* refers to the number, variety and mutuality of research-practice collaborative relationships. *Structure* pertains to the degree of systematic organization around four factors: the resource system, the user system, the dissemination-utilization strategy, and the (coherence of the) message. *Openness* is the belief that change is desirable and possible; this is accompanied by a social climate that is willing and ready for change. Highly correlated with power, size and experience, *capacity* is the capability for retrieving and marshaling resources. *Reward* has to do with the frequency, immediacy, amount, mutuality, planning and structuring of positive reinforcements. Nearness in time, place and context constitute *proximity*, which hinges on familiarity, similarity and recency. Finally, Havelock refers to *synergy* as the number, variety frequency and persistence of forces that can be mobilized to produce a knowledge-utilization effect. In the learning sciences, these issues remain highly pertinent today.

Modes of research-practice interaction

Based on this synthesis, Havelock identified several modes in which those factors can be seen: social interaction; research, development and diffusion (RDD); and problem solving. More recently, attention has also been given not only to the use of scientific knowledge for educational practice (e.g. Hargreaves, 1999; Levin, 2004), but also to how it is produced (Vanderlinde & van Braak, 2010). Specifically, there is growing attention for how researchers and practitioners can collaboratively bear the responsibility for both producing and using relevant knowledge in education. Burkhardt and Schoenfeld (2003) identify seven models to describe the relationship between research and practice, five of which feature strong divisions of labor, relate more to evidence-based

practice and align well with the RDD model described by Havelock (the reading model; the summary model; the professional development model; the policy model; the ‘long route’) and two of which show more characteristics of Havelock’s problem solving model (design experiments; and the engineering model). De Vries and Pieters (2007) add an eighth model which shares elements of Havelock’s social interaction model and highlights equal collaboration: knowledge communities. Each of these models denotes different assumptions and expectations regarding the roles of practitioners and researchers in the generation and application of theoretical understanding. Taken together, three broad types of research-practice interactions during knowledge production can be distinguished in education today: RDD, design-based (implementation) research DB(I)R and teacher communities (Ormel, Pareja Roblin, McKenney, Voogt, & Pieters, 2012; Pareja Roblin, Ormel, McKenney, Voogt, & Pieters, 2014; Voogt, McKenney, Pareja Roblin, Ormel, & Pieters, 2012). The most dominant model in scientific research, RDD is based on the notion that researchers deliver knowledge, intermediaries translate this knowledge into usable products for practice, and that professionals use knowledge in the form of the products. Here, DB(I)R refers to a family of research approaches including design-based research (DBRC, 2003) and design-based implementation research (Penuel, Fishman, Cheng, & Sabelli, 2011), that share the dual aims of (1) deriving new knowledge through (2) collaboration between researchers and practitioners, to iteratively design and implement durable solutions to real-world problems. In *teacher communities*, researchers and educators work together to understand and improve teachers’ existing practice, mostly through iterative cycles of observation and reflection.

Toward publically-accessible learning sciences insights

Internationally, enormous efforts have been launched to improve the practical relevance and actual use of research knowledge, especially in the fields of education and health care. As demonstrated above, common types of research-practice interactions have been identified, as well as factors that support success. Additionally, crucial aspects of evidence-based practice and research utilization from other fields can be applied to the learning sciences. However, both the scholarly insights and effective practices have yet to become widely spread in the learning sciences. Even though researchers are becoming increasingly stimulated to disseminate research findings among practitioners, few research programs devote serious attention to preparing their faculty for the task, and many researchers find it daunting. Further, sustainable modalities for such work are lacking, especially due to the fact that dominant promotion and tenure systems reward other output (e.g. scientific publications). For many, the result is that outreach work takes place primarily during personal time and is therefore limited. Guidance is needed to support learning scientists in the daunting but rewarding task of shaping research-practice interactions such that new scientific insights become accessible and usable in practice. The remainder of this paper describes, analyzes, and reflects upon three cases of research-practice interactions intended to facilitate outreach in the learning sciences.

Methods

Focus and approach

Ultimately, this study was undertaken to facilitate learning scientists in perceiving and shaping fruitful research-practice interactions. To reach that goal, we sought an answer to the overarching research question, “*How are the three modalities (RDD, design-based research, and teacher communities) and factors (linkage, structure, openness, capacity, reward, proximity and synergy) reified in existing learning sciences projects?*” The nature of this largely descriptive and exploratory question necessitated the articulation of a set of case studies (Yin, 2014) to help identify observed patterns in factors associated with particular modalities. As this was an initial, exploratory analysis, we choose to focus on projects we already know very well and for which information was readily available (convenience sampling). Within those, we sought one project of each type, i.e. RDD, design-based research, and teacher communities. A final selection criterion was that projects would be sufficiently mature that their core orientation toward research-practice interaction had stabilized (even if their specific activities continued to evolve). Each project selected represents a unique case conducted in three unique contexts with diverse, multifaceted designs, partners, and target populations. The modalities and factors associated with each case represent embedded units of analyses for development of an embedded, multiple-case study.

Data collection and analysis

The data collection and analysis was based on first-hand knowledge of the projects as derived from document analysis, archival records, and attending project meetings and engaging firsthand in project work as participant observers. While the present study was conducted in a post-hoc fashion (after project design and data collection), its reliability and validity are enhanced by access to similar data sources in each project. These variant data sources, as well as the role of each author in providing project leadership, enables a significant degree of data

triangulation for embedded analyses within each case. The data analysis was guided by the theoretical framework above, and three sub-questions in particular. Specifically, we sought to understand:

- What are the goals of research-practice interactions in this project, and how explicitly were outreach goals part of the project agenda?
- What features are salient to understanding this project as an example of its type (RDD, design-based research or teacher communities)?
- How does this project realize the factors that are crucial for uptake and use of new knowledge (linkage, structure, openness, capacity, reward, proximity and synergy)?

Each factor (linkage, structure, openness, capacity, reward, proximity and synergy) was examined for each project. The presence of each factor was ranked as being high, moderate, low or variable.

Results

RDD: Science Literacy Initiative

How can we successfully feed 2.4 billion more people by 2050? How can we utilize the Earth's natural resources in a sustainable manner to do so? These are critical questions that face the global community in the 21st Century and lie at the heart of notions of science literacy. Ultimately, scientists, policymakers, and members of the public must work together to find solutions to these significant challenges and pave the way to a sustainable future. This involves an enhanced capacity, both at the individual and collective levels, to make effective decisions grounded in STEM-informed analyses of complex, real-world challenges associated with agricultural, environmental, natural resource, and technology issues. We confront these challenges and address this need through the Science Literacy Initiative (SLI) at the University of Nebraska-Lincoln, which uses food, energy, and water systems as a platform for a comprehensive suite of programs designed to foster science literacy among PK-16 students, the public, and stakeholders.

A critical dimension of the SLI involves the development of education and communication strategies that effectively reach a diverse array of audiences. This research-practice interaction involves translation of rigorous empirical research on both a) food, energy, and water systems (science) and b) human learning and behavior (social science) into program models and products that effectively foster intended outcomes with target populations. This includes theoretical perspectives and research from basic and applied scientific disciplines, cognitive psychology, the decision sciences, organizational theory, and STEM education. Rather than serve merely as a translational intermediary, however, the SLI spans the research, development, and dissemination domains of the RDD process through transdisciplinary partnerships involving individuals with a range of expertise in relation to specific programs and audiences. While accounting for stakeholder needs, goals, and objectives is an important part of this process, it is nonetheless largely one-directional in design. This linearity presents significant challenges for all research organizations involved in dissemination (e.g., Lavis et al., 2003).

To address this challenge, we theorize and operationalize RDD as a form of 'decision support' for engagement with STEM-based dimensions of everyday life as consumer, citizens, and employees. This is a difficult process, however, as individuals are prone to snap judgments that are quick, irrational, and subject to error. Research and theory from the decision sciences provides a multitude of perspectives on how humans arrive at decisions for complex systems and their associated challenges. These perspectives differ across a number of dimensions that define decision-making along a continuum, such as a) the temporal scales within which decisions are made and b) the number of factors for which decision-makers must account. Some perspectives emphasize intuitive, experience- and instinct-driven decision-making in real-time in high-stress situations (naturalistic decision-making). Others foreground weighing multiple options based upon a complex set of interacting and overlapping criteria. This second perspective, often referred to as multi-criteria decision-making (MCDM), accounts for how decisions about complex issues are made over longer periods of time. We argue that making high-quality, STEM-informed, actionable decisions about complex issues associated with diet, food production, natural resources, transportation, and other contemporary challenges involves being deliberate, rational, and attuned to uncertainties, a process more aptly characterized by MCDM.

To optimize the process and outcomes of the SLI, we emphasize procedural factors identified by Havelock (1971). *Linkage* is a defining element of the initiative itself, being designed to leverage significant and diverse expertise to engage in RDD through multiple programmatic channels to reach a broad set of constituents. Contributors to SLI include faculty from STEM disciplines, social, behavioral, and learning scientists, K-16 educators, and external stakeholders from government and industry. To facilitate initiative *structure*, we have benchmarked all programs (including the resource and user systems) in a shared conceptual framework and set of heuristics for science literacy defined by enhanced proficiency to engage in STEM-informed decision-making.

These heuristics are grounded in theoretical perspectives on multi-criteria decision-making drawn from the decision sciences that foreground framing and problematizing issues, defining their boundaries, identifying and interrogating viable options and strategies for action, and justification of decisions. By grounding programming in contemporary socio-environmental issues, such as water resource use, environmental degradation, food production, and energy, initiative programming also foregrounds *proximity* and close relevance to the lives of individuals served. However, these factors also help illuminate challenges associated with SLI. For example, we remain highly attuned to *capacity*. Personnel resources are crucial not only to knowledge-based contributions to SLI, but also to translational development and engagement. However, the SLI often brings to the fore organizational tensions associated with the allocation of personnel commitments across multiple and competing efforts. *Openness* and *reward*, being largely qualities of the target audience, remain important considerations for program design and for embedded research to understand how initiative efforts and enhanced science literacy influence social, cultural, economic, and civic dynamics. SLI programming must provide a ‘value proposition’ with intrinsic benefits to its target audiences who, themselves, must be willing to envision change. Each of these factors influence the overall *synergy* contributing to the knowledge-utilization effect underlying the initiative and its emphasis on STEM-informed analyses of complex, real-world challenges associated with agricultural, environmental, natural resource, and technology issues.

DB(I)R: PictoPal

PictoPal is the name of a technology-rich learning environment for early literacy. The primary goal of the learning environment is to teach kindergarteners about the nature of written language. Supporting this goal, the environment also helps children recognize the relationship between spoken and written language; and various functions of print. PictoPal consists of connected on-computer and off-computer activities. The on-computer activities scaffold the creation of written products (e.g. letters, poems, lists). The off-computer activities use prints of the products created with computer support as literary props or for authentic purposes (e.g. letters are mailed, poems are read aloud, groceries are ‘bought’ in the store corner of the classroom, etc.). Teachers can easily tailor the contents of on-computer and off-computer tasks within each PictoPal module.

The PictoPal learning environment and related scientific understanding have evolved through collaboration between researchers and practitioners. The collaboration revolved around the iterative development of PictoPal prototypes, thereby addressing two challenges experienced by kindergarten teachers. Namely, designing and using PictoPal helped them (1) integrate activities that addressed crucial but typically under-attended learning goals in the Dutch language curriculum for kindergarteners; and (2) develop their competencies for using technology and understanding early literacy. The PictoPal work has yielded both practical and scientific outputs. From a practical standpoint, this work has produced a usable and effective tool that can easily be adapted for a wide range of kindergarten classrooms. From a scientific standpoint, this work has produced insights related to (Citations removed for review):

- Developing (language) software for young children
- Teacher roles during the design of technology-rich curriculum materials
- Teacher knowledge, beliefs and reasoning as revealed through conversations during design
- Teacher beliefs about early literacy curriculum implementation

Through retrospective analysis, we see that this project did realize linkage, structure, openness, capacity, reward, proximity and synergy. *Linkage* was realized through the collaboration with over a dozen participant groups, ranging from individuals to kindergarten teachers within a school, to entire school districts. All participation was voluntary, and mutual benefits were present (data informed theoretical understanding and school decision making, the schools kept the PictoPal resources they helped create and gained access to ones made by others). The *structure* featured systematic organization with regard to expectations and tasks from the research team (resource system), the teacher teams (user system), the implementation and use of PictoPal (dissemination-utilization strategy), and what had been learned as a result (message). *Openness* was largely served by the sampling procedures. Even though individuals varied in their opinions of how desirable and feasible the change would be in their own classrooms, the choice to work with schools that volunteered because they saw intrinsic value in PictoPal typically yielded a social climate ready for change. All iterations of PictoPal design research benefitted from *capacity* within the research team, and most also included involvement of participants with authority and/or experience that further helped marshal resources. Early positive findings related to pupil learning gain from use of PictoPal increased the *reward* for all those involved. Further, the rich insights and usable materials served as positive reinforcements. *Proximity* was achieved by the explicit choice to work face-to-face with teachers in their own schools on a regular and (in bursts) intense schedule. For most participating teachers,

this may have been increased by early exposure to materials previously designed by other Dutch kindergarten teachers (familiarity, recency). Finally, *synergy* was present in all iterations through the extremely powerful, shared commitment to supporting pupil learning about early literacy (with different contributions from those involved). For most iterations, this was also the case for supporting teacher learning about technology and early literacy through participation in design teams.

Teacher communities: Knowledge Building

Knowledge building (Scardamalia & Bereiter, 2014) is an educational innovation that emphasizes working as a community to advance the state of knowledge in that community (typically a class). It differs in important ways from other approaches that the learning sciences have produced; although a knowledge-building experience always addresses external requirements such as educational standards, it does not come with a scoped and sequenced curriculum, instructions for teachers, and an imagined endpoint. Instead, the teacher in a knowledge-building classroom needs to learn to help students to: explore their knowledge and interests in a domain, find important questions in that domain to pursue, and to investigate and discuss these questions—and ideas about them—to advance the community’s state of knowledge as far as possible. There are principles and heuristics but few prescriptions.

To address the above challenge, we used a community-based approach to advance knowledge building locally by creating synergy between two types of activity: graduate teaching, and collaboration with teachers who are working on knowledge building in their classrooms (Chan & van Aalst, 2006; van Aalst & Chan, 2001). Thus when postgraduate students take our course on knowledge building, the “learning environment” includes teachers working on knowledge building in their classrooms—some already having 8 to 10 years of experience. The graduate students can observe the classes of those teachers and their online work or collaborate with them to try a new idea out in the classroom; some of the teachers also visit the graduate course. This strategy brings together the academic knowledge that graduate students acquire from their course, and which the teachers have limited time to acquire, with the practical knowledge of the teachers. Both the graduate students and teachers benefit from this interaction. In addition, we periodically organize events where graduate students, graduates and teachers come together to share their personal advances and learn about the latest research on knowledge building internationally. We also collaborate with the government to address problems that are important to the community at large (Chan, 2011). As some examples, very early we organized symposia in which graduate students, teachers, and some students from their classrooms worked together on important questions that were holding knowledge building back at the time—how the teacher introduces a question in the online environment. Later on, K-12 student panels shared about their experience with knowledge building. When we carried out a review of current theories of learning for the local government, the most experienced teachers became important collaborators, providing detailed examples of how their work exemplified the theories.

Our group has many of the hallmarks of a learning community (Bielaczyc & Collins, 1999), including shared goals, a sense of belonging to the community, ways of sharing accomplishments, and ways to regenerate membership. The course and symposia provide regeneration. One important signal that we have a community is when teachers who have relied on our assistance in the past become more self-directed and only ask for advice or limited assistance for their own initiatives. Another is that through the teachers own professional work we increase the footprint of knowledge building; for instance, one teacher became a member of the Curriculum Development Committee of the government; another hosted a group of teachers from Singapore who were on a study tour on knowledge building. While this community approach is not scalable to very large numbers of teachers it appears to be sustainable and makes an important contribution to developing scalable practices. Recently we have begun to establish formal partnerships with some of the schools where the teachers work to provide a larger role in professional development for those schools.

Which of the aforementioned elements are present in this approach? *Linkages*: There are a variety of short and long-term collaborations: between graduate students, between graduate students and teachers, among teachers, between researchers and teachers, and between the researchers, teachers and the government. *Structure*: We maintain websites to share resources for and on behalf of the teachers, create opportunities for collaboration, and opportunities for teachers to share their stories. Apart from symposia, we organize some social events that help to build community, and have recently begun to publish a newsletter. *Openness*: We believe that our model is one in which various participants have voice. Examples are when we have done research that responded to teacher questions. We have created opportunities for graduate students to carry out trials of knowledge building in their own classes, and thereby prepare them for continued work in this area after they graduate. Teachers and graduate students bring in knowledge of other technologies than Knowledge Forum® that enriches what we could offer via professional development. *Capacity*: We have benefitted from government initiatives that align with

knowledge building and related funding schemes; for example, a major curriculum reform in the first decade of this century emphasized working with ICT, collaborative learning, inquiry-based learning, and formative assessment, and knowledge building became a way to address all these goals. Currently, local emphasis on knowledge exchange (or valorization of research) stimulates our work. *Reward*: Locally, working with a university-based group to improve learning is desirable for schools. Opportunities to share work and be recognized for achievements also are seen as rewards. *Proximity*: The pedagogical requirements of knowledge building are novel—thus less familiar—but the current climate of educational change is an asset to our work. *Synergy*: Everything is knowledge building—our research, the work of graduate students, and that of teachers and their students.

Cross-case analysis

Table 1 provides an indication of the extent to which each feature is present in each case.

Across these three cases, we can observe both parallels and significant variation in terms of Havelock's procedural factors. Each, for example, relies heavily on diverse and varied collaborative relationships (linkage) within parameters established to define the numbers, types, and norms of community interactions (structure). These factors serve to cultivate a productive organizational environment in which high levels of synergy can be achieved. For example, a teacher community, because it has a large number of participants with varying expertise and interests, can have high degrees of linkage and synergy. Teacher communities also can require a great deal of structure, including the maintenance of technical infrastructure. Some communities such as large groups in social media, however, also can be effective at sharing knowledge but perhaps less so at knowledge construction that builds shared understanding of research results, agendas, and practical problems (Author, 2009). However, variations in individuals' belief and commitment to changes in organizational activity (openness) and/or the social/cultural capital associated with change through these activities (reward) can limit the capacity of research dissemination to impact change. Efforts within teacher communities may not be focused on a specific short-term target, and there may not be clear rewards apart from being recognized in the community for one's contributions. In non-professional contexts, SLI-focused activities have the potential to negatively impact one's standing within a community by challenging prevailing norms or established identities.

Table 1. Overview of features in each case

Feature		RDD	DB(I)R	Teacher Community
Linkage	Number and types of collaborative relationships	High—researchers, educators, policymakers, stakeholders	Variable—multi-year and multi-month partnerships between researchers and practitioners	High—short and long-term; researchers, teachers, government
Structure	Degree of systematic organization	High—grounding in theory-based program model for decision-making	High—organized by university team	High—organized by university team
Openness	Belief that change is desirable and possible	Variable—dependent on 'change-readiness' of target audiences	High—due to voluntary participation	Moderate—need to change partly accepted
Capacity	Ability to retrieve and marshal resources	Moderate—needs typically exceed capacity, particularly for personnel	Moderate—high in research team, high in those school teams that included educational leaders	High—if linked to a government change agenda
Reward	Frequency, immediacy, amount of positive reinforcements	Variable—dependent on cultural costs associated with target audiences	Variable—learning gains and teacher professional development generally positive but in varying degrees	Moderate—informal recognition of achievements
Proximity	Nearness in time, place and context	High—direct connections to expertise and lived experience of target audiences	High—direct connections to expertise and practices of participating teachers	Moderate—ongoing so without clear endpoint

Synergy	Number and variety of persistent forces that can be mobilized	Moderate—variable lasting impact on behavior beyond finite programming	High—Shared commitment to pupil learning, with different contributions from those involved	High—Participants contribute in variety of ways
----------------	---	--	--	---

Conclusion and implications

Enhancing the impact of research beyond academia—i.e., on policy-making, educational practice, and society, has never been more important. In an era in which society is much more interested in the return on investment of research funding than in previous decades, governments, research institutions, funding agencies, and the public at large are paying more attention to the benefits of research to society. One aspect of this is increased attention to research integrity (Macrina, 2014), but another is a call to the research enterprise to enhance its impact. For example, in the United Kingdom, the 2014 Research Excellence Framework—its external assessment of funded research—for the first time required “case studies” demonstrating the impact of research on policy-making (<http://www.ref.ac.uk/>). Indeed, across Europe, the last decade has witnessed increased attention to valorization. Here, valorization refers to the process of value-creation out of knowledge. Specifically, it is concerned with making new knowledge suitable and available for economic or societal use, usually by translating it into high-potential products, services, processes and industrial activity. These considerations seem especially applicable to the RDD modality. In thinking about the products, services, processes and industrial activity in the field of education, it would seem especially important to consider how to establish and maintain connections with the primary creators of broadly used curricula, assessments, and professional development opportunities, because such educational designers wield powerful influence on teaching and learning enactment (Author). These are key linkages and synergies underlying RDD efforts. However, even when these program- and organizational-level factors are high, characteristics of the population to be served are critical, including perceived need for and within-community incentives for change. Findings from learning sciences research can be directly translated into the design of educational programs and resources with potential for significant reach and impact when these factors are aligned effectively.

For DB(I)R and teacher communities, in which the research activities themselves simultaneously contribute to knowledge building among practitioners, Levin’s (2013) notion of knowledge mobilization may more accurately describe the vehicles at play, because it stresses the interactive, social and gradual nature of the bilateral connections between research and practice in the field of education. In DB(I)R and teacher communities, researchers and practitioners take each other seriously in genuine collaboration that yields both scientific and practical benefits.

From its inception, the field of the learning sciences has aimed to impact teaching, learning, and behavior in schools, universities and everyday contexts where learning happens. Despite some exceptions, the field has yet to make substantial progress in this direction. Furthermore, public understanding of what learning scientists know about how people learn remains extremely limited. As such, there remain significant opportunities for the field to tackle the problems of research, practice and public understanding of learning. Findings presented here begin to address these challenges. This paper makes a useful contribution by considering three common extant approaches to realizing research-practice connections. It describes various manifestations of the factors that are crucial for uptake and use of new knowledge (linkage, structure, openness, capacity, reward, proximity and synergy). By offering framework for analysis and initial findings based on three projects in the learning sciences, this paper makes a modest but clear contribution toward helping learning scientists conceptualize a range of research-practice connection modalities, and envision how they can be shaped in learning sciences projects. Subsequent work could further operationalize the factors for each modality, as those shaping research-practice connections seek not only conceptual tools, but also instruments that can help investigate and monitor the nature and functionality of specific research-practice connections.

References

- Bauer, K. & Fischer, F. (2007). The Educational Research-Practice Interface Revisited: A scripting perspective. *Educational Research and Evaluation*, 13(3), 221-236.
- Bielaczyc, K., & Collins, A. (1999). Learning communities in classrooms: A reconceptualization of educational practice. In C. M. Reigeluth (Ed.), *Instructional design theories and models, Vol II* (pp. 269-292). Mahwah, NJ: Lawrence Erlbaum Associates.
- Broekkamp, H. & Van Hout-Wolters, B. (2007) The gap between educational research and practice: A literature review, symposium and questionnaire, *Educational Research and Evaluation*, 13, 203–220.

- Burkhardt, H., & Schoenfeld, A. (2003). Improving educational research: Toward a more useful more influential and better-funded enterprise. *Educational Researcher*, 32(9), 3-14.
- Chan, C. K. K. (2011). Bridging research and practice: Implementing and sustaining knowledge building in Hong Kong classrooms. *International Journal of Computer-Supported Collaborative Learning*, 6, 147-186. doi: 10.1007/s11412-011-9121-0
- Chan, C. K. K., & Van Aalst, J. (2006). Teacher development through computer supported Knowledge Building: Experience from Hong Kong and Canadian teachers. *Teaching Education*, 17(1), 7-26.
- De Vries, B., & Pieters, J. (2007). Bridging the gap between research and practice: Exploring the role of knowledge communities in educational design. *European Educational Research Journal*, 6(4), 382-392.
- Design-Based Research Collective (DBRC) (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5-8.
- Hargreaves, D. H. (1999). The knowledge-creating school. *British Journal of Educational Studies*, 47(2), 122-144.
- Havelock, R. (1971). *Planning for innovation through dissemination and utilization of knowledge*. Ann Arbor: Center for Research on Utilization of Scientific Knowledge.
- Lavis, J., Robertson, D., Woodside, J., McLeod, C., Abelson, J., and the Knowledge Transfer Study Group. (2003). How can research organizations more effectively transfer research knowledge to decision makers? *The Milbank Quarterly*, 81(2), 221-248.
- Levin, B. (2004). Making research matter more. *Education Policy Analysis Archives*, 12(56), 1-22.
- Levin, B. (2013). To know is not enough: Research knowledge and its use. *Review of Education*, 1(1), 2-31.
- Macrina, F. L. (2014). *Scientific integrity: Text and cases in responsible conduct in research* (4th ed.). Washington, DC: ASM Press.
- Ormel, B., Pareja Roblin, N., McKenney, S., Voogt, J., & Pieters, J. (2012). Research-practice interactions as reported in recent design studies: Still promising, still hazy. *Educational Technology Research & Development*, 60(6), 967-986.
- Pareja Roblin, N. N., Ormel, B. J., McKenney, S. E., Voogt, J. M., & Pieters, J. M. (2014). Linking research and practice through teacher communities: a place where formal and practical knowledge meet? *European journal of teacher education*, 37(2), 183-203.
- Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, 40(7), 331-337.
- Rogers, E. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.
- Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 397-417). New York: Cambridge University Press.
- Thomas, G. & Pring, R. (2004). *Evidence-based practice in education*. Maidenhead: Open University Press.
- Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal*, 36(2), 299-316.
- Van Aalst, J., & Chan, C. K. K. (2001). "Beyond sitting next to each other:" A design experiment on knowledge building in teacher education. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *European perspectives on computer-supported collaborative learning: Proceedings of the first European Conference on Computer-Supported Collaborative Learning* (pp. 20-28). Maastricht, the Netherlands: ISLS.
- Voogt, J., McKenney, S., Pareja Roblin, N., Ormel, B., & Pieters, J. (2012). De R&D functie in het onderwijs: Drie modellen voor kennisbenutting en -productie. *Pedagogische Studiën*, 89(6), 335-337.
- Yin, R. (2014). *Case study research: Design and methods*. London: Sage.

Acknowledgments

This paper results from and builds on work from a pre-conference workshop held at ICLS 2014 (reference removed for blind review). The authors express their gratitude to Wil Doane, Kim Gomez, and Brian Reiser for helping establish the cooperation through which this paper came to fruition.

Design Collaborative Formative Assessment for Sustained Knowledge Building Using Idea Thread Mapper

Jingping Chen, University at Albany and University of Electronic Science and Technology of China,
chgping@gmail.com

Jianwei Zhang, University at Albany, jzhang1@albany.edu

Abstract: This design-based study investigated collaborative formative assessment for knowledge building in two comparable Grade 6 science classrooms. Students assessed their collective knowledge progress using the Idea Thread Mapper (ITM)—a timeline-based collective discourse mapping tool—and planned for further efforts to address deeper issues. Analysis of students' online discourse and individual portfolio notes suggests the positive impact of the assessment on the community's knowledge building discourse as reflected in students' idea-deepening and idea-elaborating questions, refined explanations, and build-on connections; and on student scientific understandings as documented in their individual portfolio notes.

Introduction

Designing new assessments in line with evolving conceptions of how people learn is a central challenge in the learning sciences (Pellegrino, 2014). Research on collaborative learning and knowledge building has demonstrated new learning designs and classroom practices to develop deep knowledge and high-order competencies, including knowledge-creating capabilities (Scardamalia & Bereiter, 2006; Stahl, 2006). The pedagogy of collaborative knowledge building requires new learning assessments to measure both individual and collective knowledge advances and feedback on students' inquiry and collaboration as the process unfolds. The purpose of this study is to design and examine formative, collaborative assessment in the context of knowledge building communities. In a knowledge building community, members continually advance the state-of-the-art understanding of core issues in a focal area(s) through interactive idea input. They build collective knowledge as a social product of their community beyond individual notions and concepts (Scardamalia & Bereiter, 2006). The community's collective knowledge is represented through the conceptual artifacts developed and shared by the community members as they engage in idea-transforming discourse. Using collaborative online platforms such as Knowledge Forum, students contribute and build on one another's ideas over time to address deepening issues and develop increasingly sophisticated understandings. Deeper challenges are identified as progress is made, leading to sustained cycles of inquiry and progressive discourse (Hakkarainen, 2003; Zhang et al., 2007). Existing assessments primarily focus on individual learning processes and outcomes, and capture very little of the sustained, collaborative knowledge building processes that continually deepen and evolve. Therefore, designing new assessment for sustained, collaborative knowledge building becomes a critical challenge (Scardamalia et al, 2010; van Aalst & Chan, 2007).

The current study explores the design of assessment for sustained knowledge building in light of a set of key principles derived from the literature. These include:

(a) Student-directed assessment with high-level collective responsibility: Sustained knowledge building requires students to take on collective responsibility for high-level decisions, including setting goals, long range planning, and progress tracking (Scardamalia, 2002). Assessment for knowledge building hence needs to be student-directed (van Aalst & Chan, 2007). Student make collective decisions about what should be investigated and assessed, based on what evidence, and what further actions should be carried out based on the results of the assessment. They are active agents for the internal assessments of their own work instead of passive test-takers in the external assessments (Scardamalia & Bereiter, 2006).

(b) Collaborative assessment of collective knowledge in relation to individual learning: Assessment for knowledge building needs to capture both collective and personal knowledge advancement (Scardamalia et al, 2010; van Aalst & Chan, 2007). A community's progress and practices in collective knowledge advancements are evident through its trajectories and patterns of knowledge building discourse (e.g. deepening ideas and questions) and knowledge artifacts generated. Students' individual knowledge advancements are reflected in their contributions to the community's discourse as well as their personal artifacts generated to support their own learning and reflection.

(c) Formative and transformative assessment to inform sustained idea improvement: Beyond existing assessments that characterize student' performance and progress in past learning (Mislevy & Haertel, 2006), assessment for knowledge building needs to provide ongoing feedback to support sustained idea improvement

and collective engagement (Scardamalia & Bereiter, 2006) and scaffold future learning and inquiry (Schwartz & Arena, 2013; Shepard, 2000). The collection and use of assessment information becomes an integral part of the ongoing learning process (Pellegrino, 2014).

(d) Technology-supported assessment using analytic tools: Assessment for knowledge building based on ongoing discourse and student artifacts requires technology support to analyze rich data and provide easily interpretable results and visualizations. Various analytic tools have been developed to capture cognitive and social dynamics of collaborative learning and knowledge building using semantic, lexical, and social network analysis (Scardamalia et al., 2010; Zhang et al., 2009). However, these new assessment and analysis tools mostly remain as research tools; they are often too complicated for students to use and interpret (Zhang & Chen, 2012).

Research by van Aalst, Chan and colleagues (Lee, Chan, & van Aalst, 2006; van Aalst & Chan, 2007) showed the positive impact of student-directed, formative assessment on collaborative knowledge building. In their research students created e-portfolios to identify productive examples of knowledge building contributions and discourse episodes. However, with the lack of effective means to making collective knowledge progress visible in current online discourse environments (Zhang, 2009), it is cognitively challenging for students to monitor and assess collective knowledge progress based on distributed discourse in long-term inquiry. Guided by the above-mentioned principles, this research tests a design of collaborative formative assessment for knowledge building supported by a timeline-based collective discourse mapping tool: the Idea Thread Mapper (ITM) (Chen, Zhang, & Lee, 2013). ITM interoperates with Knowledge Forum (Scardamalia & Bereiter, 2006) and potentially other collaborative learning platforms. In these online environments, student ideas are presented in distributed postings (e.g. notes) and responses (build-ons) in extended online discourse. To help students monitor what is going on in the discourse and interpret the collective focuses and progress, ITM incorporates conceptual threads of inquiry--“idea threads” (Zhang et al., 2007)--as a larger, emergent unit of ideas in online discourse. ITM allows students to create idea threads through selecting certain discourse contributions (notes) for various focal objects of inquiry. Each idea thread is composed of a sequence of discourse entries (possibly several build-on trees) contributed by a subset of the members of a community to address a shared problem or conceptual topic. The collective knowledge of the community in a whole inquiry-based initiative is further represented as clusters of idea threads that address interrelated problems. With the authors and build-on connections identified, the idea threads are displayed on a timeline as an “idea thread map” (left side of Figure 1). The progress in each idea thread is further made transparent by students through co-authoring a “Journey of Thinking” synthesis that includes three sections: *We want to understand*, *We used to think...and we now understand...*, *We need to do more* (right side of Figure 1). Idea threads and thread-based syntheses are co-editable by members of the community, with each version recorded for later review. New analytics tools in ITM support auto-clustering of notes based on thread topics and analysis of discourse contribution types (e.g. questioning, explaining) in each idea thread to support student reflection.



Figure 1. A map of idea threads (left) and a “Journey of Thinking” synthesis (right) created by one of the Grade 6 classrooms in this research studying biodiversity. Each colored stripe represents an idea thread extending from the first till the last note contributed addressing a shared focal problem. Each square represents a note. A dotted vertical line shows notes shared between different threads discussing interrelated issues. The “Journey of Thinking” page (right) shows students’ co-authored reflection on what we need to understand, “big ideas” learned, and what we need to do in the next step.

Existing research suggests that ITM can support collective meta-discourse and reflection among students: to review ongoing discourse contributions to formulate shared focuses and goals of inquiry, monitor how ideas have been advanced in unfolding lines of inquiry, synthesize insights, idea connections and deeper actions to be taken by community members (Zhang et al., 2015). The current research further extends the ITM-aided discourse review into a systematic design of collaborative assessment to leverage sustained knowledge building. In light of the related literature (Scardamalia et al., 2010; van Aalst & Chan, 2007; Zhang et al., 2009), the collaborative assessment design focuses on three essential aspects of knowledge building: collective knowledge, evidenced through sustained and progressive discourse contributions to advancing various lines of work in the community; social dynamics of collaboration, evidenced through the active participation, idea connection, and distributed engagement of the members; and student individual understanding, revealed through their personal reflection on what they have learned. Students reflectively assess each aspect of knowledge building focusing on three questions underpinning formative assessment (Pellegrino, 2014; William & Thompson, 2007): (a) to define where we need to go; (b) to reflect on where we are now; and (c) to reflect on how we will go there by making productive moves.

To be specific, in this study students used ITM to review their discourse to identify collective themes and goals of knowledge building on the basis of the diverse range of questions and interests represented in their discourse and work. They also synthesized the “big ideas” learned and gaps using ITM’s Journey of Thinking feature. To assess their social dynamics, they reviewed their participation in each idea thread and build-on connections as reflected in the social network graphs. As an extension of the collaborative assessment, each student further self-assessed his/her individual knowledge development through writing portfolio notes that summarize what has been learned from the whole community’s work. By examining the idea thread maps, social network graphs, and portfolio notes, students identified strengths and advances of their work as well as weak areas and potential connections to be addressed in future inquiry. Based on discussions of these findings, they constructed plans to deepen their collective inquiry, refine collaboration, and increase personal contribution.

This design-based research was conducted to test and refine the above assessment design in two elementary classrooms. The process of the assessment, including the role of the teacher, was analyzed using video analysis and reported in Chen (2015). The current paper focuses on the role of the assessment in sustaining knowledge building. Our research question asks: In what ways does the collaborative, formative assessment support collaborative deepening moves in the community’s online discourse and advancement of student personal understanding?

Method

Classroom contexts

This study was conducted in two comparable Grade 6 classrooms at Zongbei Elementary School in Chengdu, China. The two classrooms had been implementing knowledge building pedagogy using Knowledge Forum (Scardamalia & Bereiter, 2006) for two years. Students in each classroom (39 students in class A, and 42 in B) studied two science units—energy and biodiversity—over a four-month period. The two classrooms were taught by the same science teacher in cooperation with an ICT (Information and Communications Technologies) teacher who supported the online discussions.

Research design

This design-based study (Collins, Joseph, & Bielaczyc, 2004) adopted a two-phase, time-lag design. In phase/Unit 1 focusing on energy, only class A conducted ITM-aided collaborative assessment; in phase/Unit 2 focusing on biodiversity, the assessment was refined and extended to both classrooms. The impact of collaborative assessment in the knowledge building practice was examined by comparing Class A to B in phase 1 and comparing Class B’ performance between the two phases.

The inquiry in each unit lasted for eight weeks. Each week they had four 40-minute lessons, typically two focusing on face-to-face activities and two for online discussions. The assessment was first implemented in Week 4. The community engaged in the following activities to assess its collective knowledge advancement:

(a) Setting focus: The community members reviewed their ongoing discussions to identify what they need to understand, represented as shared, high-interest themes of inquiry. The students first generated a “theme list” in their notebooks, and then proposed their themes to the community. Through a whole class discussion the community reviewed the themes and their connections, and co-created a list of shared themes of inquiry, as the assessment focus.

(b) Collecting evidence: Based on the shared focal themes of inquiry, the students formed into groups each of which used ITM to identify important Knowledge Forum notes for each focal theme of inquiry, as an idea

thread. The idea threads were further co-refined through group discussions to remove the unrelated notes, add more notes, highlight key notes. Through reviewing the notes in each idea thread, each group co-authored a “Journey of Thinking” synthesis to highlight the important ideas and questions.

(c) Generating feedback: With the map of idea threads projected on a screen, the community collaboratively interpreted the processes and progress of collective knowledge advancement, using intensive discourse contributions, connections, highlighted notes, and “Journey of Thinking” syntheses in the different idea threads as indicators of collective advancements. Through the collective review, students identified important insights gained as well as gaps and challenges to be addressed in each line of inquiry and connections to be built across the inquiry themes.

(d) Planning: Based on the feedback, the community then made two types of plans to deepen their knowledge building work: in groups they made plans for deeper inquiry in their focal idea threads, and as a whole class they generated plans and suggestions for their whole inquiry initiative to guide the work of all members.

Extending the reflection on collective knowledge, the community reflected on their ways of collaboration with the social network analysis tool that shows who had read and who had built onto whose notes. Through interpreting the social network graphs, they reflected on possible ways to improve their participation and social connections. Informed by the collaborative assessment, each student further wrote a portfolio note to reflect on what he/she had learned about the various inquiry themes of the community and how they would better contribute to the collective knowledge work and improve their own learning.

Guided by the collaborative and individual plans, students conducted further inquiry for two more weeks. Near the end of the unit, the students conducted another round of assessment to update their idea threads (e.g. adding new Knowledge Forum notes), revisit the idea thread maps, and update their individual portfolio notes. The assessment design was refined based on its implementation and data analysis in Unit 1. The major changes were to create a more connected and smooth flow of the assessment activities, help students understand the purposes of the activities through discussions, and integrated the assessment information (e.g. social network graphs and ITM maps).

Data sources and analyses

The data sources included online discourse, personal portfolio notes, social network data of the online discussions, and the graphical and textual representations on ITM including idea threads, idea thread maps, and Journey of Thinking syntheses.

Content analysis (Chi, 1997) was conducted to gauge the quality of the knowledge building discourse before and after the assessment. Knowledge building discourse is characterized by progressive moves to identify deepening problems and develop increasingly sophisticated explanations (Hakkarainen, 2003; Scardamalia & Bereiter, 2006; van Aalst, 2009). Based on our prior studies (Zhang et al., 2007; 2009), our coding scheme first identified initial notes that posted questions, and classified the questions as basic fact-seeking questions (e.g. what) vs. deeper explanation-seeking questions (why and how); idea-initiating wonderments vs. idea-deepening questions and idea-clarifying questions. We then coded student build-on notes to examine students’ interactive input based on theorizing and explaining (T), using evidence (E), referencing sources (R), connecting and integrating (C), and designing and applying (D). Theorizing/explaining included five sub-categories: intuitive explanations (T1i), alternative explanations (T1a), refined explanations (T1r), clarifying explanations (T1c), and suggestions (T1s) (see the coding scheme at <http://tccl.rit.albany.edu/papers/Coding-scheme.pdf>). The density of the build-on interaction (i.e. who had built on whose notes) was further analyzed using social network analysis.

Students’ individual portfolio notes were coded by two independent coders through content analysis focusing on the scope of knowledge, depth of understanding, and connectedness of ideas. To examine the scope of student personal understanding, each portfolio note was coded based on the list of the themes covered by the community’s online discussions. Student ideas related to each theme were coded for depth of understanding using two well tested four-point scales (Zhang et al., 2007; 2009): scientific sophistication (1=pre-scientific, 2=hybrid, 3=basically scientific, to 4= scientific) and epistemic complexity (1=unelaborated facts, 2=elaborated facts, 3=unelaborated explanations, and 4=elaborated explanations). Student ideas related to each theme were further coded based on connected/coherent vs. unconnected/scattered. A satisfactory inter-rater reliability was achieved: Cohen’s Kappa = .84 for the coding of themes, .83 for scientificness, .82 for complexity, and .89 for idea connectedness.

Results

Content analyses of the knowledge building discourse

To investigate the role of the collaborative formative assessment in sustaining collaborative knowledge building, our analyses examined the changes in the knowledge building discourse before and after the assessment between the two classrooms based on questions raised and build-on moves to generate and improve ideas.

Questioning moves in the online discourse

Table 1 reports the distribution of the different types of questions in the online discourse of the classroom A and B before and after the assessment activities. In Unit 1 for the energy study, only class A implemented the assessment design. In Unit 2 for the biodiversity study, both classrooms implemented a refined version of the assessment design.

Table 1: The frequencies and percentages of different types of questions raised in the discourse of the two communities before and after the assessment in the two units of study.

		Fact-seeking	Explanation-seeking	Idea-initiating	Idea-deepening	Idea-clarifying
Class A Unit 1 (ITM assessment)	Before	30	4	5	3	3
		66.67%	8.89%	11.11%	6.67%	6.67%
	After	2	5	3	10	3
		8%	20%	12%	40%	12%
Class B Unit 1 (no ITM assessment)	Before class A's assessment	32	2	3	7	15
		52.46%	3.28%	4.92%	11.48%	24.59%
	After class A's assessment	5	3	0	2	7
		26.32%	15.79%	0%	10.53%	36.84%
Class A Unit 2 (refined ITM assessment)	Before	25	14	3	7	6
		45.45%	25.45%	5.45%	12.73%	10.91%
	After	3	8	3	9	3
		10.34%	27.59%	10.34%	31.03%	10.34%
Class B Unit 2 (refined ITM assessment)	Before	28	6	9	2	3
		58.33%	12.5%	18.75%	4.17%	6.25%
	After	5	6	5	8	8
		15.63%	18.73%	15.63%	25%	25%*

Note. The percentages of some contribution types add up to over 100% because each note may be coded for multiple categories.

In Unit 1 before the assessment, fact-seeking questions (e.g. what, how many) dominated the discourse in both classrooms (66.67% for A and 52.46% for B), with very few higher-level questions such as explanation-seeking questions. A Chi-Square test shows no significant difference between the two classes ($p > .05$). After class A's assessment, class A generated more explanation-seeking (20%) and idea-deepening questions (40%) to explain reasons, relationships, specific mechanisms while class B generated more idea-clarifying questions (36.84%) and fact-seeking questions (26.32%) to search for and clarify information ($\chi^2(12) = 132.76, p < .05$).

In the Unit 2, raising factual questions was the dominant pattern of discourse in both classes before the assessment (45.45% for A and 58.33% for B). After the assessment, both classroom had a drop in factual questions and an increase in explanation-seeking and idea-deepening questions that are essential to sustained, progressive discourse for knowledge building.

Patterns of build-on notes

The network density of who had built on whose notes increased over time for each classroom in each unit of study. In Unit 1, a greater increase in the density was observed in class A (15.38%) after its assessment than in classroom B (10.19%) that did not use the ITM-aided assessment. In Unit 2 when both classrooms conducted a refined version of the ITM-aided assessment, both classes had a higher increase in the density (Class A: 17.29%; Class B: 18.54%).

Figure 1 shows the patterns of the build-on notes to generate and improve ideas in the online discourse before and after the assessment in the two classrooms. In Unit 1, before the assessment of class A, a majority of students' build-on notes in both classes contributed intuitive explanations based on personal experience (58.11%

for A and 41.58% for B) and referred to sources of information from readings (24.32% for A and 29.21% for B). After the assessment, class A contributed less intuitive explanations (decreased from 58.11% to 36%) and more refined/sophisticated explanations (increased from 5.41% to 20%). In contrast, class B brought in even more intuitive explanations (increased from 41.58% to 62%) with few notes contributing refined explanations (4.95%).

In Unit 2, before the assessment a majority of students' build-on notes in both classes were also intuitive explanations (55.46% for class A and 46% for class B), followed by referencing sources of information (26.05% for class A and 34% for class B). After the assessment students in both classes contributed less intuitive explanations (38.98% for class A and 35.71% for class B) and referencing sources (16.95% for A and 22.32% for B). The percentage of refined/sophisticated explanations increased for both class A (from 8.4% to 12.71%) and B (from 2.67% to 16.07%). The changes after the ITM-aided assessment in Unit 2 in both classrooms are consistent with the changes observed in Unit 1 for class A after its implementation of the assessment.

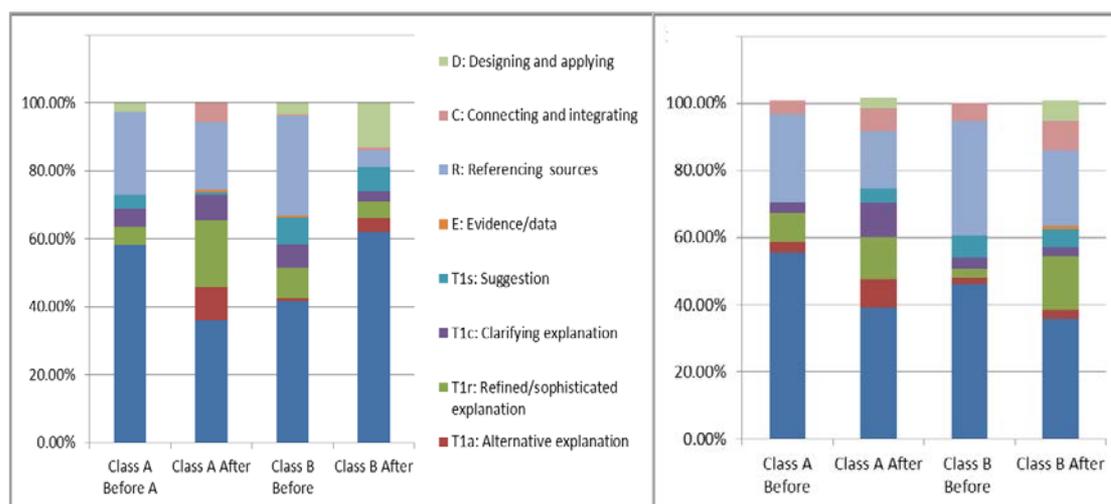


Figure 2. Patterns of the build-on notes before and after the assessment in the two classes in the two units: energy (left) and biodiversity (right).

Content analysis of individual portfolio notes

Table 2 reports the content analysis of student individual portfolio notes based on the themes of inquiry covered, depth of understanding about each theme as gauged based on the sum of two ratings, scientificness (1-4) and epistemic complexity (1-4), and percentage of inquiry themes with connected and coherent ideas.

Table 2: Evaluation of Personal Knowledge Advances Summarized in Student Portfolio Notes

	Class A			Class B		
	Themes covered M (SD)	Depth of understanding M (SD)	Connectedness of ideas M (SD)	Themes covered M (SD)	Depth of understanding M (SD)	Connectedness of ideas M (SD)
Unit 1: Energy	4.83 (1.84)	5.4 (1.89)	9.34% (18.22%)	4.26(2.11)	4.74 (1.76)	13.24 % (22.53%)
Unit 2: Biodiversity	5.11(1.57)	6.6 (1.44)	28.31% (26.70%)	5.41(1.97)	5.26 (1.7)	25.80% (27.99%)

Focusing on each of the three indicators, a two-way repeated measure ANOVA was conducted to compare the quality of the portfolio notes from the two classrooms in the two units of study. A significant between-class difference was found in depth of understanding ($F(1, 34) = 11.44, p < .05$) with class A outperforming B. The between-unit difference is significant in each of the three indicators, for the number of inquiry themes covered ($F(1,34) = 7.3, p < .05$), for the depth of understanding ($F(1, 34) = 23.49, p < .05$), and for connectedness of ideas $F(1, 34) = 18.67, p < .05$. The refinement of the collaborative assessment in Unit 2 led to improved results.

Discussion

This study tested a collaborative, formative assessment for knowledge building, which focuses on three major constructs: collective knowledge advancement, social dynamics, and individual understanding. Students played an active and collaborative role in each step of the formative assessment supported by their teacher: to identify shared focuses and deepening goals of inquiry as their discourse evolved; to collect and interpret evidence of idea contributions, advancement, and collaboration; and to generate feedback and guidance about how to address their deeper needs and gaps of knowledge through further inquiry. The ITM tool served to represent their collective focuses of inquiry, make the collective progress in various lines of inquiry transparent for collective review, and support student efforts to reflect on their Journey of Thinking in terms of where they were going, where they were now, and what deeper efforts were needed. As the results indicated, this collaborative, formative assessment played an important, positive role to improve students' collective, progressive discourse as well as individual understandings. Specifically, engaging in the collaborative formative assessment using ITM helped students to engage in deeper knowledge building discourse through generating more explanation-seeking and idea-deepening questions. The ITM-aided assessment on idea progress in each line of inquiry, including summarizing focal problems, "big ideas" learned and deeper issues, served to catalyze student intentional efforts of progressive questioning. In response to the deepening questions and ideas from their peers, students' build-on notes contributed more refined and sophisticated explanations after the formative assessment.

The positive association between the assessment and the enhanced discourse quality should largely be attributed to the purposeful design of the collaborative formative assessment: to enhance student monitoring of collective knowledge and generate ongoing feedback to guide sustained idea improvement (Zhang et al., 2009, Zhang et al., 2013). Through the assessment aided by ITM's visualization and analytic tools, students made more informed reflection on which lines of ideas needed to be improved through what deeper actions and contributions. Small groups then planned on how to further advance each idea thread, and the class, as a whole, discussed collective efforts to improve the whole inquiry. Supported by the collaborative reflection, individual students made informed decisions about how to connect their individual interest, strength and resources with the community's needs.

The results also revealed a positive role played by the ITM-aided collaborative assessment to enhance students' personal knowledge gains as they advanced their collective knowledge. Class A with the collaborative formative assessment in Unit 1 demonstrated deeper understandings in student portfolio notes than class B. Significant improvements were observed from Unit 1 to 2. In Unit 2 that implemented a refined design of the assessment, students' portfolio notes addressed a broader range of inquiry themes that were explained using more scientific, complex, and connected ideas. The assessment helped students to have a reflective awareness of the important inquiry themes and idea advancements across their community's knowledge space, understand discourse contributions and connections in a temporal context, and connect ideas to generate coherent understandings. These results suggest that the refinement of the assessment in Unit 2 to support a more coherent flow of the assessment activities and integrated view of the assessment information helped to increase the effectiveness of the assessment. Of course, this two-phase research design cannot rule out the possible impact of the content topics (e.g. energy vs. biodiversity) on the complexity of student portfolio texts, although both topics are open-ended and relevant to student interest.

In conclusion, this research tested a collaborative formative assessment for knowledge building communities, with ITM mapping out the community's online discourse as unfolding conceptual trajectories of inquiry. In line with existing research on student-directed, technology-supported assessment for knowledge building (van Aalst, J., & Chan, 2007), this study found a positive role of the assessment in enhancing the progressive discourse of the community with deepening questions and refined explanations while supporting individual efforts to generate deeper and connected understandings. A condition to increase the effectiveness of such assessment is to integrate the assessment process coherently with the knowledge building process without causing much additional work among students. We are conducting research to refine the design of this assessment to integrate the activities and assessment information with an easy flow, supported by automated analyses in ITM that can assist students to cluster discourse contributions based on inquiry topics and review discourse contributions based on different types. Further research also needs to test this assessment design in other classroom contexts and content areas to elaborate the design and examine its impacts.

References

- Chen, J. (2015). *Formative assessment for collaborative knowledge building in Chinese elementary science classrooms*. Doctoral dissertation submitted to the University at Albany. Retrieved from ProQuest Dissertations and Theses (Accession Order No. AAT 3722030).

- Chen, M.-H., Zhang, J. & Lee, J. (2013). Making collective progress visible for sustained knowledge building. In N. Rummel, M., Kapur, M. Nathan, & S. Puntambekar (Eds.), *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings Volume 1* (pp.81-88). Madison, WI: International Society of the Learning Sciences.
- Chi, M. T. H. (1997). Quantifying qualitative analysis of verbal data: A practical guide. *Journal of the Learning Sciences*, 6, 271–315.
- Collins, A., Joseph, D. & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences*, 13(1), 15–42.
- Hakkarainen, K. (2003). Progressive inquiry in a computer-supported biology class. *Journal of Research in Science Teaching*, 40(10), 1072–1088. doi:10.1002/tea.10121
- Lee, E.Y.C., Chan, C.K.K., & van Aalst, J. (2006). Students assessing their own collaborative knowledge building. *International Journal for Computer-Supported Collaborative Learning*, 1, 277-307.
- Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centred design for educational testing. *Educational Measurement: Issues and Practice*, 25(4), 6-20.
- Pellegrino, J. (2014). A learning sciences perspective on the design and use of assessment in education. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (2nd edition) (pp. 262-282). New York, NY: Cambridge University Press.
- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 97-115). New York, NY: Cambridge University Press.
- Scardamalia, M., & Bereiter, C. (2010). A Brief History of Knowledge Building. *Canadian Journal of Learning and Technology/La Revue Canadienne de l'Apprentissage et de la Technologie* [online], 36(1). Retrieved from <http://www.cjlt.ca/index.php/cjlt/article/view/574/276>
- Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29(7), 4–14.
- Stahl, G. (2006). *Group cognition*. Cambridge, MA: MIT Press.
- van Aalst, J. (2009). Distinguishing knowledge sharing, knowledge construction, and knowledge creation discourses. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 259-287.
- van Aalst, J., & Chan, C. K. K. (2007). Student directed assessment of knowledge building using electronic portfolios in Knowledge Forum. *Journal of the Learning Sciences*, 16, 175–220.
- Wiliam, D. & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C.A. Dwyer (Ed.) *The future of assessment: Shaping teaching and learning*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Zhang, J., & Chen, J. (2012). Building knowledge through the cyber space. In: Z. Yan (Ed.), *Encyclopedia of cyber behavior* (pp. 383-393). Hershey PA: IGI Global. (I, R)
- Zhang, J., Chen, M.-H., Tao, D., Lee, J. Sun, Y., & Judson, D. (2015). Fostering sustained knowledge building through metadiscourse aided by the Idea Thread Mapper. In O. Lindwall & S. Ludvigsen (Eds.), *Exploring the material conditions of learning: Proceedings of the 11th International Conference on Computer Supported Collaborative Learning* (Vol. 1, pp. 166-173). Gothenburg, Sweden: International Society of the Learning Sciences.
- Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of nine- and ten-year-olds. *Educational Technology Research and Development*, 55(2), 117–145.
- Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge building communities. *Journal of the Learning Sciences*, 18(1), 7–44.

Acknowledgments

This research was sponsored by the National Science Foundation (IIS #1122573, IIS #1441479). We would like to thank Dr. Heidi Andrade and Dr. Alandeom Oliveira for their advisory input. We owe special thanks to the participating teachers and students for their classroom work enabling this research; and to the members of our team who contributed to the design and development of Idea Thread Mapper.

“I Think We Were Pretty Powerful This Summer as Scientists”: Generating New Possibilities for Youth of Color in Science

Tammie Visintainer, TERC, tammievis@gmail.com

Abstract. This research explores processes of learning and identity construction for high school students of color as they engage in a community-based scientific research project as part of a summer science program. This research utilizes qualitative (e.g. interviews) and quantitative (e.g. surveys) data sources. Findings illustrate the emergence of an identity generative process: that engaging in science practices (e.g. presenting research) and the accompanying program resources (e.g. placing students in a position of power) generated new possibilities for students as capable science learners, scientists, and change agents. Furthermore, findings show that the instructor’s perspective of science and vision for his students guided the design of program resources (e.g. pedagogy, instruction) made available that helped create these shifts. Findings show that youth of color can imagine new possibilities for who they can be in science if their science and racial identities are supported in science programs.

Keywords: identity, science practices, youth of color, community-based research, instructor vision

Introduction

The underrepresentation of high school students of color in advanced science courses and the need to increase diversity in science fields is widely agreed upon (Oakes, Ormseth, Bell & Camp 1990; Darling-Hammond, 2010). In addition, developing the next generation of innovators in science, technology, engineering, and math (STEM) is deemed essential because it is linked to the nation’s economic prosperity (National Research Council, 2011). This highlights two issues: a need to increase diversity and foster innovation in science fields.

The challenge of increasing racial/ethnic diversity in science is complicated by two dominant representations: 1) *what* science is (i.e. what gets counted as science practices), and 2) *who* does/can do science. These representations too often provide narrow images of science (e.g. prescribed labs) and scientists (e.g. white, male) that are problematic for youth of color who don’t see themselves within the bounds of these narrowly defined categories. A typical solution to increasing diversity is to focus on reframing *what* science is through curricular and policy reforms (e.g. National Research Council, 2011). While this is an important part of the solution, race-based achievements gaps are persistent (National Science Board, 2012). This pushes us to consider how we think about equity; as not only access to authentic science experiences but also opportunities to see oneself as a capable science learner. I argue that in order to make progress towards increasing diversity in the sciences, we need to extend beyond a focus on *what* science is and pay equal attention to the messages youth of color receive about *who* they can become in science based on the resources made available in science learning environments.

Objective

This research takes a holistic approach to understanding processes of learning and identity construction for high school students of color as they engage in a community-based scientific research project as part of a summer science program. I explore how the possibilities high school students of color view as available to them in science are reimagined and transformed through program participation. The goal of this research is to better understand the mechanisms that support the generation of new possibilities: how engaging in science practices and the programs resources made available for the practices (e.g. instruction, pedagogy, designed experiences) generates new possibilities for *who* youth of color can become in science. To do this, I explore how the science instructor’s perspective of science (e.g. purpose, utility) and his students (e.g. social positioning) and the vision he has for how students can use science in their lives guided the design of the program resources made available while engaging youth in scientific research. In addition, I document the shifts that occur for students and how the program resources made available support holistic identity construction (i.e. students’ science and racial identities) and allow students’ agency and power to come to life.

The aspirational thinking expressed by two scholars encompasses what it means for educators to have a vision specifically in relation to working with youth of color from historically marginalized populations. In an interview broadcast on a local radio program, Dr. Cornel West encouraged teachers from schools serving predominantly youth of color to develop opportunities for their students to become creators. In this way, students are empowered as agents who are able to create change in their lives. Along similar lines, during a presentation to undergraduates at the University of California Berkeley, Dr. Kris Gutierrez encouraged aspiring educators to

approach learning with new social and pedagogical imaginations, to promote what is possible, and to provide tools for youth of color to become social dreamers and designers of their own futures (Gutierrez, 2008). Both of these scholars speak of providing tools that allow students to hope, dream, design and create change. This research builds on this visionary thinking to illustrate what is possible in a science program when there are opportunities to generate hope and possibility and when an instructor's vision shapes program resources that position students as doers of science and agents of change in their community. I ask three main questions:

- What shifts for high school students of color in regards to how they see themselves in relation to science following program participation? What are the mechanisms involved?
- How do instructors' perspectives of science (e.g. purpose, utility) and their students (e.g. social positioning) guide the design of program resources (i.e. instruction, pedagogy and designed experiences) made available while engaging youth in scientific research?
- What types of program resources support the co-development of students' science and racial identities and generate new possibilities for *who* they can become in science?

Theoretical framework

This research draws on sociocultural theories of learning and identity (Wenger, 1998). I view identities as fluid, dynamic and socially constructed from the resources available (Holland et al., 1998) as students are positioned in science learning environments (Harre, 2008). This research views learning as not only taking on new knowledge structures but guiding personal transformations of "becoming" (Nasir & Hand, 2006). I define *identity* as how youth of color perceive science, how they see themselves in relation to science, and how the interactions they have and believe they can have shape *who* they become in science. This involves their ideas about *who* does/can do science and their ideas about *what* scientist do.

Scholars examining learning and identity construction provide important insights into the types of identities offered and authored by youth of color in science learning environments across formal and informal contexts (e.g. Aschbacher et al., 2010; Calabrese Barton, Tan & Rivet, 2008; Polman & Miller, 2010). In addition, research shows how engaging youth in science practices and scientific investigations supports processes of learning and identity construction and provides unique opportunities for students to: develop deep understanding of content (Lehrer, 2009), construct unique epistemological stances (Sandoval, 2005), perceive themselves as real scientists and identify as capable learners (e.g. Bowen & Roth, 2007). As one type of investigation, research shows that community-based research can empower youth as change agents (e.g. Bouillion & Gomez, 2001). I define *science practices* as types of activities utilized to conduct scientific research and/or produce scientific knowledge through the research process.

While research shows that racial positioning matters for the long-term trajectories of women of color in science (Carlone & Johnson, 2007; Johnson et al, 2011, Malone & Barabino, 2008), treatments of race in relation to science are limited (Parsons, 2014). More research is needed that examines intersections of students' racial and disciplinary identities (Varelas, Martin, Kane, 2012). In addition, research is needed that offers insights into how students' science and racial identities *develop* together through participation in science practices.

Conceptualizing program resources

This research explores how the types of program resources made available in science learning environments supports youth of colors' practice-linked identities in science (i.e. viewing participation in a practice as central to who one is) (Nasir & Cooks, 2009; Nasir 2012). I define *program resources* as the types of instruction, pedagogy and designed experiences that instructors make available to students in science learning environments as they engage in science practices. I utilize the "practice-linked identity" framework and its supporting identity constructing resources (Nasir & Cooks, 2009; Nasir, 2012) in order to explore how learning, positioning and engagement in science develop together in context as students participate in the summer science research program.

Conceptualizing pedagogical vision

As one source of power in science classrooms, teachers have the ability to offer alternative possibilities and roles for their students through the organization of institutional structures, cultural practices, social interactions, and relationships (Nasir, 2004). This research explores how an instructor's "pedagogical vision" guides the organization and structuring of activities in a science program in ways that empower youth of color as doers of science and agents of change in their community. I define "*pedagogical vision*" as the ways that teachers' own backgrounds and experiences inform their goals and purposes for teaching science, what they envision their students doing with science, and the possibilities they create for *who* their students can become in science. The conceptualization of a "pedagogical vision" builds from Cole's (1996) idea of ideal artifacts and the process of

“prolepsis”. Cole (1996) conceptualizes ideal artifacts as ideas, experiences and cultural pasts that shape our imagining of what is possible in the future. These artifacts carry meaning across time through prolepsis (Cole, 1996). Through the process of prolepsis ideal artifacts from cultural pasts are projected into the future in ways that structure and mediate activities at present (Cole, 1996). Bringing the idea of prolepsis to schooling, Nasir (2004) describes how educator’s cultural pasts, experiences, and world views (i.e. ideal artifacts) shape the futures they imagine for their students and the organization of school practices at present (e.g. norms, activities). Of particular importance is that ideal artifacts align with actions to influence the structuring of the classroom environment and interactions between teachers and students in the present (Nasir, 2004). The type of vision, I explore in this study centers race, is historically embedded, and politically and racially conscious.

Methods

This research employs qualitative (e.g. interviews) and quantitative (e.g. surveys) data sources with students as the unit of analysis. I utilize a multiple case sampling approach (Miles & Huberman, 1994) with variability on student’s incoming identity and perceived ability in science to capture a range of student experiences.

Program

The Westport summer science program engaged youth from underserved communities in community-based scientific research. Participants attended the program for four hours per day, five days per week for a total of seven weeks during the summer. During program participation, students conducted a seven-week long air quality research project involving a local transportation agency. Students participated in and contributed to all aspects of the research process including: the generation of overarching research questions, experimental design, and data analysis.

Participants

Students in the program attended grades 9-12 from Northern California schools. All student participants (n=11) identified racially as youth of color; the majority identified as Latino. Students entered the Westport program with low average perceived ability in science scores based on pre program surveys and commonly described experiencing little success in science class in school. The majority of the students participated in the summer program because they had failed science class or needed to make up science credits though some students voluntarily participated.

The lead instructor for the Westport summer science research program, “Matt”, graduated with an environmental science degree and had a strong disciplinary and science research background. He identified racially as white. He had been working as an instructor for the summer program and was a veteran at guiding students from underserved communities in the design and implementation of community-based research.

Data sources and analysis

This study utilized four data sources: 1) Pre and post program student surveys, 2) Pre and post program focal students interviews, 3) instructor interview, and 4) program observations.

Pre and post program student surveys

9 of the 11 students in the Westport program completed pre and post program surveys. The goal of pre and post program surveys was to establish a baseline (pre) and determine changes (post) in students’ attitudes towards, ideas about, and perceptions of ability in science as well their ideas about the scientific research process. Pre program surveys were administered on the first day of the program; post program surveys on the last day of the program.

Survey Analysis. Attitudinal questions appeared as Likert scale items on pre and post program surveys. Likert scale items were scored from 1-5 with 5 representing the most positive response and 1 representing the most negative response. All items for a validated scale were considered a subcategory (e.g. science identity). To determine differences between students, data was averaged by subcategory for individual students.

Interviews

All interviews were semi-structured and audio recorded. *Focal students.* Seven focal students completed pre and post program interviews. Interviews lasted ~45 minutes. Pre program interviews were designed to capture students’ ideas about science practices, who does/can do science, how students saw themselves in relation to science and other aspects of their science and racial identities. Post program interviews asked students to describe their research projects, what they did to conduct scientific research, and about shifts in identity.

Scientist Instructor. The lead instructor completed an interview following completion of the program. The interview lasted ~45 minutes. Interview questions captured: goals for instruction, utility of scientific research, science practices promoted and sense making about racial disparities in the professional science community. Interviews also captured instructors' ideas about their students that included: racial backgrounds, societal positioning and experiences engaging youth in scientific research.

Interview Analysis. Focal Students. Transcripts were coded through an iterative process for ideas about science practices and identity that included: engagement in science practices, uses for scientific research, and aspects of science/racial identity. All coding categories emerged from the data and were not predetermined. In some cases direct links could be made between science practices (e.g. collecting data) and identity statements (e.g. I am good at science). When a science practice could be directly linked to an identity statement (e.g. collecting data made me feel good at science), this was coded together as an identity/practice statement. Practice/identity statements were analyzed as a unit to determine how engaging in science practices "functioned" for students with regards to identity construction. *Scientist Instructor.* Transcripts were read and coded for perspectives of science and students and goals for instruction. Together with program observations, interviews were used to determine how an instructor's "pedagogical vision" shaped the program resources they made available.

Program observations

Program observations were made two to three times per week for the duration of the program. Field notes were recorded and were used to capture the mechanisms associated with program participation and the types, availability and uptake of resources in real-time. Program observations were designed to capture moments of positioning between participants and instructors and the types of resources made available and exchanged between participants (e.g. students, instructors) during the programs.

Observation Analysis. Field note coding was done inductively and through an iterative process to remain open to novel constructs, interactions, and resources (Miles & Huberman, 1994). Field notes were coded for types of program resources made available (e.g. instruction), how the resources were taken up and utilized or resisted by participants and moments of positioning between students and instructors.

Findings

Findings show that students' ideas about *what* science is (e.g. science practices) and who can do science shifted together while engaging in scientific research. Participation in community-based research: 1) allowed students to add new types of science practices to their repertoires, and 2) broadened the meanings students' associated with science practices from restrictive (e.g. prescribed, right/wrong) to expansive (e.g. science as iterative process). Developing expansive meanings promoted learning, allowed students to make connection between practices, helped make science more accessible and allowed students to identify as capable science learners. In addition, almost all Westport students identified themselves and their peers as scientists following program participation. Furthermore some Westport students specifically identified youth of color as scientists following program participation.

Findings show that Westport students made gains in perceived ability (e.g. how they saw themselves as science learners) and identified themselves and their peers as scientists because of what they were able to *do* as scientists. For example, different aspects of the data collection process created new ways to engage with science, made science accessible and cultivated the identity of a capable science doer and learner. In addition, engaging in multiple aspects of the research process generated new meanings for science practices and uses for scientific research that elevated the power and agency associated with the identities generated. In addition, findings show how the opportunity to construct new knowledge about their community (i.e. collect/analyze data) and to share this information with people in positions of power (i.e. present research) broadened students' ideas about the utility of scientific research, *where* research can be done, and *who* it can impact.

To illustrate these findings, I (briefly) present a case study of a focal science practice: *presenting research*. I examine relationships between: 1) the program resources (i.e. instruction, pedagogy, designed experiences) made available for the science practices, 2) the nature of the shifts that occurred for students while engaging in these practices and accompanying resources, and 3) the types of identities that were constructed through this process.

Case example: Presenting research

Westport students collected air quality data involving a local transportation agency over the course of the summer. As a culminating presentation, Matt arranged the opportunity for students to present their research to administrators from the transportation agency. The experience of presenting research in a high stakes context

empowered the students and generated a new possibility for *who* they could become in science: *change agents* in their community.

Program resources: Vision and positioning as change agents

Matt had a particular perspective of science and of his students, as youth of color from an underserved community, that informed his programmatic goals and the design of instruction, pedagogy, and experiences he made available. His years of experience working with underserved youth shaped his perspective of his students and his ideas about his role as a science instructor. He describes his perspective:

Science helps reveal the relationship a lot of young people and their families have with the larger system that has alienated them so much from power or having any sort of dignity.

His perspective merges ideas about hierarchical power structures in society, science and the positioning of his students. He views his students and their families as lacking power and science as a tool that can make this transparent. As described above, many of the Westport students were in the summer program because they had failed their science class in school. Matt builds on his idea about how his students are positioned: “So many of the students that we work with have extremely low self-efficacy, self-esteem and a lot of that stems just from their social and class position in society.” Here, Matt makes connections between this lack of power, societal positioning of his students, and their social-emotional well-being. Instead of seeing his students as failing, he sees them as youth who need social-emotional support and encouragement.

Matt’s goal was to create an opportunity for students to see that the air quality data they collected could lead to change and to disrupt the lack of power he felt that they experienced in their daily lives. He describes his perspective of science:

My larger goal is to have them be able to be scientist in all aspects of their lives, whatever they end of up doing. So that they can have a critical lens on the world and so they’re not just, assuming everything is just the ways it is, but they can use science to create change in their lives and communities.

Matt views science as a tool that his students can use to create change in their lives. To support this and his goal to empower his students as change agents, Matt designed an experience that specifically placed his students in a position of power: presenting their research to representatives from the transportation agency where students had collected their data. Matt described his vision for the research presentations that guided the resources he made available for this practice: “They (the students) have to see that all the data collection and analysis that they’ve done, actually leads to...some sort of changes.” Collecting and analyzing data provided an opportunity for students to construct new knowledge relevant to their community and presenting their findings provided a platform to share this new knowledge in an empowering way. Throughout this process, Matt positioned his students as agents of change and provided opportunities that created the potential for this to become a reality.

What shifts for students? Generating identities as change agents

The experience of presenting research functioned in two main ways that allowed students to see themselves as agents of change in their community. First, the experience of communicating their findings through research presentations gave new meaning to this practice and shifted students’ ideas about uses for scientific research. Second, power shifted to the students, as youth of color, as scientists.

Becoming change agents: Broadening ideas about uses for scientific research

Presenting research broadened students’ ideas about the meaning of science practices. During his post program interview, when asked what he learned about the scientific research process, Fernando, a 12th grade student who identified racially as Latino describes his ideas:

Science is not just like...finding out facts. You actually need to know how to present what you are going to do, like to be able to get your data or whatever you researched out to people, you gotta talk to people, like look this is what my project was, this is what I found out, so the majority of it was finding the data. But the other part was like actually knowing how to present it out.

Fernando describes a shift in the meaning he associates with science practices from restrictive and static (i.e. finding facts) to expansive and active (e.g. collect data, present research). He also made a connection between the practices of collecting data and presenting. In this way, the purpose of collecting data was generated through the context of presenting research (i.e. “get data out to people”). In addition, Fernando’s statement indicates a shift in

agency. When he says, “look this is what my project was, this is what I found out”, he asserts that he has something to tell people and that their research is important.

In addition, students’ ideas about uses for and the importance of scientific research shifted because of their experience presenting in a high stakes context. Collecting data in their community and presenting their research findings broadened students’ ideas about *where* scientific research can be done and *who* it can impact. Natasha, a 12th grade student who identifies racially as Chicana provides an example. She described the importance of their air quality research they conducted: “Nobody’s...doing stuff out here, like in low-income communities to make the air better. I guess that's why we're doing the research.” Natasha places a particular importance on *where* she and her peers are collecting air quality data. In addition, Cid, a 12th grade student who identifies racially as Latino makes a similar connection. On a post program survey he stated that his ideas about uses for science were “completely different now” after presenting their air quality research. He explained why his ideas shifted so dramatically: “Because the data can actually be used to change people's mind.” In addition, he stated on the survey that their research was “very important” and when asked to explain to whom he stated: “to people in our community, Black and Brown.” Cid makes an important connection between data and the impact data and research findings can have on people when shared publicly, especially to people in positions of power. As Fernando described above, the students learned that presenting was a way to get their research out to people. Here, Cid describes the impact this information can have to “change people’s minds”. Fernando, Natasha and Cid suggest that their research is important to people in their community and that their findings can be used to create change. In this way, the identity of change agent was constructed.

Shifting power to youth of color

Students expressed a significant transformation and shift in power that occurred over the course of the summer. For example, when Melanie, a 10th grade student who identified racially as Filipina entered the program she described scientists as people who were high status and not from her racial background:

I've never heard of a professional Filipino scientist. Maybe it's just a lot of the other races don't feel comfortable. Here, it's a lot of low-income or working class people and they probably don't feel like they're good enough to become a scientist, cause it's like high.

She describes not knowing any Filipina scientists and reasons that people from certain races and classes might not be “comfortable” or feel that they can achieve the high status of a scientist because of their societal positioning. She also described scientists as “powerful” and when I asked her why she stated: “Because they are labeled scientist, and that makes them smarter than everyone else.” In this way, Melanie describes the power she associates with the position of scientist and how this power is linked to knowledge.

After participation in the summer program, Melanie still views scientists as powerful. She explains:
I think scientist are pretty powerful. I think we were pretty powerful this summer as scientist and through out the year it paid off because we got to present in front of these representatives and I think that was powerful for us to do.

Here, Melanie describes a shift in power: who is a scientist has shifted from people that are “not her race” to herself and her peers as powerful scientists. In addition, she expresses a shift in power to herself and her peers because of what they were able to do as scientists (i.e. present to the representatives). In addition, she describes the importance of working together as a community of youth of color:

I think its really important to represent and show people that scientists aren't powerful because of their race, but because of who they are and what they want to accomplish as a scientist and what they want to contribute to their community so I feel like its really important for people of color to represent scientist.

Presenting their research in a high stakes context provided an opportunity for Melanie and her peers to disrupt historically embedded power structures. Melanie was able to construct a new image of *who* can occupy the powerful position of scientist that included herself and her peers. Together the students experience generated a new possibility; that of becoming *scientists* and *change agents* in their community.

Findings illustrate how the opportunity to construct new knowledge about their community and to share this information with people in positions of power broadened Westport students’ ideas about the utility of scientific research, where research can be done, and *who* it can impact. Matt’s goals were to create an opportunity for students to see that the data they collected could lead to change and to disrupt the lack of power he felt that they experienced in their daily lives. The opportunity to present their research in a high stakes context made these

goals a reality. Engaging in the science practice of presenting research shifted the purposes and uses for scientific research, generated new possibilities for *who* students could become in science, and shifted power to youth of color.

Conclusions and implications

Findings illustrate the emergence of an *identity generative process*: that engaging in science practices (e.g. presenting research) and the accompanying program resources generated new possibilities for Westport students as *capable science learners, scientists, and change agents*. Findings show that the instructional and pedagogical resources made available for science practices determined *how* the practices “functioned” for students. In this way, the generation of new possibilities for focal students was program resource dependent. Furthermore, findings show that the Westport instructor’s vision of science (e.g. purpose, utility) and his students (e.g. racial background, social positioning) guided the design of the program resources he made available. Because he viewed his students as lacking power and having low self-efficacy due to their race/class positioning in society and he viewed science as a tool to create change, he made unique resources available for particular science practices (e.g. collecting data) that empowered students as change agents in their community. Findings show that youth of color can imagine and take up new possibilities for who they can be in science when their identities are holistically supported in science programs.

This research illuminates mechanisms that support the generation of new possibilities for youth of color in science. Exploring this identity generative process as a unit of analysis highlights that simultaneously expanding students’ ideas about *what* science is and *who* can do science generates new opportunities for youth of color in science. In addition, this research extends our understanding of the significance of an instructor’s “pedagogical vision” in guiding the design of program resources (i.e. pedagogy, instruction, designed experiences) that support holistic identity construction.

Scholarly significance

Findings inform the design of learning environments that create multiple pathways for learning and identity construction in science. In addition, findings can be leveraged across contexts to construct educational resources that support meaningful learning opportunities in schools. As states adopt the Next Generation Science Standards (NGSS, 2013), a unique opportunity exists to develop educational resources that engage students in science practices and support identity construction for youth of color in holistic ways. Findings can be applied to the creation of opportunities in science programs, classrooms, and teacher education that foster successful and meaningful engagement with science practices and empower youth of color as capable learners, doers and change agents in science. In this way, we can truly create space for youth of color in science learning environments and broaden opportunities for participation in science.

References

- Aschbacher, P.A., Ellen, E.L., Roth, J. (2010). Is science me? High school students’ identities, participation and aspirations in science, engineering, and medicine. *Journal of Research Science Teaching*, 47(5), 564-582.
- Bouillion, L.M. & Gomez, L.M. (2001). Connecting School and Community with Science Learning: Real World Problems and School-Community Partnerships. *Journal of Research in Science Teaching*, 38(8), 878-898.
- Bowen, G.M. & Roth, W.M. (2007). The practice of field ecology: Insights for science education. *Research in Science Education*, 37, 171-187.
- Calabrese Barton, A., Tan, E., & Rivet, A. (2008). Creating hybrid spaces for engaging school science among urban middle school girls. *American Educational Research Journal*, 45, 68-103.
- Carlone, H.B. & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, 44(8), 1187-1218.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Belknap Press of Harvard University Press.
- Darling-Hammond, L. (2010) *The flat world and education: How America’s commitment to equity will determine our future*. New York: Economic Policy Institute and Teachers College Press.
- Gutierrez, K. (2008). Developing a sociocritical literacy in the third space. *Reading Research Quarterly*, 43, 148-164.
- Harre, R. (2008). Positioning Theory. *Self-Care and Dependent-Care Nursing*, 16(1), 28-32.
- Holland, D., Skinner, D., Lachicotte, W., and Cain, C. (1998). *Identity and Agency in Cultural Worlds*. First Harvard University Press.

- Johnson, A., Brown, J., Carlone, H., and A.K. Cuevas. (2011). Authoring identity amidst the treacherous terrain of science: A multiracial feminist examination of the journeys of three women of color in science. *Journal of Research in Science Teaching*, 48(4), 339-366.
- Lehrer, R. (2009). Designing to develop disciplinary dispositions: Modeling natural systems. *American Psychologist*, 64(8), 759-771
- Malone, K.R., & Barabino, G. (2008). Narrations of race in STEM research settings: Identity formation and its discontents. *Science Education*, 93(3), 485-510.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Second edition. Thousand Oaks: Sage.
- Nasir, N.S. (2004). "Halal-ing" the child: Reframing identities of resistance in an urban Muslim school. *Harvard Education Review*, 74(2), 153-174.
- Nasir, N.S. (2012). *Racialized Identities: Race and Achievement Among African American Youth*. Stanford University Press, Stanford, CA.
- Nasir, N.S. & Cooks, J. (2009). Becoming a hurdler: How learning settings afford identities. *Anthropology & Education Quarterly*, 40(1), 41-61.
- Nasir, N.S. & Hand, V. (2006). Exploring sociocultural perspectives on race, culture, and learning. *Review of Educational Research* 76(4), 449-475.
- National Science Board (2012). *Science and Engineering Indicators Digest 2012*. Arlington VA: National Science Foundation (NSB 12-02).
- National Research Council. (2011). *Successful K-12 STEM Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics*. Washington, DC: National Academies Press.
- Oakes, J., Ormseth, Bell & Camp (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science*. RAND Report. Pp. 1-152.
- Parsons, E. C. (2014). Unpacking and critically synthesizing the literature on race and ethnicity in science education. In N.G. Lederman & S.K. Abell (Eds.), *Handbook of research on science teaching*, volume II (167-186). New York: Routledge.
- Poleman, J.L. & Miller, D. (2010). Changing stories: Trajectories of identification Among African American youth in a science outreach apprenticeship. *American Educational Research Journal*, 47, 879-918.
- Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, 89, 634-656.
- Verelas, M., Martin, D.B, Kane J.M. (2012). Content learning and identity construction: A framework to strengthen African American students' mathematical and science learning in urban elementary schools. *Human Development* 55, 319-339.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. New York: Cambridge University Press.

“I’m Not Just a Mom”: Parents Developing Multiple Roles in Creative Computing

Ricarose Roque, Karina Lin, and Richard Liuzzi
ria@media.mit.edu, karinalin@gmail.com, richard.liuzzi@gmail.com
MIT Media Lab

Abstract: Creating, designing, and building with computing have gained recognition as important activities for children, but few actually engage in these creative opportunities. Social support from parents can be essential to engaging children, but for parents with limited backgrounds in computing, they are often unsure what roles they can play. For parents to develop supportive roles, they must gain first-hand experience in creative computing for themselves and with their children. In this paper, we examine the experiences of parents participating in a community-based program called Family Creative Learning, where families design and invent together using creative technologies like Scratch. We describe three case studies of how parents developed multiple roles such as collaborator, teacher, and learner as they created and collaborated with their children on technology-based projects. We discuss design opportunities to provide meaningful experiences for parents and children to build projects and build supportive roles around computing.

Keywords: parents, families, roles, creative learning, computing, Scratch, MaKey MaKey

Introduction

While narratives of the “digital native” might suggest that young people are using computer programming and design-based computing activities on their own, many studies find a different story. In a study of youth who engaged in creative activities with computing, Barron and colleagues (2009) identified the many roles that parents play to support their children. Parents enact roles like collaborator, teacher, and learner to encourage, sustain, and deepen their children’s learning experiences with computing. Although these roles do not require parents to possess domain knowledge, all of the families that Barron and colleagues (2009) studied had at least one parent with a technology-related background.

Parents without backgrounds and insights into the changing landscape of technology struggle to negotiate what roles they can play, such as how to work together in computing activities or how to find learning opportunities for their children (DiSalvo, Ried, & Roshan, 2014; Roque, 2013). At the same time, traditional roles are becoming complicated. Children develop roles as “technology brokers” with expertise beyond their parents (Correa et al., 2013). As a result, parents and other adult caretakers may assume more distant roles where their children will “just pick things up.” However, parent co-activity and discussions can nurture their children’s learning with technology (Simpkins, Davis-Kean, & Eccles, 2005). To support families to negotiate these emerging challenges with technology, studies suggest designing learning experiences where families can engage in joint activities around computing and parents can develop their expertise (Livingstone, Mascheroni, Drier, Chaudron & Lagae, 2015; Takeuchi, & Stevens, 2011). While there are many opportunities to broaden participation in computing, many initiatives engage young people or adults exclusively and often focus on individual skill development.

To examine how we can support parents who have limited expertise with technology to develop more active roles, we present a study of parents’ experiences participating in a community-based program called Family Creative Learning (FCL) (Roque, 2016). Through FCL, parents and children learn to use new technologies and design projects together based on their interests. The program targets families with limited access to resources and social support around computing. Parents gain hands-on experience designing their own projects with creative technologies, while experimenting with roles that they can enact to support their children in this context. We argue that examining this process is a crucial first step towards understanding what supports and environments parents need to construct roles that align with their families’ needs and backgrounds.

We ask (1) what kinds of roles did parents enact as they designed projects and worked with their children and (2) how did this creative computing experience support parents’ development of these roles? We focus our study on three parents, each with children who are excited about the possibilities with computing. These parents ranged in technical expertise and came into the experience with varied goals, interests, and histories. We describe how families collaborated on projects and how different parental roles developed. Finally, we reflect on the

opportunities to further design experiences for parents and children to develop as learning partners with computing.

Background

In recent years, there has been a growing recognition of the importance for children to create, express, and build with technology. They can develop as computational thinkers as they engage with concepts, practices, and perspectives supported by computing (Brennan & Resnick, 2013; Wing, 2006). However, studies highlight a troubling participation gap in who engages in these creative activities, particularly among young women and ethnic minorities (Livingstone & Helsper, 2010; Margolis & Fisher, 2003). To understand how we can support broader participation in creative activities with computing, many argue that we need to move beyond thinking about access to technology and consider the broader ecology of social support and opportunities that surround a young person (Ito et al., 2013; Barron 2004).

Brigid Barron and colleagues (2009) took an important step to identify the kinds of roles that parents can play to support their children in the development of technological fluency. They identified seven roles: teacher, collaborator, resource provider, learning broker, non-technical consultant, employer, and learner. These roles demonstrate the ways that parents support children such as sharing expertise, finding opportunities, and providing encouragement. However, there are limitations to the general applicability of these roles. At least one parent in each household had a technology-related background with an appreciation and vision for the role of technology in their children's lives.

While Barron and colleagues identified roles, we also need to understand parents' motivations and values that underlie these roles. Clark (2012) identified two guiding principles that influenced parents' and children's uses of technology: an "ethic of expressive empowerment," or uses of technology for educational and enrichment purposes, primarily among middle to upper income families; and an "ethic of respectful connectedness," or uses of technology to maintain family connections, primarily among lower-income families. Research strongly emphasizes the influence of parents' support in children's learning experiences and pathways with technology (Simpkins, Davis-Kean, & Eccles, 2005). However, what supports are available for parents who have diverse motivations and backgrounds to develop positive and constructive roles? Parents also need learning opportunities and first-hand experience to understand the kinds of roles they can play.

While there have been a number of successful programs for families from under-represented groups in informal STEM learning activities (Weisbaum, 1990; Heil et al., 2012; Rivas & Olmsted, 2013), these initiatives have generally not focused on computational activities. Other programs have engaged parents and children in computer-based learning experiences, such as Tech Goes Home (<http://techgoeshome.org>) and Computers for Youth (<http://cfy.org>), but these initiatives have typically focused only on basic computer literacy for parents, such as looking up information online and using office applications. While these are important first steps in computational fluency, these efforts still fall short in providing opportunities that go beyond using and interacting and into creative and expressive uses of computing.

Design of Family Creative Learning

Family Creative Learning (FCL) is a community-based program that invites families to design and invent together using creative technologies (Roque, 2016). FCL has five workshops and are held in a community center once a week for two hours each. FCL is collaboratively implemented with staff from community centers, such as Boys and Girls Clubs and centers at housing developments.

The design of FCL draws on constructionist traditions of learning, which argue that people learn best when they are building things that are personally and socially meaningful (Papert, 1980; Kafai, 2006). Constructionism builds upon constructivist traditions that knowledge is not something that is transmitted or acquired, but something that is actively constructed through experience (Piaget, 1976). As people build projects, they build ideas. To be personally meaningful, the design of FCL invites families to build on their diverse "repertoires of practices" and "funds of knowledge" (Gutiérrez & Rogoff, 2003; Moll, Amanti, Nef & Gonzalez, 1992). To be socially meaningful, the design of FCL has also leveraged learning theories that emphasize the social aspects of learning (Brown, Duguid, & Collins, 1989; Lave & Wenger, 1991). Families are encouraged to work together as well as interact with other families participating in FCL.

The design of FCL goes beyond instructionist learning environments where there is a central instructor and pre-determined activities. FCL has four aspects that define its structure: tools, activities, facilitation, and environment. *Tools* include the Scratch programming language (Resnick et al., 2009) and the MaKey MaKey physical invention kit (Silver, Rosenbaum, & Shaw, 2012). *Activities* structure how families learn, interact, and collaborate on projects. *Facilitation* consists of research and community center staff and volunteers who provide

families with social support and encouragement. *Environment* includes both the physical arrangements and the socio-emotional space, which define how easily participants can interact and feel safe to take creative risks.

The five workshops are organized as follows: In workshops 1 and 2, parents and children independently learn how to use Scratch and MaKey Makey. Scratch enables families to create a broad range of interactive media such as games, animations, and stories, while MaKey Makey enables people to simulate a keyboard using everyday materials such as aluminum foil and Play-Doh. In workshops 3 and 4, they collaborate on a family project. This arrangement allows parents and children to focus on being creators first before they engage with each other as collaborators. In workshop 5, they share this family project in a community showcase.

Each workshop is divided into four parts: Eat, Meet, Make, and Share. In *Eat*, workshops begin with a meal from a local restaurant, allowing all participating families and facilitators to eat with each other. In *Meet*, families split up into two groups and facilitators meet separately with parents and children to talk about their experiences in the workshops. In *Make*, parents and children engage in design activities with Scratch and MaKey Makey. In *Share*, families talk about their projects to other families and receive feedback and questions.

In FCL, family members have the opportunities to take on roles that Barron and colleagues (2009) described, such as teacher, learner, collaborator, or non-technical consultants, as they participate in workshops that enable and support them to work on their own projects as well as a family project. The next sections describe how we studied families' experiences and the ways that these roles emerged.

Studying families' experiences

Settings and participants

We implemented the workshops in a Computer Clubhouse located within a Boys and Girls Club in an urban community in the northeastern United States. Computer Clubhouses are informal learning centers designed for low-income youth to engage in creative activities with technology (Kafai, Pepler, & Chapman, 2009). We recruited children between the ages of 7 and 12, but welcomed their younger and older siblings to participate.

We observed families in two program implementations in Spring and Fall 2013. In Spring 2013, five families participated for a total of fifteen unique participants with five parents (four mothers, one father; ages thirty-one to fifty-eight) and ten children (four girls, six boys; ages four to thirteen). Two of the parents were single mothers. Three of the families self-reported as Hispanic/Latino and these parents were immigrants. The other two families self-reported as White. All five families continued participating through the community showcase. In Fall 2013, six families initially participated, but two dropped out because of a family emergency and conflicting commitments. Among the remaining four families, there were nine unique participants with four parents (two fathers, one mother, and one grandfather; ages thirty-seven to fifty-nine) and five children (one girl, four boys, ages seven to eleven). One of the families self-reported as Hispanic/Latino, one self-reported as Black/African-American, and two families self-reported as White.

Data collection and analysis

To understand families' experiences, we primarily used ethnographic methods and collected multiple forms of data to triangulate our observations, in the form of field notes, individual and group interviews, and video recordings. During the Meet sessions, we asked parents and children questions such as, "What was it like to see your parent/child create a project with Scratch and MaKey Makey?" and "What was challenging in working together?" Facilitators, which included the authors, wrote field notes. The Meet sessions and interviews were recorded and transcribed. To better understand participants' experiences during and after the workshop series, we conducted interviews with parents and children within one month after the series ended. We asked questions such as "How did you help your family member?" and "Why did you continue participating?" Qualitative data consisted of approximately 8 hours of video recordings, 70 observed hours total from 6 facilitator-researchers, 20 group interviews during the Meet sessions, and 14 individual follow-up interviews from 60-90 minutes.

We analyzed our data using grounded theory strategies (Charmaz, 2006) to uncover processes that contribute to parents development of roles by examining what parents did, how they interacted with their children, and what they said about their experience. We used the roles identified by Barron and colleagues (2009) in our coding and used the grounded approach to discover any other roles. All three authors participated in data collection, transcription, coding, and analysis. Through a six month period, we met weekly to discuss our data analysis and cases. These discussions served as our inter-coder reliability checks, making sure that each analysis was approved by all team members.

Of the original nine families, we focused on three families. These families were selected using the following criteria: (1) completion of a project for the community showcase and (2) availability of all participating family members for interviews after the workshop series. The three families show the different ways that parents

experienced this program with their children, but they are by no means representative of all parents' and families' experiences in the workshops.

Results

Our coding and analysis led to the identification of how parents developed different roles to support their children. We focused on the kinds of practices that parents enacted (e.g. brainstorming and asking questions), perspectives that parents said about themselves and their children (e.g. feeling uncertain about their skills, seeing how creative they can be), and interactions that emerged from parent-child dynamics as they participated in FCL (e.g. responding to children's struggles, building on their children's interests). The experiences of these three families highlight broader trends in other parents' development of roles, particularly the ways that parents built on their existing practices, connected to new perspectives about themselves, their children, and technology, and developed their interactions with their children in creative and positive ways around computing.

Sandy and Pete: Leveraging each other's interests and strengths

Sandy, a 41-year-old hair stylist, was excited to work on a project with her 10-year-old son, Pete. Pete had been learning to use Scratch before the workshops in the Computer Clubhouse, but MaKey MaKey was new to him. Sandy felt anxious after she made her Scratch project in workshop 1, as she considered herself a newcomer to technology. She looked forward to working with Pete and felt less overwhelmed knowing that he could help her.

During the child Meet session in workshop 4, Pete shared how tech savvy he was compared to his mother and how excited he was to serve as her "Scratch parent." In the parent Meet session, Sandy also shared how they often worked together for Pete's school projects, although she had never collaborated with him on a technology-based project. Sandy looked to Pete as her technical consultant and shied away from learning anything related to technology on her own. As they started their project and brainstormed ideas, Sandy was inspired by both the craft materials in the room and Pete's upcoming violin performance, and suggested they make a violin for the family project. Pete liked the suggestion and they continued to exchange ideas as they designed elements of a violin with MaKey MaKey and Scratch.

Sandy delegated tasks between them based on their interests and strengths, and told Pete, "I'll make it [the violin], you make sure it works." Pete used his abilities in computing to set up the Scratch code, and Sandy used her abilities in drawing to make the violin out of materials they found. Despite having delineated tasks, Pete and Sandy checked in with each other consistently throughout their process. As Sandy drew a violin on cardboard, she asked Pete to cut it out. As Pete looked through sounds on Scratch, he made sure to get feedback from Sandy for every sound he programmed. Throughout the workshop, Sandy asked Pete to let her try to troubleshoot when Pete struggled with the MaKey MaKey. Although Pete was unwilling at first, Pete and Sandy eventually shared the duty of attaching wires to the MaKey MaKey. In a post-workshop interview, Sandy indicated, "I felt more comfortable [using the MaKey MaKey] 'cause I could kind of bring my skills to it."

Between the workshops, Sandy and Pete discussed the project at home and how they could make it better. They brought Pete's violin to the workshop and used Scratch to record him playing violin. A facilitator suggested to Pete to include an image of his mother somewhere in the Scratch project just like he included a picture of himself. Sandy was initially excited about this idea, but when she noticed Pete's hesitation, she instead declared that they were satisfied with the project and called it "done."

When they shared their project at the community showcase, Pete actively included his mother when describing their design process. Both Sandy and Pete were proud and excited, and were eager to continue to use MaKey MaKey and Scratch. After the workshop, Sandy shared how pleased she was that her son was able to see her as "not just a mom" and that she finally "got to see what [Pete] was actually learning...and he wasn't just sitting there and playing games."

Case analysis

Sandy initially expected to be hands-off during the Make sessions because she never saw herself playing a significant role in technology-related projects. Pete echoed these thoughts when he expressed his intentions to assume a "Scratch parent" role to his mother. Before coming to the workshops, Sandy and Pete often collaborated on Pete's school projects. Within FCL, Sandy continued to practice the same roles, collaborator and non-technical consultant, that she typically used with Pete in other activities (Barron et al., 2009). Sandy remained flexible and made sure to balance the roles she enacted between her child's interests and her own capacities (Drummond & Stipek, 2004). Sandy moderated her collaborator role as she weighed her growing excitement in the creative process and Pete's interests, which she prioritized over her own. In leveraging each other's strengths, Sandy often deferred more technical tasks to Pete, limiting her opportunities to continue developing her own expertise with the technologies. By finding multiple meaningful entryways for involvement with Pete in a technology-related

project, Sandy took a significant step towards active involvement where she values the activities Pete engages in and encourages his continued participation (Peterson, 2013).

The craft materials with the MaKey MaKey were important FCL tools that connected Sandy's creative side, a side of herself she places values in and finds meaningful. The ability to be creative with the tools alleviated Sandy's anxiety regarding technology and brought her towards a wider perspective of the roles she could play with Pete and technology.

Rosa, Sonia, and Clara: Stepping back and stepping in

Rosa, a 49-year-old physical therapist assistant and mother of two daughters—Clara, 13-years-old, and Sonia, 9-years-old—looked forward to learning more about computers in the workshops and spending time with her daughters. However, she knew that unlike her daughters, who seem to pick up computer-related things very quickly, it would take her some time to do so.

During workshop 3, to create their community showcase projects, Rosa worked with Clara, while Sonia worked on her own on a table next to her family. Facilitators suggested this grouping after sensing Sonia's independent personality. Like Sandy and Pete, Rosa and Clara were aware of each other's abilities and interests and collaborated on two musical instruments: a cardboard guitar and foil drum set. Rosa got the physical materials ready, while Clara programmed Scratch. Together, they connected the instruments to MaKey MaKey.

Meanwhile, Sonia was working next to them with the help of a facilitator Alex. While her mother and older sister were already assembling their two instruments, Sonia was still trying to get her MaKey MaKey project to work with pieces of Play-Doh. Alex helped her through her various issues such as connecting MaKey MaKey to her computer. Once she got some musical notes on Scratch to play with the Play-Doh through the MaKey MaKey, Sonia decided to make a drum with foil plates and took apart her working project. At this point, Alex left to check-in on other families while Rosa checked-in with Sonia. Rosa expressed enthusiasm for Sonia's drum project and helped her in disconnecting MaKey MaKey from the Play-Doh. When they shared projects at the end of the workshop, Clara and Rosa shared their nearly finished project, while Sonia had difficulty getting hers to work, insisting that it worked before. She felt discouraged when she was unable to fix it. Rosa gave Sonia a hug, reassuring her that she did a great job and that her project just needed to be fixed.

In the parent Meet session during workshop 4, Rosa expressed how much she enjoyed working with MaKey MaKey, but wondered if she was being helpful to her daughters who seemed to know more than her. A facilitator heard these concerns and assured Rosa that her practices in providing feedback and encouragement for her children were helping Clara and Sonia work through their projects. The facilitator then asked Rosa how she felt about Sonia working on her own. Rosa felt that because of Sonia's independent spirit, it made sense for her to work on her own project, but she planned to pay more attention to her in this workshop.

In the Make session, Rosa made sure to situate herself between her two girls so she could also check on Sonia. While Rosa tried to be more involved in Sonia's project, Sonia already had a clear vision and wanted more control over the process, trying to program the project and build the MaKey MaKey connections herself. Rosa decided to step back and watched Sonia, providing support when Sonia asked her for help. However, as Sonia continued to struggle, Rosa became more involved, helping to figure out issues with facilitators and making suggestions when Sonia seemed unsure. For example, knowing Sonia's sense of style, Rosa suggested decorating the cardboard with pink leopard print duct tape, which Sonia happily incorporated. As Sonia finished her project, Rosa made sure to give her positive comments on what she had accomplished.

Case analysis

Rosa's responsiveness to her two children reflected the kinds of roles that she took on to support them. With Clara, she seamlessly became a collaborator and engaged in different design practices such as building on each other's interests, splitting up tasks, and working together to integrate the different parts of the project. With Sonia, Rosa had to adapt her roles based on Sonia's needs. Sonia wanted to work independently and Rosa initially stepped back becoming more of a non-technical consultant as she encouraged her and gave her reassurance, especially since a facilitator was helping Sonia. However, after noticing Sonia's struggles to work independently, Rosa responded to her daughter, stepping in and out more to scaffold the creative process for her like a facilitator. Although Brigid Barron and colleagues (2009) described the teacher role as someone who could pass expertise to their child, Rosa scaffolded her daughter's design experience based on what she felt her daughter needed, despite still developing her own expertise. Rosa's case represents the fluidity of parent roles and how they respond to the situation and what they believe their children need.

Meanwhile, facilitators were also observing Rosa supporting her children. In workshop 1, Rosa initially expressed uncertainty in her ability to support her children. While Rosa displayed both design practices and supportive practices during the workshops, she still admitted feeling like she was not contributing. Facilitator

feedback and joint reflection with other parents helped Rosa make connections from her experience to the roles she can play with her children.

Tim and Ethan: From regulating to facilitating

When Tim, a 38-year-old father of three, first came to the FCL workshops with his 11-year-old son Ethan, he was unsure what the program was even about. While Tim used computers for his mapping work for the city, he was insecure about his abilities to learn how to code. He was also wary of Ethan's growing interest in computers, and monitored the amount of time Ethan and the rest of his children could use the computer, even referring to his role as "the guardian, to kind of regulate." Despite these reservations, Tim wanted to better understand Ethan's interests. He understood his other children's interests in dance and hockey, but was unsure of what Ethan was "into." Ethan, on the other hand, was very excited to attend the workshops. He was already learning how to use Scratch at the Computer Clubhouse and wanted to learn more.

During the Make session in workshop 1, Tim did not ask many questions, but instead followed directions in an activity guide provided to him by facilitators. He soon explored the blocks on his own. By the end of the evening, Tim created a project that animated his initials and played a recording of his voice saying hello to his children. After the workshop, Tim pursued his growing curiosity with Scratch and created a Scratch account for himself. During workshop 2 Make session, he explored more advanced features such as cloning. During the Share session, a facilitator asked him to explain what he did in his project. Tim claimed to not know, but a facilitator encouraged him to try and he described his process successfully.

While Ethan created more advanced Scratch projects than Tim, Ethan started to feel eclipsed by his father's growing proficiency and enthusiasm. However, when it came time for them to work together on a project in workshop 3, Ethan was excited when facilitators asked the children to use their expertise with Scratch to support their parents. He was looking forward to sharing what he can do with Scratch with his father, and he hoped to learn something from his father as well.

When it came time for father and son to work together, Tim's growing confidence with Scratch and Ethan's eagerness to be a Scratch mentor to his father did not materialize in the type of project collaboration that facilitators had expected. Over the course of the Make session, Ethan primarily took charge of the project, while Tim assisted and observed. As Ethan became familiar with MaKey MaKey, Tim took items from the materials table and presented them to Ethan to experiment with. He would occasionally give Ethan suggestions on how to connect them, reminding Ethan what he needed to do to make his project work. As Ethan took greater ownership of the project, Tim became less involved and observed more as Ethan made a "banana mouse" that used the mouse movement inputs on MaKey MaKey. During Share, Ethan talked about his design process and project without mentioning his father, who watched as he presented the project.

Case Analysis

While Tim's limited involvement in their shared project could be explained as disengagement, Tim was an active facilitator rather than a passive observer. He consciously chose to let his son take the lead, while he supported him—reversing typical parent and child roles, with the son as the leader and the father as the assistant. Tim observed what Ethan was doing, provided assistance in learning how to use MaKey MaKey without giving direct answers, and encouraged Ethan to experiment with materials and design their project. Tim wanted to support his son's passion for technology and skill with Scratch by allowing Ethan to take charge. As Ethan took more ownership, Tim gradually became less involved, but remained observant, ready to step in if needed. In a post-workshop interview, Tim said:

He [Ethan] knows what he wants to do and he'll let me know if I can help. For the most part, he's going to run the show and that's fine with me...I'm pretty secure with myself. I think it empowers them... I think they're proud to be able to say "Oh, listen, I know how to do this. Let me show you."

Through his desire to uncover what his son was "into" and how his son defined himself, Tim immersed himself in learning Scratch and MaKey MaKey in the first two workshops when he was working on his own. In doing so, Tim had the opportunity to consider the parental role he played as a gatekeeper regarding Ethan's use of technology. For some parents, anxieties around technology use can sometimes lead to decisions that can limit computing opportunities for their children (Tripp, 2011). By experiencing creating with Scratch and MaKey MaKey himself and watching his son enthusiastically build a project, Tim came to better appreciate what his son is "into" and how his son identified with this interest. This shift in perspective creates potential for Tim to assume a learning broker role that finds opportunities for his children.

Discussion: Enabling and supporting parents as learning partners

A goal of Family Creative Learning was to create an environment where parents can both engage with their children in design-based computing activities as well as explore roles to support their children in an unfamiliar but increasingly important context within our digital society. We wanted parents to not only walk away with content and skills, but also to experiment with roles to support their children around computing.

The case studies demonstrate the ways that parents' participation in these design-based activities with their children enabled and supported the roles that they played in this program. At the beginning of the workshops, parents expressed concerns about what their children do with technology and anxieties around supporting their children with computing activities. Some parents called themselves "computer illiterate," while others saw themselves as regulators of their children's computer use. However, over the course of the workshops, the parents in these case studies demonstrated their commitment to supporting their children through a variety of practices and roles. Sandy and Pete leveraged each other's strengths, ideas, and interests. Rosa tried to step in and out throughout Sonia's successes and struggles with her project. Tim relinquished creative control to Ethan, but watched him closely, made suggestions, and gave encouragement along the way. These differences demonstrate the importance in designing environments that invite parents to explore and discover roles that build upon their strengths, children's needs, and family's goals — rather than prescribing defined and constrained roles without changing trajectories of participation.

There were also tensions present with parents and children working together on projects within a context the children often felt a strong ownership over. While they liked seeing their children take charge and display their growing expertise, parents sometimes had challenges integrating their own interests into their shared experiences. For example, even though Sandy and Pete had a generally productive collaboration, Sandy often had to ask Pete to give her a chance at working with MaKey MaKey, too. We found that every family experienced these kinds of tensions, but these tensions were part of the family's continual negotiation of roles around computing. More importantly, these negotiations enabled parents to explore and experiment with practices and see for themselves what worked and what did not to support their children. Facilitators and designers of these environments should pay attention to when these negotiations and tensions lead to disempowerment or marginalized participation. While not documented in these case study families, there were some families that could not work well together in the workshops. Facilitators often intervened to help them think aloud their ideas together or to suggest that they work on separate but nearby computers.

Additionally, children experienced first-hand the benefits of working with their parents on design-based projects. Because of children's greater familiarity with and excitement around computing, the typical teaching roles between parents and children were reversed, with children often leading the creative process and explaining the technologies to their parents. This reversal allowed children to try out new roles, such as a teacher of technology to their family members, in addition to being a creator. Their parents also reinforced these new roles and practices by asking their children to help them or take the creative leadership.

We are also interested in how to design environments that can support families to develop as creative learning partners around computing. It is not enough to merely enable parents and children to create together. *How* we design the learning experience matters. The tools, activities, facilitation, and environment should be designed both for expressive empowerment and respectful connectedness (Clark, 2013) and to support the multiplicity of motivations, goals, and priorities of families. In these case studies, we found that these features of FCL contributed to parents' engagement as learning partners: the flexibility and ease of use of the tools allowed families to create different kinds of projects that connected to the different kinds of interests in the room. The use of crafts and everyday materials also became a bridge for parents from one familiar medium into a new one. The activities fostered interactions within and across families through shared meals, meetings with their peer groups, collaborative interactions, and opportunities to talk about and share their work. Facilitation positioned the participants as creators and collaborators in addition to supporting the development of their expertise. The environment created a physical space where they could work together. It also created a socio-emotional space where they could feel respected and safe within a technology-based context where they were often unsure about their capacities and the roles they could enact for their children. These features are not independent, but interact throughout the workshops to support parents in negotiating roles and practices.

By engaging parents in a creative learning environment with their children, parents and children had opportunities to see both themselves and each other take on more empowered roles as learning partners. Parents could see the positive and creative things that their children could do with computers—an object that was often a source of contention between family members. Children could see their parents as creative designers with computers and experience working on projects together as a family—activities which often fell in the domain of games, crafts, and homework. Computing outreach programs often serve children, without integrating other people in the larger learning ecology (Barron, 2004). Children are left to explain and advocate for their interests

to their peers, parents, and teachers (DiSalvo et al., 2013). Through a shared experience of designing and creating their own projects, families could apply practices that they used in other activities, such as homework help, and adapt it to the context of computing. Families can build connections to this important context in their lives while building relationships within their families and connecting to other families in their community. By engaging in design-based computing activities at their own community center, parents come to understand the wider learning ecology around their children's developing interests and see the kinds of people, activities, and interactions that can support their children—and develop a variety of ways to participate in these worlds as well.

References

- Barron, B. (2004). Learning ecologies for technological fluency: Gender and experience differences. *Journal of Educational Computing Research*, 31(1), 1–36.
- Barron, B., Martin, C. K., Takeuchi, L., & Fithian, R. (2009). Parents as learning partners in the development of technological fluency. *International Journal of Learning and Media*, 1(2), 55–77.
- Barton, A. C., Drake, C., Perez, J. G., Louis, K. S., & George, M. (2004). Ecologies of parental engagement in urban education. *Educational Researcher*, 33(4), 3-12.
- Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *Presented at the American Education Researchers Association*, Canada.
- Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Clark, L.S. (2013) *The parent app*. Oxford, United Kingdom: Oxford University Press.
- Correa, T., Straubhaar, J.D., Chen, W. & Spence, J. (2013) Brokering new technologies: The role of children in their parents' usage of the internet. *New Media and Society*, 0(0), 1-18.
- DiSalvo, B., Guzdial, M., Bruckman, A., & McKlin, T. (2014). Saving face while geeking out: Video game testing as a justification for learning computer science. *Journal of the Learning Sciences*.
- DiSalvo, B., Reid, C., & Roshan, P. K. (2014). They can't find us: the search for informal CS education. *In Proceedings of the 45th ACM SIGCSE* (pp. 487-492). ACM.
- Gutiérrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher*, 32(5), 19–25.
- Ito, M., Baumer, S., Bittanti, M., Boyd, D., Cody, R., Herr-Stephenson, B., Horst, H. A., et al. (2009). *Hanging Out, Messing Around, and Geeking Out: Kids Living and Learning with New Media*. MIT Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Livingstone, S., & Helsper, E. (2010). Balancing opportunities and risks in teenagers' use of the internet: The role of online skills and internet self-efficacy. *New Media & Society*, 12(2), 309–329.
- Livingstone, S., Mascheroni, G., Drier, M., Chaudron, S., & Lagae, K. (2015) How parents of young children manage digital devices at home: The role of income, education and parental style. Retrieved from: <http://eprints.lse.ac.uk/63378/>
- Margolis, J., & Fisher, A. (2003). *Unlocking the clubhouse: Women in computing*. Cambridge, MA: MIT press.
- Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice*, 31(2), 132–141.
- Papert, S. (1980). *Mindstorms: children, computers, and powerful ideas*. New York, NY, USA: Basic Books.
- Peterson, T. K. (2013). *Expanding minds and opportunities: Leveraging the power of afterschool and summer learning for student success*. Washington, DC: Collaborative Communications Group.
- Resnick, M., Maloney, J., Monroy-Hernandez, A, et al. (2009). Scratch: Programming for all. *CACM*, 52, 60-67.
- Roque, R. (2013, July 15) Collateral benefits: Focus groups as social support groups. [Web blog]. Retrieved from: <http://ethnographymatters.net/2013/07/15/collateral-benefits/>
- Roque, R. (2016, *Forthcoming*) Family Creative Learning. In Peppler, K., Kafai, Y., & Halverson, E. (Eds.) *Makeology in K-12, Higher, and Informal Education*. New York, NY: Routledge.
- Silver, J., Rosenbaum, E., & Shaw, D. (2012, February 19–22). Makey Makey: Improvising tangible and nature based user interfaces. *Proceedings of the International Conference on TEEI*. New York: ACM.
- Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2005). Parents' socializing behavior and children's participation in math, science, and computer out-of-school activities. *Applied Developmental Science*, 9(1), 14-30.
- Takeuchi, L & Reed, S. (2011) The new co-viewing: Designing for learning through joint media engagement. Retrieved from: http://www.joanganzcooneycenter.org/wp-content/uploads/2011/12/jgc_coviewing_desktop.pdf
- Wing, J.M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33-35.

Scaffolding Into Ambitious Teaching: Representations of Practice in Teacher Workgroups

Jason Brasel, Brette Garner, and Ilana Seidel Horn
jason.brasel@vanderbilt.edu, brette.garner@vanderbilt.edu, ilana.horn@vanderbilt.edu
Peabody College, Vanderbilt University

Abstract: Teacher workgroups are increasingly mandated in policy as a strategy for instructional improvement, but without accounting for teacher learning. In this comparative case study, we take a situative view of learning to examine the learning opportunities afforded in workgroup meetings with middle-school math teachers and pedagogically expert facilitators. Using interaction analysis, we analyze the role of representations in affording learning opportunities. The facilitators framed workgroup activity differently than the teachers; learning opportunities were limited when teachers were not able to engage the facilitators' framings. But by introducing rich representations of practice, the facilitators attuned teachers' attention to key conceptual resources. This scaffolded the workgroup into new forms of practice, thereby supporting teachers' opportunities to learn more ambitious forms of professional practice. This work highlights the need for facilitators and instructional leaders to attend to issues of teacher learning in collaborative activity.

Keywords: teacher workgroups, representations of practice, framings, learning opportunities, ambitious instruction

Introduction

A substantial body of research has demonstrated that common approaches for teachers' professional development—e.g., short-term workshops—lack the power to transform instruction (Borko, 2004; S. M. Wilson & Berne, 1999). A promising proposal is to move professional development closer to teachers' practice (Ball & Cohen, 1999); one way to do this is through the development of teacher workgroups like content-based professional learning communities or grade-level planning teams. Teacher communities can support teachers' learning because they are embedded in school contexts and give teachers space to discuss problems of practice. Indeed, high correlations between strong teacher communities and greater-than-expected student achievement suggest that such communities are productive sites for teachers' learning (Bryk, Sebring, Allensworth, Easton, & Luppescu, 2010; McLaughlin & Talbert, 2001). This has led school leaders to mandate teacher communities as a mechanism for educational improvement. This strategy is often implemented under the assumption that if teachers have time and space to meet, then instructional improvement will happen (Little, 2003).

Many scholars have interrogated this assumption. They have identified aspects of workgroup activity—for example, participation structures (Horn, 2010) and epistemic stances (Hall & Horn, 2012)—that shape the nature of teachers' opportunities to learn. One endemic challenge for teacher workgroups is representing practice. Teachers typically discuss their work asynchronously from actual instruction, which makes it difficult for workgroup members to establish intersubjectivity. This drives the need for representations that adequately illustrate problems of practice (Hall & Horn, 2012).

Though other conversational features also shape teachers' learning opportunities in workgroup conversations, this study attends specifically to how representations of practice can be used in math teacher workgroups to support instructional improvement. We argue that participants can use representations of practice to attune others' attention to key conceptual resources, thereby scaffolding them into new ways of participating in workgroup activities. Teacher workgroups can thus become productive sites for instructional improvement as teachers are supported in engaging in more ambitious (Lampert, Boerst, & Graziani, 2011) forms of practice that support procedural fluency and conceptual understanding for all students.

Conceptual framework

We take a situated view of learning and define learning as a change in participation in a community of practice (Lave & Wenger, 1991). Since our analysis focuses on joint interaction, we view participation as presenting *opportunities* to learn (OTL) and refrain from making claims about individuals' learning. Following the work of Horn and colleagues {Hall:2012fg}, we operationalize OTLs by examining how interactions (a) marshal conceptual resources for teachers and (b) mobilize teachers for future work. In discussions with rich learning opportunities, teachers collectively develop concepts about pedagogical issues and connect these to their future

instruction. Such discussions afford learning opportunities as they position teachers to engage in new forms of professional practice.

Horn, Kane, and Wilson (2015) identified four interrelated conversational features that shape learning opportunities in teacher workgroup conversations: activity structures, frames, epistemic claims, and representations of instructional practice (Figure 1). These features are further shaped by the epistemic community and workplace culture they are situated in. Because of limited space, we describe the conversational features that are most relevant to this analysis, which are *frames* and *representations*; we refer the reader to Horn and colleagues (2015) for a fuller explanation of the other conversational features.

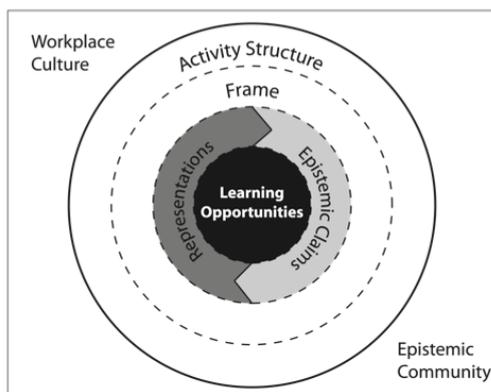


Figure 1. Conversational features that support learning opportunities (From Horn et al., 2015).

Frames refer to the ways that issues are defined in interaction (Goffman, 1974). The framings that participants employ shape the meaning of activities (Horn, 2007), changing what is salient in addressing problems of practice and thereby influencing learning opportunities. For instance, one teacher might frame a lesson-planning activity as a chance to identify questions for a whole-class discussion. A different teacher might frame the same activity as a chance to identify problems to include on a worksheet. These framings highlight different aspects of planning, and are thus consequential for OTLs.

Representations of practice refer to the ways that teachers make aspects of their teaching visible (Little, 2003). Common representations of practice include lesson plans, descriptions of past and future instruction (Horn, 2010), student work, and assessment data. Since workgroup conversations typically happen outside of instructional time, representations of practice support intersubjectivity in workgroup conversations, providing participants with a shared reference for their discussion.

We focus here on the relationship between *framings* and *representations of practice*. Participants in workgroup meetings may employ different framings of an activity, depending on their goals for the activity, their role in the meeting, and their level of expertise. On one hand, different framings may disrupt learning opportunities when participants do not negotiate shared meanings and instead talk past each other. On the other hand, different framings can support learning opportunities as teachers’ ideas are negotiated, particularly when workgroups have a trusting culture (Horn, 2007) or “re-visioning” routines (Horn, 2010). In other words, workplace culture and conversational routines can be used to scaffold teachers into new forms of practice.

When participants are scaffolded into new forms of practice, they are operating in the *zone of proximal development* (ZPD) (Vygotsky, 1978/1997). The ZPD lies at the boundary between what a person can accomplish with scaffolding versus what they can accomplish on their own. Commensurate with our focus on joint interaction, we extend this definition to the group setting (Nyikos & Hashimoto, 1997). That is, a group’s ZPD is the difference between what the group can accomplish with scaffolding and what the group can accomplish on its own. When the workgroup is scaffolded into new forms of practice—for example, by jointly engaging in a new way of lesson planning—the joint accomplishment constitutes an OTL for workgroup members because it affords changes to individuals’ practice (Greeno & Gresalfi, 2008).

If participants in teacher workgroups employ different framings of activity, it can be difficult to establish joint interaction and collectively operate in the group’s ZPD. Facilitators can attend to this issue by supporting teachers to engage the same framing. In this analysis, we address the following question: *How are representations of practice used to scaffold teachers’ participation in workgroup conversations?* Our central claim is that facilitators can leverage representations of practice to scaffold workgroups in new ways of participating in activities. The use of representations of practice as a scaffold is particularly salient when teachers and facilitators employ different framings of an activity.

Methods

Research context

The data under analysis come from an eight-year design-research study of how districts support improvement of mathematics instruction at scale (Cobb, Jackson, Smith, Sorum, & Henrick, 2013). Beginning in 2007, our team purposively sampled four large, urban districts for their commitment to improve the quality of middle-school mathematics instruction. Within each district, schools were sampled representatively with respect to their capacity for instructional improvement. In order to study teachers' learning opportunities, we selected teacher workgroups for close study based on the presence of a catalyst for teachers' learning (e.g., individuals with pedagogical expertise or other unusual resources). We conjectured that such catalysts would spark additional opportunities for teachers to learn about improving their instruction. The focal workgroups in this study were selected for their facilitators' unusual mathematics and pedagogical expertise.

Data and analysis

The data analyzed in this study come from videotaped meetings of two focal workgroups at different middle schools in the same district. The first focal group consisted of the 7th-grade math teachers at Silver Pond Middle School: Ngozi, Maria, and Peter (all proper names are pseudonyms). Jane Shepley, a school-based instructional coach and former math teacher, facilitated these meetings. Over the course of two meetings, teachers planned and debriefed a lesson on proportional reasoning. The second focal group consisted of the 6th-grade math teachers at Riverview Middle School (Rachel, Crystal, and Devon). Principal Vera Cardwell, a former math teacher and instructional coach, facilitated these meetings. During these sessions, teachers used data from a mid-year assessment to plan instruction.

The facilitators for each of these focal groups, Jane and Vera, demonstrated unusually high mathematical and pedagogical expertise. Both were middle school mathematics teachers (for 6 and 10 years, respectively) and had experience supporting math teachers as instructional leaders (for 8 and 6 years). In interviews, they also described ambitious visions of high-quality math instruction (Munter, 2014) and expressed goals for supporting teachers' learning to improve their practice. The teachers in each focal group, however, demonstrated moderate mathematical and pedagogical expertise. They tended to engage in traditional forms of instruction as measured by the Instructional Quality Assessment (Boston & Wolf, 2005) and described less-ambitious visions of high-quality math instruction than their facilitators. We characterize both workgroups as cases in which we expect to see rich learning opportunities because of the expertise of the facilitators.

Qualitative analysis

Our primary unit of analysis is an episode of pedagogical reasoning (EPR; Horn, 2007). We define EPRs as topically-bounded units of talk where participants reason about an issue of instruction. We first identified EPRs within each workgroup's meetings. We then noted that some EPRs afforded richer OTLs than others (Horn & Kane, 2015). Using multimodal interaction analysis (Jordan & Henderson, 1995), we analyzed OTLs, paying attention to the conversational features available: activity structures, frames, epistemic claims, and representations of practice (Figure 1). In many EPRs, the teachers and facilitators employed different frames as they engaged in workgroup activities. Across these episodes, we noted that there was wide variation in learning opportunities; we examined patterns of other conversational features to explain this phenomenon.

Findings

In the episodes excerpted below, we illustrate how facilitators drew on representations of practice to scaffold teachers into more ambitious forms of participation. In all four episodes, the facilitators' framings emphasized the importance of attending to students' thinking. In EPRs 1 and 3, the teachers' learning opportunities were limited by different framings and inadequate representations. However, in EPRs 2 and 4, stronger representations of practice were used to scaffold teachers to engage with the facilitator's framings.

EPR 1: Misaligned framing for lesson planning

EPR 1 illustrates the limited learning opportunities that can occur when participants employ different framings. At Silver Pond, Jane led the 7th-grade math teachers in planning the "Orangey" task, a lesson on proportional reasoning (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2009). The task asks students to compare different mixtures of water and orange juice concentrate (e.g., Mixes A, B, C, and D; Figure 2) to decide which mixture is most "orangey."

Mix A		Mix B		Mix C		Mix D	
2 cups	3 cups	5 cups	9 cups	1 cup	2 cups	3 cups	5 cups
concentrate	cold water						

Figure 2. The “Orangey” task from the Connected Math Project (Lappan et al., 2009).

Jane suggested that the teachers introduce the task by having students taste undiluted concentrate. By introducing the task in this way, teachers would elicit students’ understanding of the contextual features of the task—what is meant by “orangey.” Jane asserted this framing in Turn 1.1 (framings underlined for emphasis throughout), drawing attention to the pedagogical function of the “launch” or introduction: to prepare students to engage in the mathematical work of the task (Jackson, Shahan, Gibbons, & Cobb, 2012).

- 1.1 Jane: By letting ‘em taste. They’re like [*makes sour face*] That’ll help them think about what a concentrate is. And then, move into the discussion of “What is the process to make this drinkable? We have to start adding water. How much water do we need to add?” Well, these students did this, this, and this. Is that gonna be enough of a launch to get us to where they need to be in order to solve the questions that we’ve already, well, we know they’re gonna end up having to do? What you think?

Jane’s framing of the launch emphasized helping students understand the contextual features of the task as a way of preparing them to engage in the pending mathematical task. Framed in this way, the launch is an important aspect of ambitious teaching as it supports all students to engage in mathematics by lowering barriers for participation. However, the teachers employed a framing that emphasized logistical concerns:

- 1.2 Ngozi: Um, really by time they make up these mixtures in real life, it takes a lot of time. Because then, everybody wants to go through each station to taste because then, the colors are gonna look the same because they don’t taste it.
- 1.3 Jane: Okay.
- 1.4 Ngozi: So, what I would suggest or what I did in the past was we, probably, this is like a reward. This is our after thing.

Ngozi’s response to Jane’s questions asserted a framing that emphasized logistical concerns. Instead of letting students taste mixtures or the concentrate as part of the launch (which “takes a lot of time”), Ngozi suggested that it become a “reward” activity to do after the main lesson. This framing fundamentally shifts the meaning of the planning activity by emphasizing logistics over student thinking.

Jane responded by reasserting a framing that emphasized preparing students for the mathematical task by familiarizing them with its contextual features. She clarified her proposal: kids should get a small taste of the concentrate “so that they know what [concentrate] means and what that is.” However, the other teachers engaged with Ngozi’s framing, raising additional logistical concerns like whether diabetic students should taste the concentrate or what brand of concentrate they should buy.

Taken together, these data illustrate a misalignment between the framings employed by Jane and the teachers. Though Jane consistently pressed the teachers to attend to pedagogical concerns, the teachers consistently responded with logistical concerns. The teachers raised valid issues—certainly, students’ health and efficient use of time are important—but their responses shifted the meaning of the activity in ways that were unlikely to lead to instructional improvement. We saw this pattern throughout many of the Silver Pond’s 7th grade planning meetings. This limited the workgroup’s participation in a new form of lesson planning that actively anticipates students’ understanding of the contextual features of tasks.

EPR 2: Using student work as a scaffold into ambitious practice

Whereas EPR 1 showed how misaligned framings can disrupt teachers’ learning opportunities, EPR 2 illustrates how representations of practice can be used to overcome misalignment and scaffold a workgroup’s participation in forms of practice that lie in its ZPD. A few days after the planning meeting described in EPR 1, the Silver Pond teachers taught the Orangey task; during the meeting described here, they debriefed the lesson. Consistent with the framing she employed in EPR 1, Jane asked the teachers to reflect on whether the introductory activities (e.g., the launch) were “enough to tell if our students were ready” to successfully participate in the mathematical task.

Initially, teachers gave superficial replays (Horn, 2010) of instruction that elided students' contributions. Unsatisfied with their responses, Jane reasserted her framing and drew teachers' attention to a concrete representation of practice: posters of student work.

- 2.3 Jane: We want to look at this work using the lens of a teacher. What does it tell me, and what do I need to do moving forward? Let's take a few minutes to look at the work samples, and see if we're ready to move forward. [*teachers turn to face student work*] Do we need to remediate? Reteach something? [*walks toward posters, gestures to them*]

In this turn, Jane used the framing of "the lens of a teacher," highlighting the importance of interpreting students' work before planning future instruction. Once attuned to the representations of student thinking, the teachers took up Jane's framing and focused on discerning what students understood about their solution strategies. Figure 3 shows one student group's solution. Maria noted that the students found a common denominator for 3, 9, 2, and 5 and scaled up each ratio to a denominator of 90. Ngozi agreed and then extended Maria's observation, wondering whether her students understood how they were comparing the ratios:

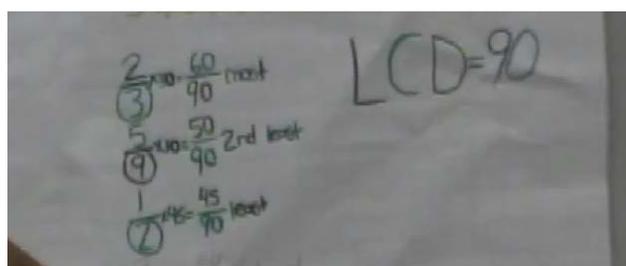


Figure 3. Poster showing students' solution strategy for comparing the ratios 2:3, 5:9, 1:2, and 3:5 (not shown).

- 2.8 Ngozi: Yeah. They got the same number of cups of water, but do my students know that? I don't know.
- 2.9 Jane: That's a good question.
- 2.10 Maria: They categorized them. Which one was the most.
- 2.11 Jane: She's saying, do they understand that when they scaled up here—
- 2.12 Maria: Why did they do it?
- 2.13 Jane: When they scaled up, do they notice that they all have the same number of cups of water, and the one that I'm picking has the most cups of concentrate.

In this exchange, Jane ratified and encouraged the workgroup's sensemaking. Through their line of inquiry (Turns 2.8, 2.12), Ngozi and Maria engaged Jane's framing of the activity. They went on to discuss whether students knew they were using a part-to-part comparison (i.e., concentrate to water, as opposed to concentrate to total volume). The teachers then prepared for future work by rehearsing (Horn, 2010) questions to ask students in upcoming lessons. Ngozi further clarified that she would "expect them to say concentrate/water." Jane agreed, gesturing at the posters and wondering, "So do they know what this stuff is?"

In this EPR, teachers used representations of practice (i.e., student work) to make inferences about students' understanding. In doing so, they jointly participated in the activity of lesson debriefing in a way that engaged Jane's framing. Alternatively, one might conjecture that this form of participation can be attributed to the fact that the teachers frame planning and debriefing differently. Yet in other meetings, the workgroup participated in debriefing activities without adequate representations, and the teachers' engaged framings that emphasized logistical concerns. We argue that joint participation in the activity of debriefing (EPR 2) was fundamentally altered by the group's use of representationally adequate artifacts of student thinking. The posters of student work scaffolded the workgroup's participation in new forms of debriefing that were in the group's ZPD. This participation constituted learning opportunities for teachers that support instructional improvement.

EPR 3: Inadequate framing for data use

Our third EPR took place at Riverview during a meeting to discuss data from a multiple-choice assessment. Vera, the principal, provided teachers with a variety of representations, including a list of assessment items showing the percentage of students that answered each item correctly, the distribution of students' responses on each item,

copies of the assessment, copies of the state math standards, and a lesson planning template. To start the meeting, Vera framed the activity by telling teachers to use assessment data to plan instruction:

Vera: Look at the item, study the [standard], what part of the [standard] was addressed, what did kids struggle with, BAM, that's your finding.

In this short sentence, Vera framed a sophisticated data use process that emphasized determining student thinking (“what did kids struggle with”) to design an instructional response (“your finding”) to support student learning. To participate fully in this practice, teachers needed to synthesize across a variety of representations to interpret how students may have approached assessment items on many different mathematical concepts.

As the teachers began analyzing data, they engaged with Vera’s framing of the activity. They began this episode by identifying an item that students did poorly on. The standard (fraction, decimal, and percent conversion) is “one that [students] are still struggling on,” according to Devon. He told Crystal to read the item:

- 3.5 Crystal: “Jeremy bought a skateboard on sale for \$28, which was 12.5% off the original price. What was the discount as a fraction of the total price?” Well all they need to know is the 12 and a half percent.
- 3.6 Devon: The majority of students chose B. [12/5]
- 3.7 Crystal: Why?
- 3.8 Devon: 12 over 5
- 3.9 Crystal: That's crazy
- 3.10 Rachel: All they did was take the numbers and—that makes me mad [*laughs*]

Though the teachers engaged in Vera’s framing of the data use activity—they looked at the item, the standard, and how students responded to it—their explanation for the source of students’ difficulties was not strongly linked to students’ mathematical thinking or issues of instruction. Crystal went on to say that the students “are not confident, and they don’t know it.” Rachel concluded that “We need to do some problems with a percent that has a decimal in it” before moving on to consider other test items. This instructional response (providing more practice problems) is unlikely to support more ambitious forms of instruction that address students’ conceptual understandings of fraction, decimal, and percent relationships.

The Riverview teachers had many material representations available to analyze assessment data, but were not able to participate in more ambitious forms of data use. They engaged with a thin, almost mechanistic version of activity as Vera framed it. Admittedly, this framing was described in vague and general terms; Vera did not specify what it meant to “look” at an item or how to determine what “kids struggled with.” We argue that her framing was insufficient to scaffold teachers into new forms of practice. Furthermore, the representations of student learning available to the Riverview teachers (tables of assessment data) were distal to instructional practice, particularly when compared to the representations available in EPR 2 (posters of student work). This made it more challenging for teachers to plan future instruction in ways that accounted for student thinking. Therefore, participation in ambitious data use practices was limited in this conceptually sparse space.

Episode 4: Coordinating evidence to scaffold participation

In EPR 4, Vera joined the teachers’ conversation and provided additional representational resources that scaffolded the workgroup’s participation into more sophisticated data use practices. She elaborated on her framing—much like Jane’s elaboration on “using the lens of a teacher” in EPR 2—by modeling her method of data use and coordinating multiple representations of practice as evidence of student thinking.

Devon asked Vera to review the same assessment item the teachers examined in EPR 3. When Vera joined the conversation, she immediately asked a series of questions to orient herself, identifying the standard, the item, teachers’ instructional approaches, and the most commonly selected answer choices. Much of this paralleled the workgroup’s discussion above. But as Vera interpreted the data, she added additional representations of practice (replays and rehearsals) to connect the material representations to instruction:

- 4.31 Vera: Okay, so I would say, then, this is probably, I mean, my gut it could be a combination of the .5, however, I, it's probably rooted in the question.
- 4.32 Crystal: I thought, that's what I thought, because
- 4.33 Vera: You know what I mean? They didn't realize that you were finding the equivalent form of a number. They might even try to

- 4.34 Crystal: Do something with the 28, they don't know how to omit the non-necessary information

Vera considered multiple sources of student misunderstanding; she acknowledged that both the decimal in the percent (i.e., 12.5% as opposed to 12%) and the wording of the question may have been difficult for students (Turn 4.31). Crystal elaborated on Vera's explanation, suggesting that students "don't know how to omit the non-necessary information" (Turn 4.34). However, Rachel pushed back, asserting that "they have not dealt with decimals in percents enough." Vera then offered a rehearsal of a way to solicit more evidence about students' understanding (Turn 4.37) and supported it with a replay of her own instruction (Turn 4.39):

- 4.37 Vera: So what I would ask them, then, if you want to deduce that right? And you want, you know, to determine that that definitely is, then the best way is to have every kid take okay, 12.5%, represent it as a fraction. Take all the words out.
- 4.38 Rachel: Mmhmm. And they don't know how to do that.
- 4.39 Vera: And that, and that will tell you immediately whether or not that was one of the layers of problem. But I know when we we talked about, um, an-an-and those are common ones that we usually go over? Like, okay, when I was a 6th grade teacher, at this point in the time, I don't know if I would say that that was overwhelmingly all of my kids' issues. Because, I had a chart they filled out every single day, like, starting in January, that showed 12.5, I want it as a fraction, I want a decimal, you know, I want it as uh, uh, a percent, or whatever, my conversions. So we would include those.

In this exchange, Vera built a nuanced representation of multiple ways in which students may have approached the test item. She drew upon conceptual representations of instruction (replays and rehearsals) to augment and connect to physical representations of student learning. When Rachel re-asserted concerns about the decimal in the percent, Vera used the rehearsals (Turns 37, 39) to scaffold the workgroups' participation in developing an alternate instructional response that addressed multiple "layers of problem." By using conceptual representations of practice (replays and rehearsals) to augment thin physical representations of practice (tables of assessment data), Vera scaffolded the workgroup into more ambitious forms of data use practice and mobilized the teachers for future instruction that accounted for students' thinking.

Conclusions and implications

Across these episodes, we see that misaligned or inadequate framings of workgroup activities can disrupt teachers' opportunities to learn more ambitious forms of practice (EPRs 1, 3). However, facilitators can leverage representations of practice to scaffold teacher workgroups into forms of participation that are within the group's ZPD (EPRs 2, 4). As workgroups engage in new forms of practice, participants can develop concepts of pedagogy and prepare to engage in more ambitious instruction that attends to students' mathematical thinking. Thus, representations of practice can serve as a resource for teacher workgroups to use in efforts to improve instruction.

This study builds on previous literature that has shown how certain workgroup features, like workgroup culture (Horn, 2007) or conversational routines (Horn & Little, 2010), can be used to support teachers' learning opportunities in workgroups where participants employ different framings. When these conversational features align (i.e., afford the same goals of the activity), workgroups can engage in new forms of practice. When there is misalignment across the conversational features learning opportunities are significantly more limited.

This work has important implications for the development of teacher workgroups as sites for professional learning. In particular, facilitators must attend to issues of teacher learning in collaborative activity. Facilitators can support teachers' learning opportunities by framing workgroup activities in terms of ambitious mathematics instruction (e.g., considering student thinking in planning for instruction), and representations of practice can serve as key supports to attune teachers' attention to issues of pedagogy and to scaffold workgroups' engagement in new forms of practice. Future work is needed to investigate how to support facilitators to (a) recognize when workgroup participants employ framings that are not aligned with goals for instructional improvement and (b) employ activity structures and conversational features that can scaffold workgroups into more ambitious forms of practice.

References

Ball, D. L., & Cohen, D. (1999). *Developing Practice, Developing Practitioners: toward a practice-based theory*

- of professional education. In G. Sykes & L. Darling-Hammond, *Teaching as the learning profession: Handbook of policy and practice* (pp. 3–32). San Francisco: Jossey Bass.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3–15.
- Boston, M., & Wolf, M. K. (2005). *Assessing Academic Rigor in Mathematics Instruction*. Center for the Study of Evaluation, National center for Research on Evaluation, Standards, and Student Testing.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing Schools for Improvement*. University of Chicago Press.
- Cobb, P., Jackson, K., Smith, T., Sorum, M., & Henrick, E. (2013). Design research with educational systems: Investigating and supporting improvements in the quality of mathematics teaching and learning at scale. In W. Penuel, B. J. Fishman, & B. H. Cheng, *National Society for the Study of Education Yearbook on Design-Based Implementation Research*.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.
- Greeno, J. G., & Gresalfi, M. S. (2008). Opportunities to Learn in Practice and Identity. In Assessment, Equity, and Opportunity to Learn. Cambridge University Press.
- Hall, R., & Horn, I. S. (2012). Talk and Conceptual Change at Work: Adequate Representation and Epistemic Stance in a Comparative Analysis of Statistical Consulting and Teacher Workgroups. *Mind, Culture, and Activity*, 19(3), 240–258. doi:10.1080/10749039.2012.688233
- Horn, I. S. (2007). Fast kids, slow kids, lazy kids: Framing the mismatch problem in mathematics teachers' conversations. *The Journal of the Learning Sciences*, 16(1), 37–79.
- Horn, I. S. (2010). Teaching replays, teaching rehearsals, and re-revisions of practice: Learning from colleagues in a mathematics teacher community. *The Teachers College Record*, 112(1).
- Horn, I. S., & Little, J. W. (2010). Attending to Problems of Practice: Routines and Resources for Professional Learning in Teachers' Workplace Interactions. *American Educational Research Journal*, 47(1), 181–217. doi:10.3102/0002831209345158
- Horn, I. S., & Kane, B. D. (2015). Opportunities for Professional Learning in Mathematics Teacher Workgroup Conversations: Relationships to Instructional Expertise. *Journal of the Learning Sciences*, 24(3), 150415095906006. doi:10.1080/10508406.2015.1034865
- Horn, I. S., Kane, B. D., & Wilson, J. (2015). Making Sense of Student Performance Data: Data Use Logics and Mathematics Teachers' Learning Opportunities. *American Educational Research Journal*, 52(2), 208–242. doi:10.3102/0002831215573773
- Jackson, K. J., Shahan, E. C., Gibbons, L. K., & Cobb, P. A. (2012). Launching complex tasks. *Mathematics Teaching in the Middle School*, 18(1), 24–29.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39–103.
- Lampert, M., Boerst, T. A., & Graziani, F. (2011). Organizational resources in the service of school-wide ambitious teaching practice. *The Teachers College Record*, 113(7), 1361–1400.
- Lappan, G., Fey, J. T., Fitzgerald, W., Friel, S., & Phillips, B. (2009). *Comparing and Scaling: Ratio, proportion, and percent*. Boston, MA: Pearson.
- Little, J. (2003). Inside teacher community: Representations of classroom practice. *The Teachers College Record*.
- McLaughlin, M. W., & Talbert, J. E. (2001). *Professional Communities and the Work of High School Teaching*. University of Chicago Press.
- Munter, C. (2014). Developing Visions of High-Quality Mathematics Instruction. *Journal for Research in Mathematics Education*, 45(5), 584–635.
- Nyikos, M., & Hashimoto, R. (1997). Constructivist Theory Applied to Collaborative Learning in Teacher Education: In Search of ZPD. *The Modern Language Journal*, 81(4), 506–517.
- Supovitz, J. (2012). Getting at Student Understanding—The Key to Teachers' Use of Test Data. *Teachers College Record*, 114, 1–29.
- Vygotsky, L. (1997). 4: Interaction Between Learning and Development. In *Readings on the Development of Children* (pp. 1–9). Macmillan.
- Wilson, S. M., & Berne, J. (1999). Teacher Learning and the Acquisition of Professional Knowledge: An Examination of Research on Contemporary Professional Development. *Review of Research in Education*, 24, 173–209.

Acknowledgments

This work funded by the NSF under grant numbers ESI-0554535 and DRL-1119122, and the Spencer Foundation under grant number

Self-Directed Learning in Science Education: Explicating the Enabling Factors

Khe Foon Hew, Nancy Law, Jarrad Wan, Yeung Lee, and Amy Kwok
kfhew@hku.hk, nlaw@hku.hk, jarrad@ilap-sdls.cite.hku.hk, yeunglee@hku.hk, amykwok@hku.hk
The University of Hong Kong

Abstract: The notion of self-directed learning (SDL) has gained prominence in many education institutes around the world. However, few studies detail the specific enabling factors that could foster SDL in primary and secondary school science education. Focusing on this gap, we engaged in a mixed methods study that examined 17 schools in Hong Kong, consisting of 10 primary and 7 secondary schools. Altogether, 1,538 students from 55 classes and involving 53 teachers participated in the study which ran from September 2014 to June 2015. Building upon four key SDL representation models, we first developed our three-dimensional SDL conceptual framework to guide our analysis of the specific leadership-, teacher-, and student-related factors that could foster SDL in science education. Five key enabling factors are described.

Introduction

The notion of self-directed learning (SDL) is not new. Originally conceptualized as a desired competence in adult education, Knowles (1975, p. 18) defined SDL “a process in which the individual takes the initiative, with or without the help of others, in diagnosing their learning needs, formulating learning goals, identifying human and material resources for learning, choosing and implementing appropriate learning strategies, and evaluating learning outcomes”. In recent years, however SDL has been recognized as an important element in school age students’ learning due to the realization that many students today study merely for the sake of examinations (that is, being “exam smart”) but lack the skill to monitor one’s own learning, and rely heavily on the guidance of teachers (Hew & Cheung, 2012).

In a survey of 431 employers representing more than two million employees, almost 60% of employer respondents rated secondary school graduates as being deficient in skills related to SDL (Partnership for 21st Century Skills, 2006). The results for college graduates were even more depressing – only 25% of employers rated four-year college graduates as excellent in SDL. Even though the survey results focused on the USA workforce, and thus may not be generalized to other parts of the world (no such large scale study has been conducted in Asia), the results do highlight concerns about graduates’ ability for SDL, and consequently their readiness to enter the workforce in the long term.

In this study, we address the central question: *What are the specific factors that could foster SDL?* Previous research on SDL has often been limited to the college or university contexts. Few studies detail the specific enabling factors of SDL in primary and secondary school. Focusing on this gap, we engaged in a mixed methods study to extract specific factors that could foster SDL in science education. A total of 1,538 students and 53 teachers from 17 different schools in Hong Kong participated in this project. We chose science education as our context of implementation because of its importance to the global economy, and problem solving ability which is considered one of the key 21st skills. We argue that science education is not merely about facts acquisition, but a process of inquiry which requires students to evaluate and reflect on their own thinking related to the activities of planning, measuring, observing, analyzing data, and examining evidence. Self-directed learning as an instructional strategy provides for a large degree of student choice and allows students to formulate, investigate, and reformulate problems and solutions of their own design (van Merriënboer & Sluijsmans, 2009). Hence, when self-directed learning is implemented with appropriate support for students, it can help students develop the inquiry process critical to science education.

The rest of the article is organized as follows. First, in the conceptual framework section, we describe our three-dimensional model for understanding SDL. We used the conceptual model to guide our analysis of the specific factors that could foster SDL in science education. We then describe the methodology of the study, followed by the results and discussion. We conclude by highlighting some limitations of this study and several suggestions for future research.

Conceptual framework

Many models have been proposed to understand SDL, beginning with Mocker and Spear’s *Two Dimensional Model* in the early 1980s to a more recent model from Song and Hill (2007). We found that, often, what is lacking in one model can be found in another one. Therefore, in this section, in order to develop a more comprehensive

three-dimensional model to help us understand the factors that might foster SDL in science education, we drew largely upon the writings of Brockett and Hiemstra (1991), Candy (1991), Garrison (1997), and Song and Hill (2007). These four works were selected because they appear to be among the most cited works in SDL. However, other writings were also consulted where necessary to provide supplementary or background information. Our SDL conceptual model includes three connected dimensions: personal attributes, autonomous processes, and learning context (see Figure 1).

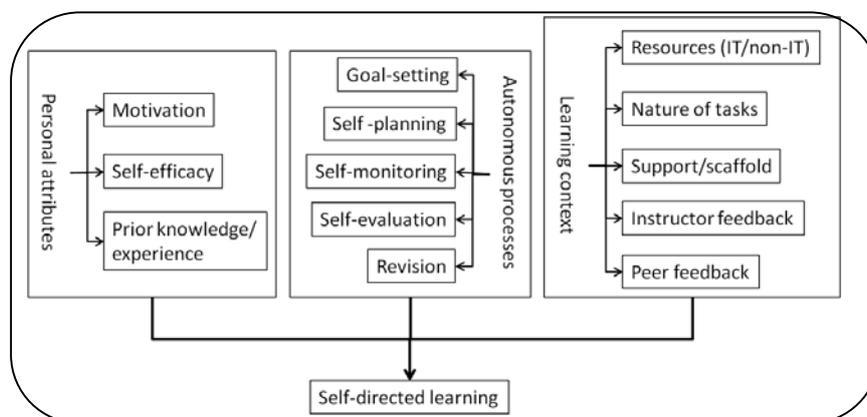


Figure 1. A three-dimensional SDL framework

Personal attributes play an important role in the initiation and maintenance of SDL. They refer to the characteristics an individual (teacher or student) bring to a learning context (Song & Hill, 2007). More specifically, personal attributes include an individual’s motivation (Brockett & Hiemstra, 1991; Garrison, 1997), self-efficacy (Brockett & Hiemstra, 1991), and prior knowledge of the content area or prior experience with the learning context (Song & Hill, 2007). Motivation may refer to intrinsic or extrinsic factors. Individuals will have a higher intrinsic motivational state if they perceive that SDL will meet their personal needs (values) and affective states (preferences) (Garrison, 1997). Extrinsic motivation (e.g. reward) is equally important, as it is found to directly enhance an individual’s intrinsic motivation to engage in activities (Yoo et al., 2012). Self-efficacy or the belief of one’s ability to accomplish a task will also help inspire and sustain an individual’s engagement in SDL (Garrison, 1997). An individual’s prior knowledge of the content area or prior experience with SDL will also affect his or her self-direction in learning (Song & Hill, 2007). Individuals who have engaged in SDL before will find it easier to diagnose their learning needs, and formulate learning goals as compared to those who are unfamiliar with it.

Autonomous processes refer to the individual’s freedom of choice in SDL (Candy, 1991). Specifically, learner autonomy can be manifested in the processes of goal-setting, self-planning, self-monitoring, self-evaluation and revision (Brockett & Hiemstra, 1991; Candy, 1991; Garrison, 1997; Song & Hill, 2007). Some possible indicators of each process are presented in Table 1.

Table 1: Possible indicators for autonomous processes

Process	Description of possible indicators
Goal setting	<ul style="list-style-type: none"> Students identify own learning goals & learning activities
Self-planning	<ul style="list-style-type: none"> Students plan for the detailed decisions and arrangements associated with own learning
Self-monitoring	<ul style="list-style-type: none"> Students manage their own time Students monitor their own learning strategies Students adjust their own learning pathway as they progress
Self-evaluation	<ul style="list-style-type: none"> Students are aware of the assessment criteria Students critically evaluate work according to set criteria
Revision	<ul style="list-style-type: none"> Students revise their work based on the feedback received from their teacher or peers Students reflect and apply what they have learnt to new contexts

Learning context focuses on environmental factors and how these may impact the level of SDL provided to the learner (Song & Hill, 2007). There are various factors in a learning context that can impact a learner’s SDL experience; these in particular include resources (IT or non-IT), nature of the learning tasks, support or scaffold provided, instructor feedback, and peer feedback (Song & Hill, 2007). Resources can take different forms which can be parsimoniously grouped into IT (information technology) resources (e.g., Internet sites), and non-IT

resources (books). Students, for example, can use Google or Yahoo! search engines to independently access additional reading materials. Another set of elements in the learning context that can impact a learner's SDL is the nature of the learning tasks, and the support given. Open-ended tasks are activities such as problems or questions that have more than one possible solution, or answer. Such activities lend themselves very well to SDL because they can help spur learners to consider various perspectives, search for different answers, and evaluate them. Since not all students are familiar with SDL, support or scaffold should be provided by the teachers. For example, asking students to create a Know-Want-Learn (KWL) chart in which they list what they know; what they wish to know; and later what they have learned can increase student learning (Campbell & Campbell, 2009), as well as metacognitive skill (Tok, 2013), an essential skill that can facilitate SDL. A KWL chart can also guide student independent thinking as it facilitates the personal pursuit of learning through exploration of the topics that students want to know. Finally, although SDL focuses primarily on the individual learner, external feedback still plays an important role. Feedback is central to student learning because it can point out errors and suggest areas for improvement so that students can move their learning forward. Both instructor and peer feedback are useful. A teacher's feedback is useful because it helps to focus students' discussion on the topic, prevent possible conflicts in the discussion, and provide useful pertinent information, while peer feedback allows students to share their views more openly (Hew, 2015).

Method

We conducted a mixed methods study using the three-dimensional SDL framework (figure 1) as an underlying factor model to analyze and extract the specific leadership-, teacher-, and student-related elements that could foster SDL. Altogether, 17 schools in Hong Kong participated in this funded project that ran from September 2014 to June 2015 (see Table 2). A total of five professional development workshops were conducted for these 17 schools. Each workshop had a specific theme for teachers to better understand the different aspects of SDL through sharing from experienced educators and hands-on activities. The specific themes included: (a) what is SDL? (b) learning and assessment design, (c) learning analytics, (d) drawing conclusion for scientific investigation, and (e) sharing of useful tools and strategies. Besides these five professional development workshops, 13 cluster meetings were also held. In these meetings, school teachers from the same cluster (e.g., north-west zone) discussed their SDL plans with the university project consultants and with other teachers.

During the cluster meetings, theoretical and pedagogical suggestions, as well as practical exemplars were given by the University project staff to teachers. In addition, teachers who were more experienced in SDL implementation acted as models to guide and review their SDL plans learning design? i.e. web-based teaching plan. In addition, various onsite training and support sessions were also conducted. From September to November 2014, the University project staff members visited participating schools and conducted teacher and student training workshops on the use of an interactive and assessment platform (iLap) (see Figure 2).

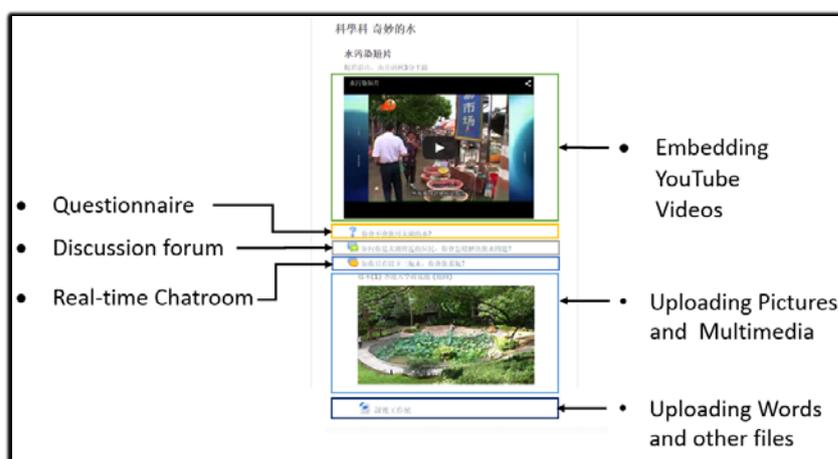


Figure 2. An interactive and assessment platform (iLap)

Specifically, iLap (interactive Learning and assessment Platform) is a Moodle-based Learning Management System (LMS) which contains various tools such as *mind maps* for students to express individual understanding for concepts learnt, *wiki pages* for students to share their learning, *blogs* for students to reflect and share what they had learnt, *discussion forums* for teachers to monitor students' understanding of concepts, and *knowledge forums* for teachers and students to engage in collative inquiry. Onsite SDL co-planning (n=24) and

lesson observation (n=18) sessions were also conducted throughout the duration of the project. These sessions focused on helping teachers in SDL implementation, notably goal setting, self-planning, self-monitoring, self-evaluation and revision. With the guidance of the University project staff, the teachers of each participating school collaborated and contributed a lesson unit that incorporated the key stages of SDL in science education. The school then hosted participants from other schools in the same cluster to attend a peer lesson observation session, as well as a debriefing meeting where principals, teachers and University project staff were present to review the SDL implementation.

Of the 17 participating schools, 10 good practice schools were further selected based on a panel of three University project staff. The three staff members were responsible for all aspects of the project from beginning to end such as providing professional development workshops for teachers, attending cluster meetings, onsite lesson observations, and interviewing principals, teachers, and students. These 10 schools were chosen based on ability of their lessons to promote scientific inquiry and sustaining SDL such as enabling students to identify learning strategies to achieve the learning goals, to set the standards for the achievement of their learning goals, to formulate questions and generate relevant inquiries, to reflect on their learning and initiate gathering of feedback from teachers and peers, and to apply what they have learnt to new contexts. In this study, we are interested to address the following question: *What are the specific factors found in these 10 good practice schools that could foster SDL, as compared to the other schools?*

Table 2: Summary of the 17 schools

School	Level	No. of classes involved in SDL	No. of students involved in SDL	No. of teachers involved in SDL	Principal participation**
School A	Primary	2	53	2	2
School B	Primary	4	98	4	3
School C*	Primary	2	51	3	5
School D*	Primary	4	120	4	5
School E*	Primary	4	94	4	2
School F	Primary	2	80	2	2
School G*	Primary	5	145	5	3
School H	Primary	3	88	3	4
School I*	Primary	5	149	5	6
School J	Primary	4	117	2	4
School K*	Secondary	5	127	2	3
School L*	Secondary	2	64	4	3
School M	Secondary	1	30	1	4
School N*	Secondary	4	116	2	4
School O*	Secondary	4	124	2	2
School P*	Secondary	2	28	4	6
School Q	Secondary	2	54	4	3

*denotes good practice school; **principal participation refers to the number of times principals participated in project activities such as attending workshops, and open class observations.

Data collection and analysis

Observation of science lessons

Altogether a total of 34 science lessons were observed, two lessons in each of the 17 schools. During observations, a record of events that included the lesson objectives, lesson sequence, types of IT and non-IT tools used, and roles of the teachers and students were kept. The analysis of data proceeded alongside the collection of observational data. Preliminary analytic notes were jotted down and provided inputs for the subsequent interview sessions after each lesson.

Interviews with teachers and principals

After each lesson observation, debriefing interviews with the teachers were held. The main purpose of such interviews was to understand the teachers' reasons for planning and implementing certain activities, reasons for using IT and non-IT tools, roles of the participants, and procedures to engage students. In addition, the teachers were asked to discuss the problems they faced in the lessons with respect to scientific inquiry and SDL, and how they might address these problems. The principals from each school were also interviewed twice for about one hour each, before and after the project. The list of topics generated for the purpose of these interviews included: objectives of the school with respect to SDL and scientific inquiry implementation, personal views of SDL and

how IT may foster it, and roles of the principals or administrative structure to facilitate the project implementation. The interview data were audio-recorded and transcribed for data analysis using Corbin and Strauss' (2008) grounded approach.

Focus group discussions with students

After each lesson observation, 5 students on average were chosen from each school for the focus group discussions. The list of topics for the interviews included: students' understanding of SDL, and the use of resources and how they facilitated SDL and scientific inquiry. The students' focus group data were also audio-recorded and transcribed for data analysis using Corbin and Strauss' (2008) grounded approach.

Findings and discussion

Taking the three-dimensional SDL framework (figure 1) as an underlying factor model, the main findings highlighted five key enabling factors that fostered SDL in science education: (a) the nature of tasks, (b) support or scaffold for teachers, (c) teacher motivation for SDL, (d) students' prior content knowledge, and (e) principal leadership behaviors.

Probably one of the most important factors that distinguished the good practice schools from the others was the nature of the student tasks. We found that schools with greater success in promoting student SDL (e.g., students identifying learning strategies to achieve the learning goals, formulating questions, or applying what they have learnt to new contexts) tended to utilize more *open-ended inquiry activities with real-life relevance*. The results of this study support Douglass and Morris' (2014) finding that teachers' use of real world experiences was a key factor that helped students' SDL. These activities were related to specific science principles found in the curriculum; however, they required students to apply these principles that have some real-world significance. For example, School C (a good practice school) gave students the task of building the strongest fan with the voltage being held constant. Students in groups were observed planning various experiments involving different materials, as well as the number of fan blades in an attempt to outdo other groups. Another good practice school, School P gave students the task of water filtration:

Teacher A: I wanted students to do SDL on water pollutant, water purification and sewage treatment. Student brought their own pollutant to school and polluted the water. They had to classify the pollutant into soluble, insoluble or microbes. The next lesson they had to *purify the water and drink it*. They pored over their textbooks and even passive students sent me WhatsApp to ask questions. They designed the setup with their own apparatus. They ran filtration followed by distillation. Students were very happy for this whole activity.

School E (another good practice school) gave students the task of building a 15cm paper bridge strong enough to hold heavy things. Students in groups had to engage in self-planning, self-monitoring and evaluation, and revising their bridge designs in order to achieve the set goal. Through these processes, students were able to learn several interesting things about paper:

Student B: I learnt that creased paper cannot be reused. Also, a light paper can be strong enough to carry the weight.
Student C: I am amazed that thin paper can withstand the heavy weight, folded paper can be strong.
Student D: We achieved the goal. We managed to place 6 cans of Cola on the bridge.
Student A: I find learning from textbooks is boring, but doing SDL by ourselves is much more interesting.

On the other hands, most schools that relied on text-book type activities were not able to promote students' interest and effort for SDL as much as the good practice schools.

Teachers play a pivotal role in determining the success of any learning endeavor. Since almost all teachers were not familiar with the notion of SDL at the beginning of the project, support for (and support among) teachers was absolutely essential. Specifically, we found that teachers in the good practice schools were more eager to work with the University project staff, as well as teachers from other schools in co-planning the SDL lessons. These co-planning activities served two important purposes: they bolstered teachers' self-confidence in running the lessons, and they helped the teachers share useful resources and ideas. During the co-planning activities, teachers particularly needed support with respect to the following three areas: (a) pedagogical knowledge of SDL concerning the processes of goal-setting, self-planning, self-monitoring, self-evaluation, revision, and ways to assess students' work, (b) technological knowledge of tools such as how discussion forums work, and (c) the interplay between SDL pedagogical knowledge and technological knowledge. For example, the

use of video recorders and forums afford the processes of self-monitoring and self-evaluation because since students could revisit the recorded materials and reflect on them:

Teacher R: Students can view the work of their classmates and learn from others. The teacher can also view the recorded materials and provide feedback. These would not be possible in a conventional lesson.

Teacher C: Using iLap, students could record their whole science experiments. This enabled students to review and amend their work.

Another factor that could promote SDL was the teachers' motivation for doing it. Not all teachers (including the good practice schools) were motivated in SDL for science education at the beginning of the project. Some teachers simply participated because they were asked by their principals to do so. Indeed, we found no significant differences between the good practice schools and the other schools in terms of the mere number of classes involved in SDL (Mann-Whitney $U = 17.5$, $Z = -1.791$, $p = 0.073$); the number of students involved in SDL (Mann-Whitney $U = 20.0$, $Z = -1.464$, $p = 0.143$); the number of teachers involved in SDL (Mann-Whitney $U = 20.0$, $Z = -1.533$, $p = 0.125$); and principal participation per se (Mann-Whitney $U = 25.5$, $Z = -0.952$, $p = 0.341$). There were also no significant differences the good practice schools and the other schools in terms of the number of teacher login (Mann-Whitney $U = 17.0$, $Z = -1.757$, $p = 0.079$); and student login to iLap (Mann-Whitney $U = 15.0$, $Z = -1.952$, $p = 0.051$).

However, we found that as the project progressed, more teachers in the good practice schools became more motivated in fostering SDL among their students. How did this come about? The reasons for these were mainly twofold: (a) an increase in the teachers' sense of competence or self-efficacy in designing SDL lessons, and (b) a desire to try something different. According to the Self-Determination Theory of motivation (Deci & Ryan, 2000), a strong sense of competence helps increase an individual's intrinsic motivation to accomplish a particular task. Teachers mainly gained the sense of competence as they worked with the University project staff, as well as teachers from other schools to co-plan the SDL lessons. Recall earlier our observation that teachers in the good practice schools were more eager to participate in such co-planning activities compared to the teachers from the other schools. These teachers were prepared to receive feedback, and in doing so gained confidence and competence in implementing the lessons in their actual classrooms.

Teachers in good practice schools were also more willing to try a different way to teach science. Essentially, these teachers were willing to release control of the lessons to their students rather than being the sole information provider. In other words, teachers who believed that their students could learn on their own had greater motivation in fostering SDL:

Teacher B: You need to trust the students and let them loose. I've learned to let the students lead the class. Using SDL, I don't have to prepare all the lesson materials, and I talk less.

Another teacher was observed allowing her students to plan and design for the science experiment by themselves in groups. The teacher did not intervene even though some of the designs were not appropriate. Often when teachers released control of the classroom, many were pleasantly surprised to see their students more engaged in the lessons, and learned better than or equally well compared to the usual teacher-conducted teachings:

Teacher H: I saw students learning on their own and questioning each other's opinions.

Teacher C: Students are generally interested in SDL. They showed greater understanding of the topic learned in the school exam. Compared to other non-SDL based topics, the students' performance in this SDL-run topic was very big.

Teacher F: Students were more engaged when they had the autonomy to do their own planning and research.

Teacher T: I found that students wish teachers to talk less and they themselves do more. I'm surprised that students prefer SDL and they learned more from this approach.

An individual's prior knowledge of the content area can also affect his or her self-direction in learning (Song & Hill, 2007). We found students who were more versed in the subject content tended to be more engaged in SDL. They were able to ask more interesting questions, monitor their learning progress, and make better sense of the science experiment findings. They were able to better evaluate the validity of the Internet resources assessed. Students who are stronger in the topics are also individuals who like the subject more; hence they are usually more proactive in class. Being proactive in classes is frequently cited as a means to promote SDL (e.g., Douglass &

Morris, 2014; Yazedjian et al., 2008). On the other hand, students who were weaker in the subject content were not very sure about how to proceed in the lessons, and they relied on their teachers for guidance and help:

Teacher S: I've a class of weaker students. Although they were interested in the topics, their SDL effort was limited. They did not know how to find the answers themselves, and they needed help from the teacher.

Teacher C: I assigned stronger students to different groups so that they can guide the weaker students.

Therefore, in order to promote SDL among students, it would be helpful for teachers to spend some time teaching students basic information or concepts about the topics concerned. We also found that certain principal leadership behaviors to be another key enabling factor associated with SDL implementation in the good practice schools. Specifically, principals and teachers in these schools had the same desire to try out the implementation of SDL in science education, and the desire to make it work. Through this shared understanding, the Principals, although not always closely involved in every detail of the students' SDL implementation, played a vital but indirect role by ensuring that teachers were given time-off to attend the workshops and co-planning lesson sessions:

Principal L: I gave them priority in arranging time table/ schedule. I helped carve out common time for the teachers do co-planning/ meeting.

Principal M: I decreased the teachers' workload. The relevant teachers were allowed to do lesson planning together or observe each other's lesson.

Principal C: Additional resources may be more time for teachers as they have fewer lessons. For hardware, we were promoting e-learning and we had iPad already, no additional resources were allocated. We had a new e-learning classroom already.

Some principals also granted teachers additional support in terms of extra teaching assistants:

Principal Y: Extra teaching assistant and technician were arranged for lessons, especially for the experiments.

We also found that having some form of accountability for teachers to plan and develop SDL lessons to be very helpful. We found that principals of good practice schools tended to engage in more frequent teacher monitoring activities such as by asking teachers to share their SDL stories (successful or otherwise) in professional development sessions within school or with other schools, observation visits to classrooms, and regular meetings with the teachers. Such activities served as a tool to exert pressure on teachers to think more carefully of the SDL lessons they intend to implement:

Principal L: Let the teachers involved to share learning with the school (be it success story or failure). I monitored the progress of the project through regular meetings with the teachers. On some occasions, I also observed their SDL lessons.

The Principals also granted teachers the autonomy to change some of the assessment methods. For example, instead of merely relying on traditional methods of assessment such as pen-and-paper examinations, some teachers utilized peer assessment and student presentations to evaluate their students' SDL work.

Conclusions

In this study, we address the central question: *What are the specific factors that could foster SDL?* We engaged in a mixed methods study to extract specific factors that could foster SDL in science education. A total of 1,538 students and 53 teachers from 17 different schools in Hong Kong participated in this project which ran from September 2014 to June 2015. Five key enabling factors were found: (a) the nature of tasks, (b) support or scaffold for teachers, (c) teacher motivation for SDL, (d) students' prior content knowledge, and (e) principal leadership behaviors. These factors relate mainly to the learning context and personal attributes elements shown in Figure 1. Interestingly, the mere use of resources (IT or non-IT) did not appear to foster SDL. Probably the most important person to ensure successful SDL implementation in a classroom is the teacher. The results of this study suggest that teachers play a critical role in planning for the right type of activities that could engage students (i.e., *open-ended inquiry activities with real-life relevance*), and teachers have to be self-motivated to implement SDL. We believe these factors are paramount to the success of SDL implementations. This is not to say that processes such as student self-planning, self-monitoring, or self-evaluation are not important; however, these processes appear to

be dependent on the nature of the tasks, as well as the teachers' motivation for SDL in the first place. For example, a task that is closed-ended with only one simple solution would not lend itself very well to the need for students to plan, monitor or evaluate other possible explanations or answers. Neither would a teacher who desires to control every element of the class activity be expected to achieve success in SDL because such a behavior hinders the students' autonomous processes of self-planning. The Principals appear to play an indirect role by developing the teachers' instructional capability, giving teachers the autonomy to implement appropriate assessment methods, and monitoring teachers' progress so that teachers at the frontline of teaching are better supported to implement SDL. Students, on the other hand, appeared to play a somewhat "passive" role as far as the *goal-setting of SDL topics* are concerned. This is not very surprising given that schools in Hong Kong are expected to complete their pre-determined science topics in the syllabi in order to prepare students for high-stakes examinations. The pressure to score high on high stakes examinations, along with the need to cover a vast scope of content material within a limited amount of time, often creates a daunting challenge for any SDL implementation (Hew & Cheung, 2012). Very often, the pressure on the students to progress to the next topic or activity supersedes the need to stay with the students' level of interest and curiosity regarding the specific topics (Bodin, 2008).

References

- Bodin, P. (2008, March 5). Standardized curriculum is stifling students. *The Free Library*. Retrieved from [http://www.thefreelibrary.com/Standardized curriculum is stifling students.-a0176714693](http://www.thefreelibrary.com/Standardized+curriculum+is+stifling+students.-a0176714693)
- Brockett, R. G., & Hiemstra, R. (1991). *Self-direction in adult learning: Perspectives on theory, research, and practice*. London and New York: Routledge.
- Candy, P. C. (1991). *Self-direction for lifelong learning: A comprehensive guide to theory and practice*. San Francisco: Jossey-Bass.
- Campbell, L., and B. Campbell. 2009. *Mindful learning: 101 proven strategies for student and teacher success*. Thousand Oaks, CA: Corwin Press.
- Corbin, J. M., & Strauss, A. L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Log Angeles, CA: Sage Publications Inc.
- Douglass, C., & Morris, S. R. (2014). Student perspectives on self-directed learning. *Journal of the Scholarship of Teaching and Learning*, 14(1), 13-35.
- Deci, E., & Ryan, R. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227-268.
- Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly*, 48(1), 18-33.
- Gibbons, M. (2002). *The self-directed learning handbook: Challenging adolescent students to excel*. San Francisco: Jossey-Bass.
- Hew, K. F. (2015). Student perceptions of peer versus instructor facilitation of asynchronous online discussions: Further findings from three cases. *Instructional Science*, 43(1), 19-38.
- Hew, K. F., & Cheung, W. S. (2012). ICT in Education. In J. Tan (Ed.), *Education in Singapore: Taking stock, Looking forward*, pp. 75-88. Pearson International.
- Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. Chicago: Follet.
- Partnership for 21st Century Skills. (2006). Are they really ready to work. Retrieved on November 21, 2006 from http://www.p21.org/documents/FINAL_REPORT_PDF09-29-06.pdf
- Song, L. & Hill, J. R. (2007). A conceptual model for understanding self-directed learning in online environments. *Journal of Interactive Online Learning*, 6(1). Retrieved on November 4, 2015 from <http://www.ncolr.org/jiol/issues/pdf/6.1.3.pdf>
- Tok, S. (2013). Effects of the know-want-learn strategy on students' mathematics achievement, anxiety and metacognitive skills. *Metacognition and Learning*, 8(2), 193-212.
- Van Merriënboer, J. J. G., & Sluijsmans, D. M. A. (2009). Toward a synthesis of cognitive load theory, four-component instructional design, and self-directed learning. *Educational Psychology Review*, 21, 55-66.
- Yazedjian, A., Toews, M., Sevin, T., & Purswell, K. (2008). It's a whole new world: A qualitative exploration of college students' definitions of and strategies for college success. *Journal of College Student Development*, 49, 141-154.
- Yoo, S. J., Han, S-h. & Huang, W. (2012). The roles of intrinsic motivators and extrinsic motivators in promoting e-learning in the workplace: A case from South Korea. *Computers in Human Behavior*, 28, 942-950.

Acknowledgments

This research was funded by the HK Education Bureau School-based research project fund. We thank all the participating schools.

The Obj–Subj Dialectic and the Co-Construction of Hierarchical Positional Identities During a Collaborative Generalization Task

José Francisco Gutiérrez, University of Wisconsin at Madison, jgutierrez7@wisc.edu

Abstract: This paper presents a theoretical model of mathematics teaching and learning that captures the reflexive relationship between processes of objectification and subjectification inherent in classroom discourse. I call this model the “obj–subj dialectic,” and ultimately I posit that students’ respective semiotic means of objectification also function, simultaneously, as semiotic means of subjectification. I use this model to analyze and describe the individual student learning of three high school students engaged in a generalization task and the emergent social power dynamics. Findings show that whereas students made statements that positioned themselves as having the correct answer (i.e., as mathematical authorities), some students resorted to a broader arsenal of semiotic resources to make their point, which resulted in differentiated status positions. Thus some students gained “mathematical ascendancy” over others, which refers to the co-construction of hierarchical positional identities emerging from the obj–subj dialectic during multimodal interactions.

Keywords: discourse, objectification, mathematical cognition, instruction, power dynamic

Introduction: The story of Thalia, Ailani, Xeni, and Mr. Lam

Three high school students, Thalia, Ailani, and Xeni, have been working together on a pattern-finding problem, when their classroom teacher, Mr. Lam, approaches their table and asks questions related to their work. The problem presents a sequence of four geometric figures, and there is a numerical pattern that enables students to make predictions about figures further along the sequence, such as Fig. 10. Taking lead of the group, Thalia has made a prediction about Fig. 10, using an arithmetic strategy that effectively determines figures in the sequence. However, her strategy achieves this goal using a recursive technique, that is, by iteratively progressing from each figure to the next, whereas optimally a strategy would determine figures by algebraic function, without recourse to known items along the sequence. Mr. Lam prompts the group to consider much larger numbers, such as Fig. 100, asking, “So we have to know the one before it [Fig. 100]? Is there a way that we don’t have to know the one before it?”

Mr. Lam may not be aware of it, but a conflict is about to erupt. Implicit in his question is the assumption that the group’s strategy is inefficient. After a 2-sec pause, Thalia stomps her feet and exclaims to Mr. Lam, “Bruh, why you asking all these questions?” Ailani stirs the conflict, interjecting loudly, “YES!” Mr. Lam, keeping his usual calm and professional demeanor, responds to Thalia: “My name is not ‘Bruh.’ My name is Mr. Lam, and I’m challenging you. That’s why I’m asking you these questions, because I want you to get smarter.” Thalia buries her face behind her hands and starts laughing sheepishly; her face still hidden behind her hands, she continues: “OK, Mr. Lam, why you asking (all these questions).”

Background and objectives

There is more to learning mathematics in a classroom than learning content. I recorded the above vignette while observing Mr. Lam’s classroom as part of my research on the challenges and opportunities for improving mathematics education for students from historically marginalized public school populations. Thalia’s statement of “Why you asking all these questions?” reflects a complex sentiment. She may perceive that she and her peers are being questioned unnecessarily, even interrogated about their participation in the task. Ailani agrees with Thalia’s sentiment and, together, their behaviors indicate their sense of Mr. Lam’s instruction as coercive at worst, or troublesome at best.

This vignette sparked a variety of questions for me. What does it mean for mathematics students to experience classroom discourse in this way? What aspects of classroom interaction give rise to this type of social-mathematical antagonism? What is the pedagogical utility of framing mathematics learning as adversarial or as changing intellectual capacity?

Thalia and Ailani’s apparent discontent with their mathematics teacher and Mr. Lam’s particular views on mathematics pedagogy could be isolated phenomena limited to the envelope of this particular classroom community; it is conceivable that this interaction between three high school students, a teacher, and mathematics content is not representative of broader trends. However, by reflecting critically on the results of previous studies that I conducted in similar classroom settings (e.g., Gutiérrez, 2010, 2013), I argue that these participants’

experiences are related to broader issues endemic to not only mathematics classrooms serving historically marginalized students, but to all classrooms. The study presented in this paper is part of a larger project that aims to show that mathematical learning is not power neutral, but rather individual learning and power relations are mutually constituted through discourse.

The broader research project investigates how relations of power, which are inherent to all educational settings, impact students' quality of engagement and therefore learning. Whereas I focus my research efforts specifically on students from historically marginalized public school populations, such as African American, Latino/a, and economically disadvantaged students, I theorize that *all* students' participation, dispositions toward classroom practice, and mathematical knowledge are mediated by relations of power that shape local instructional contexts and the social interactions therein. To illuminate these issues, I bring to bear complementary learning-sciences and sociopolitical perspectives to expose hidden structures, mechanisms, and processes of power that either enable or hinder classroom mathematics learning.

The objectives of this paper are (1) to provide a sketch of the theoretical stance underlying the larger project, and (2) to present an emerging approach that stems from this theoretical orientation. At the center of this approach is the notion of an *objectification–subjectification dialectic*, which I define, apply, and elaborate in the sections below. Specifically, I analyze Thalia, Ailani, and Xeni's multimodal interactions during a group task, through the lens of the obj–subj dialectic.

Theoretical perspectives and relevant literature

The larger research project assumes the following theoretical stance: individual learning and social power relations become imbricated through discourse so as to mutually constitute and express each other. I view discourse not merely as an analytic lens for observing the *learning–power imbrication*, but also as the medium through which mathematical knowledge and power relations simultaneously become objectified, stabilized, and reproduced. In particular, I conceptualize the learning–power imbrication as the reciprocal discursive process whereby (1) students appropriate cultural artifacts (e.g., algebraic symbols and forms) as semiotic means of objectifying personal pre-symbolic, proto-mathematical knowledge (Abrahamson, 2009; Radford, 2003); yet so doing, (2) students not only adopt a new perspective on the world but they also become cognitively “beholden” and subject to particular discursive practices that temporally and ontologically precede them (cf. “ontological imperialism,” Bamberger & diSessa, 2003) and are inherently hierarchical. Thus, acts of (1) objectification (that forge individual learning) are at the same time acts of (2) subjectification (that reify relations of power). See *Figure 1*, below.

The obj–subj dialectic in mathematical discourse is not a novel perspective per se and similar ones are found in the literature (e.g., Heyd-Mezuyanim & Sfard, 2012; see below). However, returning to the underlying theoretical stance of this line of work, I have found the construct of a learning–power imbrication especially useful, because K-12 mathematics education occurs within a broader nexus of asymmetric power relations that have not been adequately accounted for in the literature. The goal of this paper is to explore how the analytic construct of the obj–subj dialectic can shed light on hidden discursive mechanisms that give rise to the learning–power imbrication in mathematics education.

In the model, the “obj” side of the dialectic is based on Luis Radford's theory of knowledge objectification (2003), and in particular his semiotic–cultural taxonomy of students' types of generalizations—*factual*, *contextual*, and *symbolic* generalizations (“F-C-S”). The F-C-S framework distinguishes among three generalization types in accordance with their level of generality (see Radford, 2003, for more detail). I have adapted the framework and applied it in a diversity of educational settings. I argue that F-C-S represents three *modes of action* (Gutiérrez, 2010) that students appropriate as means of dealing with pattern-generalization problems. In other words, the F-C-S framework describes both the final *products* of students' algebraic reasoning, as well as the *processes* that would result in those products. Thus the F-C-S framework enables analysis of the semiotic spaces that students *and* teachers must navigate. Furthermore, I argue that these semiotic spaces both reflect and influence students' mean-making while at the same time, these semiotic spaces are differentially imbued with status and authority. That is, as students and teachers navigate or traverse different levels of generality inherent in mathematical discourse, they simultaneously navigate/traverse relations of power (subjectification), which I explain next.

Turning to the “-subj” side, I combine several perspectives from various bodies of literature. Most relevant to this paper however, is the work of Anna Sfard. Sfard and colleagues (2012) open up new analytic avenues dealing with *direct* subjectification—when certain verbal utterances directly indicate a human actor. Building on yet extending this approach, I show in my data analysis instances where the referents of verbal utterances do not directly involve human actors but nevertheless involve *indirect* subjectification. Some utterances/actions mark subjectification through tacit positioning (Harré, 2008).

Furthermore, as students make mathematical assertions during collaborative problem solving, I argue, their discursive action creates hierarchical *positional identities* that must then, in turn, either be taken up, accepted, contested, negotiated, or rejected. The construct of “positional identity” is based on the work of Marcy Wood (2013) on “micro-identity” that describes “identities enacted in a moment in time” (p. 778). However, I emphasize the “positional” aspect to describe an emerging power dynamic, and the term already points to the mechanism by which hierarchizing occurs—positioning (Harré, 2008). That is, I conceptualize positional identities as hierarchical, because co-constructed subject positions are differentially imbued with social status and authority. Furthermore, these hierarchical positional identities are associated with students’ “locations” along the F-C-S trajectory.

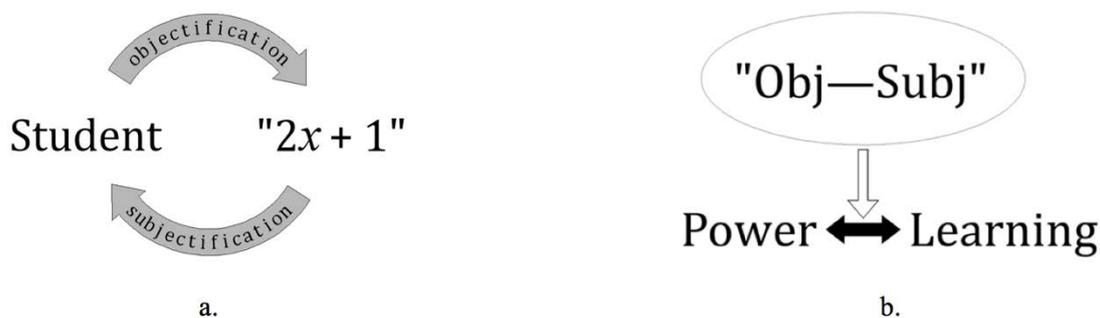


Figure 1. (a) The obj–subj dialectic: a synthesized analytic construct for observing how students appropriate cultural artifacts (e.g., algebraic symbols and forms) as semiotic means of objectifying pre-symbolic knowledge, in relation to subjectification acts that shape their positional identity. (b) As students enter each of the “F-C-S” discourse stratum and shift across them, they simultaneously create positional identities that are differentially imbued with status and authority. Thus, mathematics learning and ways of knowing become imbricated with emergent relations of power.

Methodological approach

Below I analyze two transcription segments from a much longer vignette of video data. Specifically, I explore the tension between “obj,” the semiotic resources (e.g., gesture and language, as well as conventional tools such as tables and graphs; Radford, 2003) to which students have recourse to make mathematical assertions, and “-subj,” the hierarchical positional identities that participants co-construct through these multimodal interactions (cf. Harré, 2008; Sfard & Prusak, 2005).

Data, participants, and task

The data are from a year-long participant ethnography where I immersed myself within a single mathematics class at César Chávez School for Restorative Justice (all names are pseudonyms), a small public high school located in a large, diverse urban district in northern California. At the time of the data collection (2013-2014) over 90% of the student body was Latino/a, African American, or recent immigrant students—all from the working-class and low-income neighborhoods surrounding the school.

This data collection project combined classroom observation with principles from design-based research. The focal data was collected in the Fall semester during a phase in the project where I worked closely with the teacher, Mr. Lam, to co-design and implement an experimental instructional unit. I focus my analysis on a 35-min span of video involving a group of three female students, Thalia (Grade 9), Ailani (Grade 9), and Xenii (Grade 10). The group assignment is a pattern task that combined geometric objects called “Spiralaterals” with algebraic reasoning. A Spiralateral is drawn on graph paper and is derived from a set of rules, see *Figure 2*. Frank Odds (1973) describes how to construct a $\{1, 2, 3\}$ Spiralateral:

First draw a segment of unit length to coincide with the edge of a graph square. Turn right through 90° and draw a segment two units long. Again turn right through 90° and draw a segment three units long. A basic pattern of 1-2-3 has now been established. Repeat the same steps again, continuing from the outer end of the three-unit segment. After [three] repetitions of the basic pattern, the segment will join the point at which the diagram started (p. 121).

The specific task involved a poster that presents the first four figures (i.e., “Fig. 1,” “Fig. 2,” “Fig. 3,” and “Fig. 4”) of a Spiralateral sequence, and the task objective was to express a “Code” as a set of algebraic formulas in the form of $\{f_1(n), f_2(n), f_3(n)\}$ and whose inputs are the figures’ ordinal positions (*Figure 3*). A key design feature for implementing the Spiralateral sequences was to substitute larger numbers (e.g., Fig. 10, Fig. 100), as a way for students to realize that an arithmetic-recursive strategy may be inefficient, thus motivating the need for more powerful tools and strategies such as algebraic generalizing and the use of direct formulas.

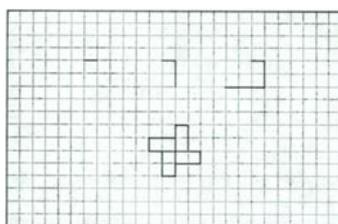


Figure 2. Steps in the construction of a “3-legged” Spiralateral (image from Odds, 1973).

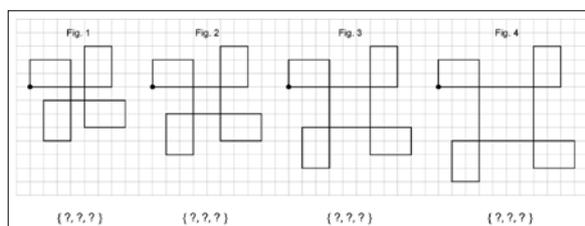


Figure 3. Poster presenting a 3-legged Spiralateral sequence, modeled as $\{2, 3, n+5\}$.

Sample data analysis: The obj–subj dialectic as an analytic frame

Here I present two transcription episodes followed by a line-by-line qualitative microgenetic (Schoenfeld et al., 1991) analysis of the semiotic resources in the students’ obj–subj processes. In this first episode Ailani adamantly claims that the solution to Fig. 10 is nineteen, whereas Thalia argues that it is fifteen; Xeni is reading her book and does not contribute mathematically in this particular transcript segment. (Note on transcript conventions: double slashes “//” mark beginning and end of overlapping utterances; two dots “..” at end of text is very slight pause, less than a second; repeated letters, e.g., “generaliiiize,” mark lengthened syllable, each repeated letter equals one “beat”; and “(??)” and “(this)” are unclear/inaudible reading or a tentative reading.)

Episode 4 (of 13) – Timestamp [08:13–09:34] – Thalia, Ailani, and Xeni work on team task

- 122 Ailani: [holds poster in front of her face with both hands, counts aloud; as she counts, her right hand comes free and she begins counting with her fingers] “Eleven twelve.. thirteen.. fourteen, fifteen, sixteen, seventeen, eighteen, nineteen.” [reaches back and taps Mr. Lam on the shoulder] “Fig. 10 will be (1-sec pause) nineteen.” (10-sec pause) [to no one in particular, she announces] “(Fig. 10 will be).. nineteen.”
- 123 Thalia: [gazes down; indicates table of values] “Fig. 10 will be fifteen.”
- 124 Ailani: [responds to Thalia] “Nineteen.”
- 125 Thalia: [gazes down; shakes head no] “Fig. 10 will be fifteen.”
- 126 Xeni: [leans her head on her hand; unclear where her gaze is].
- 127 Ailani: [loudly, almost shouting] “Fig. 10 will be nineteen!”
- 128 Thalia: [still without lifting her gaze, indicates entries with her finger as she talks in a rhythmic cadence] “It goes two-three-six, then it goes two-three-seven, then it goes two-three-eight, then it goes two-three-nine, then it goes two-three-ten, then it goes two-three-eleven, twelve, thirteen, fourteen, fifteen.” [keeps gaze down, continues working]
- 129 Xeni: [adjusts in her seat; it is possible that she responds to something Thalia or Ailani said, but one cannot tell from the footage]
- 130 Ailani: [briefly gazes in Xeni’s direction, but it is unclear if she actually addresses Xeni; likely addressing no one in particular] “I say nineteen. (4-sec pause) I’m not even sure I counted that right, but who cares?”

Analysis of Lines 122–130 reveals two important components. First, this excerpt reflects Thalia’s and Ailani’s first mathematical objectifications—both of which were articulated in the Factual mode. Ailani’s semiotic means of objectification consisted of a counting strategy and verbal speech (with a certain illocutionary force) to assertively express a partial solution, in the form of $f_3(10) = 19$. Ailani counted on from a known quantity ($f_3(4) = 6$) but she had not yet indicated that she was attempting to generalize a numerical pattern. Thalia too resorted to

verbal speech, yet she also used rhythm, gesture, and repetition, and a mathematical table as semiotic resources. Entering a rhythmic cadence (Line 128) enabled Thalia to objectify an arithmetic-recursive generalization, in the form of $\{2, 3, f_3(n-1)+1\}$ which, in this context, is more informative than Ailani's solution.

The second component this transcription reflects is a dynamic process of subjectification. To examine the dynamics of subjectification, I focus on whether/how Thalia and Ailani refer to themselves and each other. Thalia did not refer to herself or Ailani directly. Looking at Thalia's pronominal usage, what we see is that the referents of her speech did not directly involve human actors. And yet, Thalia made statements that implied *she* has the correct answer and *Ailani* does not. Ailani, too, made statements that tacitly positioned herself as having the correct answer. Thalia, however, resorted to a broader arsenal of semiotic resources to make her point. Additionally, Thalia did not make eye contact with Ailani, instead keeping her gaze on her work. This social-mathematical power encounter resulted in differentiated status positions. Ailani's effort to appropriate mathematical authority was challenged, and an opportunity was missed to engage in dialog and collaborate on the shared goal, to determine Fig. 10.

In sum, both Thalia and Ailani attempted to gain *mathematical ascendancy* over the other, which is a construct I claim that describes the hierarchical status positions that were being co-constructed in the moment. Mathematical ascendancy is based on the semiotic resources marshaled by and pitted against individual interlocutors. Thus, Thalia's and Ailani's respective semiotic means of objectification simultaneously functioned as semiotic means of subjectification.

Importantly, if we had been looking at subjectifying utterances only as defined by Sfard and colleagues, we would highlight only one single incident, when Ailani admits her counting strategy's inaccuracy (Line 130). But analysis of this single turn of talk, although a crucial one, does not capture the dynamics of implicit positioning in Lines 122–129. As the three students continue to make mathematical assertions throughout the remainder of the episode, they co-construct positional identities that must then, in turn, either be taken up, negotiated, or rejected.

Episode 5 (of 13) – Timestamp [09:34–10:56] – Mr. Lam's first visit with Thalia, Ailani, and Xeni

This episode begins as Mr. Lam returns to Team 1 in response to Ailani's tap on his shoulder. Ailani and Thalia once again assert their solutions regarding Fig. 10 and attempt to establish their mathematical ascendancy. As they argue, Mr. Lam turns to Xeni and asks her questions that would steer her toward acting in relation to the assigned task. What I aim to show in the detailed analysis, below, is that although Ailani and Thalia argue their respective solutions and appear to be reaching an impasse, their individual learning trajectories nevertheless show movement from the Factual to the Contextual modes of reasoning and, moreover, *that these advancements were forged in and through the social mathematical power dynamic.*

- 131 Mr. Lam: [to Ailani] "Alright what's the question? What was the question? You came up with how many—the code?"
- 132 Ailani: "Nineteen."
- 133 Mr. Lam: "Nineteen?"
- 134 Thalia: [loudly] "No! No. [softly, counts to herself; keeps her gaze down] Twelve.. fourteen.. [to Mr. Lam] It'll be fifteen."
- 135 Mr. Lam: [walks around to Xeni's side of the table; slides poster closer to her] "Fifteen? //Xeni, what do you think?"
- 136 Thalia: // "It'll be fifteen.// I think I did this right but I might be wrong."
- 137 Xeni: "(?)"
- 138 Mr. Lam: "Hmm."
- 139 Xeni: [turns her gaze away from Mr. Lam, down to her desk] "I don't know."
- 140 Mr. Lam [to Xeni] "OK, what might help? You think maybe writing in the code might help? [slides poster closer to her and indicates the "{?, ?, ?}" along the bottom of each figure] If you write in the code and look for a pattern?"
- 141 Thalia: "It'll be fifteen. [waves a sheet of paper in front of Mr. Lam but he ignores it and keeps his gaze on Xeni; Thalia gazes up at Mr. Lam for the first time, hands a sheet to Mr. Lam] I already wrote all the codes."
- 142 Ailani: [gaze down at her desk; she is drawing]

- 143 Mr. Lam: [grabs sheet from Thalia and places it in front of Xeni on her desk] “Oh good she’s got the codes. So maybe you can find a pattern based off of her codes. [talks in a singsong voice as he points with his fingers to values in Thalia’s table] Two three six, two three seven—”
- 144 Ailani: [interjects loudly; to Mr. Lam] “It just added a number!”
- 145 Thalia: [interjects loudly as well; to Mr. Lam] “It just goes six seven eight nine ten eleven twelve thirteen fourteen fifteen.”
- 146 Mr. Lam: [nods head, either agreeing or counting along or both] “Two three eight, two three nine, two three ten..”
- 147 Thalia: [with a tone suggesting that the pattern should be obvious to Mr. Lam] “Eleven twelve thirteen fourteen fifteen.”
- 148 Mr. Lam: “So what do you notice about the first two [referring to the first two entries of the code]?”
- 149 Thalia: “What do you mean?”
- 150 Mr. Lam: “What do you ah—”
- 151 Thalia: [makes repeating arching gesture across the figures in the poster] “It goes six seven eight nine ten eleven twelve thirteen fourteen fifteen.”
- 152 Ailani: [adjusts in her seat; without lifting her gaze, loudly] “No, it don’t!”
- 153 Xeni: [no indication she is contributing to the conversation; yet she could be observing/listening]
- 154 Thalia: “It does. It always adds//”
- 155 Ailani: // [to Thalia] “Bruh, it’s like ten [lifts gaze from her desk to Mr. Lam] but you added one or two.”
- 156 Mr. Lam: [to Thalia, indicating entries in her table of values] “Two three—no but—no but look, it says two then three then //six.”
- 157 Thalia: // “Six.”
- 158 Mr. Lam: “Then this one says two then three then seven. // Two three eight.
- 159 Thalia: // [with a tone suggesting that the pattern should be obvious to Mr. Lam; talks fast, indicating entries in the table with her finger] “It goes two three and then eight, and then it goes two three and then nine, and then it’s going to be two three and then ten, and then eleven twelve thirteen fourteen fifteen.”
- 160 Mr. Lam: [to Thalia] “OK so two three fifteen, you’re saying?” [to Ailani; makes repeated arching gesture with his right hand, marking each “entry” of the code with each motion] “OK that’s what she meant when she said fifteen. So that code, it’s staying two [gestures up-down], three [gestures up-down], and then something [gestures up-arching over to the right-down, palm up; starts to walk away]. [to the group] OK so how would you find for any—for any number?”
- 161 Thalia: [throws hands in the air; posture and facial expression suggest frustration; to Xeni] “You just keep adding the number, duhhh!” [looks to Ailani, then looks to Xeni]
- 162 A. & X.: [no response]

Earlier in Episode 4, we observed that Thalia’s and Ailani’s semiotic means of objectification also served as semiotic means of subjectification, and as such these semiotic resources formed the basis of a social-mathematical hierarchy. Here in Episode 5 the power dynamic continues and the students marshaled their semiotic resources to resolve an overt mathematical conflict, but we also observe that the teacher marshaled his semiotic resources, as well, to orient Xeni toward acting in ways that are relevant to the problem-solving domain. When Mr. Lam walked away from Ailani and Thalia’s side of the table to Xeni’s side and asked her, “What do you think?” he positioned Xeni as the center of attention and her (non)participation as the critical aspect of the situation, even as Ailani was vying for his attention regarding Fig. 10.

Thalia too fought for Mr. Lam's attention, for him to recognize that she had done all the work, when she stated "I already wrote all the codes" as she waved the sheet of paper in Mr. Lam's view. Her tone of voice suggested that he need not bother Xenia with that task because it was "already" completed. Mr. Lam maintained his usual warm, professional demeanor at this moment as well, and finally accepted Thalia's gesture to look at her codes. In that moment, Ailani interjected with, "It just added a number!" (Line 144), which is an utterance iterated in the Contextual mode as it referred to a general procedure and was not tied to concrete elements in the problem space. What is most interesting to note at this point, is that this is the first student utterance coded in the Contextual mode, whereas all previous utterances were articulated at the Factual level of generality. Thalia, too, interjected at the exact moment that Ailani articulated her Contextual statement (Line 145), with a re-articulation of her Factual recursive generalization: "It just goes six seven eight nine ten eleven twelve thirteen fourteen fifteen." Mr. Lam responded to Thalia's and not Ailani's contribution, asking a question about the pattern that Thalia was verbalizing. Ailani's contribution went unacknowledged and, as a consequence, Ailani went unrecognized as having achieved a greater level of generality than all the other student participants.

We see here that Ailani's semiotic means of objectification have increased in generality, suggesting increase in her mathematical sophistication and understanding. However, when we look closely at Ailani's remark to Thalia, "Bruh, it's like ten but you added one or two" (Line 155), we see that the mathematical basis of her assertion is uncertain. Ailani is not sure if one needs to add 1 or add 2 from the last known quantity (Fig. 4), so she adjusted the repeated summand to account for the difference between Thalia's answer of "fifteen" and her "nineteen." Despite the uncertainty Ailani is indicating here in Episode 5, and despite Ailani having admitted the possibility of a faulty counting strategy at the end of Episode 4, she nevertheless continued to hold on to this solution. She is personally invested in the task and will carry this solution through to the end. (The analysis of Episodes 8 & 11 show that her statement of "added one or two" comes back into play.)

Thalia responded to Ailani's assertions with a statement that reflects her first teeter into the Contextual mode of reasoning, at Line 154: "It does. It always adds—" but it was cut off by Ailani.

Both Ailani and Thalia made statements that express a mathematical procedure, but the illocutionary force of their utterances also positioned them as capable of asserting knowledge and thus established them as authority figures in the discussion. I claim that through these complex interactions, the threshold to operate in the Contextual mode was lowered because their status and identity were at stake. That is, the inherent power struggle of the conversation spurred them on to position themselves as authority figures within that dynamic, and what resulted were Ailani and Thalia's first Contextual statements.

With regard to Mr. Lam's role in the evolving dynamic. At Line 143, there was a notable shift in Mr. Lam's discourse from general discursive tactics (to encourage Xenia to participate) to operating in the Factual mode, when he highlighted certain aspects of the problem situation, marking them as important with his tone and cadence.

Mr. Lam revoiced Thalia's recursive generalization (Line 159), but he shifted the semiotic space from the Factual to the Contextual level of generality. So doing, Mr. Lam did not reiterate the additive/recursive component of Thalia's code, only that the first two entries remained constant while the variable was the third entry (Line 160: "It's staying two, three, and then something"). As Episode 5 came to an end, Mr. Lam walked away and tossed a final question to the team, asking them to consider cases beyond just the first ten figures (Line 160: "for any number?"). Thalia scoffs at Mr. Lam's final question, with one final statement in the Contextual mode that, for her, captures the complete solution (Line 161: "You just keep adding the number—duhhh!"). Thalia may have interpreted the subtext of Mr. Lam's question as evaluating her solution as insufficient, and she disagreed with his assessment. Mr. Lam's question is a common tactic used in these kinds of instructional contexts involving figural patterns, intended for students to realize that recursive strategies are limited when dealing with cases much further down the line. Thalia's comments point to a mismatch between her perception of the requirements of the task and Mr. Lam's expectations.

Thalia's frustration ensues, all the way until Episode 11, where an overt conflict erupts between her, Mr. Lam, and Ailani. This conflict momentarily leaves the discursive frame of "mathematical practice" to a broader ideological frame. Below, I present a synopsis of Episode 11, so as to briefly highlight other aspects of the social-mathematical power dynamic. Specifically, the teacher makes two noteworthy moves in Episode 11 that have implications for the teaching and learning of mathematics more broadly. First, Mr. Lam challenges the group to generalize their Code to a normative algebraic one that can predict much larger numbers along the sequence. Second, Mr. Lam encourages Xenia to participate in the activity, to include her in the discourse. Yet so doing, whereas Mr. Lam's moves were pedagogically effective, they reified mathematical hierarchies that prioritize certain forms of argumentation over others (e.g., formal symbolism over explanations involving gesture and other communicative measures) and thus positioned Xenia with higher status.

Episode 11 (of 13) – Timestamp [25:18–29:54] – “My name is not ‘Bruh’.”

Mr. Lam returns to Team 1 and is met by Thalia with a question: “Are we done?” Mr. Lam confirms they are nearly done. Mr. Lam turns to Xenia and kneels down next to her to talk. He reminds both Ailani and Xenia that the task at hand involves working on the assigned pattern task, not drawing pictures. Thalia defends her team, exclaiming, “We just did it, look it! We just did that.” Mr. Lam does not respond to Thalia and instead asks Xenia what she noticed about the Spiralateral pattern. Working together, Mr. Lam, Thalia, and Xenia co-construct a Code for the pattern. Xenia articulates her version of the final proposed solution using formal symbolism, stating that the last part of the Code is “ n plus one.” While Mr. Lam tries to unpack what Xenia means by “ n ,” Thalia interjects and gives indication that she is growing frustrated with the conversation and is focused on task completion... which brings us to the opening scenario at the very top of this paper.

The solution that Team 1 has articulated, for items beyond Fig. 4, achieves the goal using a recursive technique that relies on known items along the sequence. Mr. Lam prompts the group to consider much larger numbers, such as Fig. 100, asking, “So we have to know the one before it [Fig. 100]? Is there a way that we don’t have to know the one before it?” After a 2-second pause, Thalia stomps her feet and exclaims to Mr. Lam, “Bruh, why you asking all these questions?” Ailani stirs the conflict, interjecting loudly, “YES!” Mr. Lam calmly responds to Thalia: “My name is not ‘Bruh.’ My name is Mr. Lam, and I’m challenging you. That’s why I’m asking you these questions, because I want you to get smarter.” Thalia buries her face behind her hands and starts laughing sheepishly; her face still hidden behind her hands, she continues: “OK, Mr. Lam, why you asking (all these questions).”

Mr. Lam continues the mathematical conversation and challenges the group with “So could you even figure out Fig. 100 right now?” adding the parameter of “I don’t tell you what Fig. 99 looks like, can you tell me Fig. 100?” Thalia explains her method for Fig.’s 1–10, but Mr. Lam insists that her method is insufficient for items much further along the sequence. Ailani leans back in her chair, throws her arms behind her, looks to Mr. Lam and leans forward as she exclaims: “Well maybe—OK so (like) two three and you don’t have to know all of them, you know the first—you can know the original one and then (add on)! And then figure out Fig. 99.” As she speaks, Ailani pounds the table with the marker in her hand and uses a tone of voice suggesting to Mr. Lam that she aims to “settle the score” once and for all in this confrontation. Thalia interrupts Ailani, softly clapping her hands together as she speaks, accentuating each word as she talks: “Why you talking so ratchet?” Mr. Lam weighs in with: “Don’t get hyped up against me.” As Mr. Lam is speaking, Thalia points to him and says: “Ratchet. Say it, don’t get ratchet.” Mr. Lam does not respond to Thalia’s comment and instead steers the conversation back to the mathematics at hand, asking about Fig. n . The conversation continues for a few more minutes, until finally Xenia arrives at a closed solution of “ n plus five.” (See Gutiérrez, 2016, for richer description and deeper analysis of Episode 11, in which I explore a possible mechanism that links cognition to social structures and vice-versa in this episode.)

Conclusions

An analysis of mathematical ascendancy, through the lens of the obj–subj dialectic, suggests that power dynamics and new mathematical understandings are co-constitutive through public discourse. The results of the study presented here, combined with findings from the larger project, show that the co-construction of power relations and the co-construction of knowledge are not necessarily distinct processes. Surely each of these processes can be independently instantiated in practice, but I maintain that they also become co-constitutive at certain points in the discourse, thus the learning–power imbrication emerges.

References

- Gutiérrez, J.F. (2010). Distinguishing between the process and the product of student algebraic reasoning. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines (ICLS2010)*. Univ of Ill., Chicago.
- Gutiérrez, J. F. (2013). Agency as inference. *REDIMAT*, 2(1), 45-76.
- Gutiérrez, J. F. (2016, July). Exploring tensions in the obj–subj dialectic (In N. Presmeg & L. Radford [Co-Chairs], Topic Study Group 54: Semiotics in mathematics education). *13th International Congress on Mathematical Education (ICME)*. Hamburg: University of Hamburg.
- Abrahamson, D. (2009). Orchestrating semiotic leaps. *C&I*, 27, 175–224.
- Bamberger, J., & diSessa, A. (2003). Music as embodied mathematics. *IJCML*, 8, 123–160.
- Harré, R. (2008). Positioning theory. *Self-Care, Dependent-Care, and Nursing*, 16(1), 28–32.
- Heyd-Metzuyanim, E., & Sfard, A. (2012). Identity struggles in the math classroom. *IJER*, 51-52, 128–145.
- Odds, F. C. (1973). Spirolaterals. *Mathematics Teacher*, 66(2), 121–124.
- Radford, L. (2003). Gestures, speech, and the sprouting of signs. *MT&L*, 5, 37–70.
- Sfard, A., & Prusak, A. (2005). Telling identities. *Educational Researcher*, 34(4), 14–22.

Designing a Blended, Middle School Computer Science Course for Deeper Learning: A Design-Based Research Approach

Shuchi Grover, SRI International, shuchi.grover@sri.com
Roy Pea, Stanford University, roypea@stanford.edu

Abstract: As computer science (CS) thunders its way into K-12 classrooms, lack of teachers to teach CS and pedagogically sound computing curricula remain significant challenges. This paper describes a design-based research (DBR) approach to create and refine a 7-week middle school introductory CS course—*Foundations for Advancing Computational Thinking (FACT)*—informed by research in the learning sciences and computing education, and designed for blended in-class learning using online materials created on the OpenEdX platform. The paper shares FACT’s success in achieving a balanced pedagogy to meet its goals for deeper learning of algorithmic concepts, and identifies areas for improvements.

Introduction

In a world infused with computing, ‘computational thinking’ (CT) skills are seen as key for all citizens in the digital age, not only computer scientists (Wing, 2006, Grover & Pea, 2013). President Obama’s 2016 “Computer Science For All” initiative has paved the way for expansion of CS to K-12. However, few structured curricula exist for middle and elementary school levels. The middle school years are formatively central for cognitive and social development in K-12 schooling especially regarding future engagement with STEM fields (Tai et al., 2006). Middle school experiences should thus make students amenable to diverse future opportunities. Unfortunately, middle schools, like other levels of K-12, face an acute shortage of teachers to teach introductory computing curricula.

Developing well-designed online curricular materials makes possible accelerating scaling to wider audiences of students and teachers. Although recent years display growth of online CS materials on venues such as Khan Academy and Code.org, their success for development of deeper, transferable CT skills is yet to be empirically validated, and they also appear to lack robust assessment measures. Recently advancing MOOC platforms could serve this crucial need in K-12, but to be effective for younger learners, online curricula for middle school students would have to be consciously designed for active learning and engagement. What should a first-of-its-kind introductory CS course created on a MOOC platform for blended in-class learning in middle school look like? What is the best balance of pedagogies, online and offline, individual and collaborative, open-ended and directed activities for “deeper learning”? Informed by past research in the learning sciences and computing education, this research adopted an iterative process to design and refine a blended middle school introductory CS course, *Foundations for Advancing Computational Thinking (FACT)*, and to empirically answer questions on the development of deeper, transferable CT skills among teens.

This paper focuses on the use of design-based research (DBR) as a methodology to put learning theory to work for designing the learning environment for the online FACT materials and the blended in-class learning experiences as well as investigating curriculum sequences, instructional approaches, activities, and assessments. It describes how initial designs of the intervention that represented “embodied conjectures” (Sandoval & Bell, 2004) of the researchers were refined over two iterations of DBR with help from stakeholders as active “design partners”. As such, the quantitative data analyses of learning outcomes are not the main thrust of this paper. They are described in much more detail elsewhere (Grover, Pea, & Cooper, 2016).

Design-based research methodology for designing and refining FACT

The driving goal of our research was to design a curriculum to prepare middle school learners for CT. The first step in establishing the broad viability and usefulness of a curricular intervention to achieve desired outcomes as described above was to design and test how well it worked in a *real classroom setting*. Specifically, how effective is the FACT curriculum for helping middle school students develop awareness, positive attitudes and interests towards the discipline of computer science, and a deeper understanding of foundational computing concepts that can transfer to future computing experiences? Designing and studying FACT in context was arguably a ‘Type 1’ translational research effort (Pea, 2010), and thus benefited from a DBR approach, where the key stakeholders typically involved are teachers, learners, learning scientists, subject matter experts, and technology developers. The types of questions typically answered through such research are: (a) *What do students learn from this design?* (b) *How do students learn from this design?* (c) *What do problems in learning or implementation suggest about re-design of the intervention?*

DBR is increasingly being embraced for research in the learning sciences, instructional design, and curriculum development. It includes “testing theoretical accounts of learning, cognition, and development by using the theory to design or engineer instruction and conducting empirical studies of the instruction in a naturalistic setting” and it responds “to the need to study complex interventions that include a range of intentionally-designed features and materials such as curriculum sequences, technological tools, social norms, instructional approaches, activity structures, and cognitive assessments in complex settings” (Bell, Hoadley, & Linn, 2004). DBR addresses the systemic, complex nature of education by researching curricular interventions in real-life settings, often involves stakeholders such as teachers and learners in the design process, and aligns particularly well with the goal of promoting inquiry in STEM courses as well as technology-enhanced learning environments (Barab & Squire, 2004).

A key feature of DBR studies is that they are typically iterative in nature, and involve refinement of the intervention. The initial tentative set of design considerations embodies the conjectures of the researcher(s)-designer(s) that are then revised depending on their successes and challenges experienced in practice (Edelson, 2002). Bell et al. (2004) suggest that DBR tends to include compelling comparisons in which two forms of the innovation are enacted under otherwise similar conditions and variations in the innovation used during such compelling comparisons test hypotheses about learning embodied in the designs. DBR methodologies specifically targeting design of learning environments also call for designs that are *pragmatic* and *grounded* with respect to theories of how people learn and the specific contexts within which the technologies are implemented (Wang & Hannafin, 2005). The goal of design research is also to help develop domain theories, design frameworks and/or design methodologies.

Stakeholders as ‘design partners’ and iterative research design

The novelty of the curricular materials and the online platform in this context necessitated drawing from learning theory and past research on children and programming as a foundation and starting point for the **initial** designs of FACT. Using ideas as advocated in DBR for getting guidance from stakeholders on what makes sense for middle school children using a curriculum such as this one, we adopted an iterative process to refining the curriculum. Given the newness of MOOC platforms and that this effort was a first-of-its-kind blended learning course using such a platform for a K-12 setting, the idea of involving the classroom teacher and learners as ‘design partners’ seemed particularly appropriate. Adopting a DBR methodology thus involved testing the curriculum *in context*—in a real middle school classroom setting—and gathering detailed feedback from students and the classroom teacher, the key stakeholders who were ‘design partners’ in this endeavor. The original curriculum before the first iteration represented the embodied theoretical conjectures about how middle school learners can best achieve the desired outcomes, and as such, the curriculum carried “expectations about how designs should function in a setting.” (Sandoval & Bell, 2004).

Guided by findings from preliminary explorations and the design-based research (DBR) approach, empirical investigations were conducted over two iterations (hereafter referred to as Study1 and Study2) of teaching FACT in a public school classroom. In keeping with DBR philosophy, in both iterations, frequent feedback was sought from students during the intervention in addition to extensive feedback during and after the course via surveys. Students were reminded often by the classroom teacher that in addition to being learners, they also had a role in this research as ‘design partners’ and *that their inputs and feedback were crucial to refine the course for future students who would learn from an improved course that will have incorporated the students’ ideas for improvement*. The lead researcher met with the teacher several times in between the first and second iteration to go over student feedback from Study1, and to solicit her ideas for how those suggestions could translate into course features for the online materials and blended course.

The two iterative studies involving the use of FACT in a middle school classroom investigated the research questions: RQ1- What is the variation across learners in achieving desired outcomes through FACT, specifically the learning of algorithmic flow of control– (a) serial execution (b) looping constructs and (c) conditional logic, and what factors influence this variation? RQ2- How well does the curriculum promote an understanding of algorithmic concepts that goes deeper than tool-related syntax details as measured by PFL assessments? RQ3- How do students’ perception of the discipline of CS change as a result of FACT?

Teaching the curriculum face-to-face first (in Study1) *without the constraints of the online medium afforded a focus on the pedagogical content knowledge or PCK* (Shulman, 1987) *designs as well as design of assessments and surveys for gathering feedback used in the curriculum*. Furthermore, using the first iteration to pilot a portion of the curriculum as an online course on a MOOC platform in Study1 helped researchers to observe the classroom learning experience and elicit student as well as teacher feedback that informed subsequent refinements of the initial set of design and pedagogical assumptions about the use of the online version of FACT in a blended classroom setting. Specifically, the researchers set out to examine how the initial

curriculum and learning experience that embodied conjectures of the researchers/designers would be enacted in a classroom setting. In addition to questions on how students learned specific computing concepts as measured by pre-post tests, there were design questions for consideration. Our (representative) list of questions about the design and curriculum that we aimed to answer through DBR was: (1) how should the pedagogical balance between open-ended exploration, guided inquiry, instruction and directed activities be achieved? (2) how should online and offline activities and topics be sequenced? (3) what should be the ideal length of instructional videos? (4) which Scratch programming projects are the most fun for learners and which are the most difficult? (5) how can authentic activities be incorporated into the curriculum that attend to learner agency while still keeping directed programming assignments that address specific learning goals? (6) which videos are most effective, and which the least, in raising students' interest in CS? why? (7) what steps could be taken to make the online modules more engaging? (8) how should students with varying abilities be accommodated in a blended course that requires students to move every week to a new unit?

DBR with a difference

It should be noted that in designing a curriculum with clear *a priori* operationalized outcomes (represented by designed measures) and refining it to create an optimal learning experience that also results in learners achieving those desired outcomes, this approach could be viewed as somewhat problematic in some researchers' views of DBR. Engeström (2009), in particular, takes issue with the linear fashion of research that many forms of DBR efforts have adopted such as Collins, Joseph & Bielaczyc (2004) that start with researchers determining the principles and goals and going through subsequent phases of refinement leading to completion or perfection. A methodological goal of this research too was to refine the FACT in order to best achieve desired outcomes, often relevant to building an "outcomes" domain theory (Edelson, 2002). There is, however, resonance in the assertions of von Hippel and Tyre (1995) that Engeström (2008) cites to make his point about what he considers the true nature of DBR– "one can never get it right, and that innovation may best be seen as a continuous process, with particular product embodiments simply being arbitrary points along the way."

Methods

Participants and procedures

Empirical studies were conducted in a Northern California public middle school classroom. Two iterations (Study1 and Study2) of the design-based research (DBR) were conducted with two different cohorts in the same 'Computers' elective class that met for 7 weeks, with four 55-minute periods per week. The samples comprised 7th and 8th grade students (Table 1). In both iterations, approximately a fifth of the class comprised students who had been placed in the elective class by the school counselors. These students happened to be English language learners or students with other learning difficulties. Unfortunately, since this was an elective class, these students did not get the same specialist supports they received in core subject classes. The classroom teacher, who did not have a background in CS or programming, was present in the classroom at all times assisting with classroom management and "learning right alongside the students."

Table 1: Student Samples in Study1 & Study2

Study	Mean Age	Count by Gender		Count by Grade		Count in Sp. Programs	
		Male	Female	Grade 7	Grade 8	ELL	Special Ed.
1	12.9	21	5	15	11	4	2
2	12.3	20	8	16	12	3	1

In Study1, the course was taught face-to-face by the lead researcher. ***One unit in Study 1 was piloted on the MOOC platform with face-to-face instruction replaced by videos and interspersed with automated quizzes.*** Extensive feedback was sought from learners on their experiences with the online unit as a precursor to creating online curricular materials for the entire curriculum for Study2. Study2, conducted in the same classroom with a new cohort, used a designed blended learning experience with online FACT materials on Stanford's OpenEdX platform (roughly 50 videos, 1-5 minutes in length), quizzes, and Scratch activities to be done individually or collaboratively that were interspersed through the course. In addition, several refinements were made to the curriculum based on experiences and student feedback in Study1. The studies were conducted during two semesters in 2013 that included visits to the classroom preceding and following the 7-week long FACT intervention for IRB permissions, pre-post tests, as well as wrap up of final projects and presentations.

Data measures

Table 2: Instruments used to capture key data measures

Instrument	Pre	Post	Source(s)
Computational Knowledge Test	X	X	Ericson and McKlin (2012); Meerbaum-Salant et al. (2010); Zur Bargury et al. (2013)
Preparation for Future Learning (PFL) Test		X	Designed (Inspired by Schwartz and Martin (2004))
Prior Experience Survey (programming experience and technology fluency)	X		Adapted from Barron (2004)
CS Perceptions Survey	X	X	Ericson and McKlin (2012)
Online Learning Experience Survey (Study 2)	X	X	Designed (Inspired by Barron (2004))
FACT Experience survey		X	Designed (for getting student feedback)
Classroom Observations			During the interventions for DBR

FACT curriculum and pedagogy

The 7-week FACT curriculum (Table 3) included topics that were considered foundational and appropriate for middle school students and focused on algorithmic thinking required for learning programming. The curriculum design effort was guided by goals for deeper learning (Pellegrino & Hilton, 2013), which attend to the development of cognitive abilities through mastery of disciplinary learning, in addition to interpersonal and intrapersonal abilities. The pedagogy for CT followed a scaffolding (Pea, 2004) and cognitive apprenticeship (Brown, Collins, & Newman, 1989) approach. It involved the use of (worked) examples to model solutions to computational problems in a manner that revealed the underlying structure of the problem, and the process of composing the solution in pseudo-code or in the Scratch programming environment. Academic language and computing vocabulary were used throughout this scaffolding process. The course emphasized “learning by doing” for students through a mix of directed assignments as well as meaningful, open-ended projects (Barron & Darling-Hammond, 2008) in Scratch including a substantive culminating project. Frequent low-stakes and high frequency multiple-choice assessments were designed to keep learners deeply engaged with the course’s content and understanding goals, and to provide feedback for both learners and the teacher. 'Preparation for future learning' (PFL; Schwartz, Bransford, & Sears, 2005) for text-based computing contexts was mediated through expansive framing pedagogical moves (Engle et al., 2012) and providing learners opportunities to work with analogous representations (Gentner et al., 2003) of the computational solutions—plain English, pseudo-code and Scratch programs (Grover, Pea, & Cooper, 2014b). Summative assessments included a specially designed PFL test in addition to other performances of CT understanding. FACT also consciously engaged with students’ narrow perceptions of CS to help them see computing in a new light through a curated video playlist that illuminated computing as a creative, problem-solving discipline with applications in many real-world contexts (for details see: Grover, Pea, & Cooper, 2014a). The goal of DBR was to use the data measures to test the pedagogy and designed curriculum & assessments to achieve desired outcomes.

Table 3: Curriculum Sequence for FACT

<i>Unit 1</i>	Computing is Everywhere! / What is CS?
<i>Unit 2</i>	What are algorithms & programs? Computational solution = Precise sequence of instructions
<i>Unit 3</i>	Iterative/repetitive flow of control in a program: Loops and Iteration
<i>Unit 4</i>	Representation of Information (Data and variables)
<i>Unit 5</i>	Boolean Logic & Advanced Loops
<i>Unit 6</i>	Selective flow of control in a program: Conditional thinking
	Final Project (student’s own choice; could be done individually or in pairs)

Refinements to learning environment after Study 1

Key improvements to Study1 were incorporated in the FACT online/blended version for Study2:

1. Videos were shortened from ~8-10 mins to ~1-5 mins in Study 2 (to heighten learners’ engagement).
2. More time was devoted to loops and variables (since they struggled with those topics most).
3. Small modifications in content sequence that made more sense.
4. Addition of Scratch window below video for students for greater interactivity.
5. More games and creative artifacts in Scratch Assignments (based on student feedback).
6. More engaging corpus of ‘*Computing is Everywhere!*’ videos. (Additional videos to show after week 1)

7. More fun worked examples in Scratch ('Pong' for conditionals; '4-quadrant' art for Boolean logic).
8. Final project of choice (in pairs or individual), whole-class demo, and online showcase of projects.
9. A 'Word Wall' tab for computational terms.
10. Additional thought questions as well as online space for contextual questions (below each video).
11. Improved survey question design for soliciting student feedback.

Analysis and results

In order to answer the three research questions, data were analyzed separately for each study using mixed method techniques. The pre-to-posttest effect size (Cohen's *d*) on the CT test was ~2.4 in both studies. Although students scored an average of 65% on the PFL test, there was evidence of understanding of algorithmic flow of control in code written in a text-based programming language. Most of the PFL questions involved loops and variables—topics that students had the most difficulty with in the computational learning posttest. Regression analyses suggested that the curriculum *helped all students regardless of prior experience as measured by the self-report survey, however the pretest score was found to be a significant predictor for both the posttest and the PFL test*. Responses to the perceptions of computing question were analyzed using a mix of qualitative coding and quantitative methods. They revealed a significant shift from naïve “computer-centric” notions of computer scientists to a more sophisticated understanding of CS as a creative, problem-solving discipline.

Comparative analysis of Study 1 vs. Study 2

Computational learning and PFL (Transfer)

On the main outcomes of interest, namely the CT posttest score and the PFL test score, there was no significant difference among the students in Study 1 and Study 2. The pretest scores in the two studies were also not statistically different. *The learning gain, however, calculated as the difference between the posttest and pretest scores was significantly higher in Study 2 on both the t-test and the non-parametric Mann Whitney test*. The PFL test performances were also comparable and the difference between the average PFL scores in Study1 and Study2 was not significant. These results and score breakdowns by CT construct are shown in Tables 4 and 5. It should be noted that most of the questions on loops in the posttest also included conditionals and serial execution, in addition to variable manipulation—a topic with which students struggled in both studies. For the questions involving loops, students thus needed a good understanding of how different computational constructs come together in a program, and also of variable manipulation within loops. Both these aspects are particularly difficult for novice programmers (Pea, 1986; Soloway, 1986; Spohrer & Soloway, 1986).

Table 4: Student Samples in Study1 & Study2

	Study 1		Study 2		t	p < t	z	p < z
	N	Mean (SD)	N	Mean (SD)				
Pretest	24	36.33 (18.19)	28	28.06 (21.18)	1.5	0.14	1.76	0.08
Posttest	26	78.58 (17.08)	28	81.60 (21.24)	-0.58	0.56	-1.26	0.21
Learning Gain	24	43.08 (12.17)	28	53.07 (18.34)	-2.34	0.02	-2.46	0.01
PFL Test	25	63.37 (28.86)	27	65.07 (26.47)	-0.22	0.82	-0.08	0.93

Table 5: Student Samples in Study1 & Study2

Variable	Study 1	Study 2	t-Stat	p	Z-Score	p
	Mean (SD)	Mean (SD)				
Overall	78.6 (17.1)	81.6 (21.2)	-0.6	0.56	-1.3	0.21
<u>By CS Topic</u>						
Serial Execution	97.4 (13.1)	91.1 (20.7)	1.4	0.18	1.6	0.12
Conditionals	84.5 (19.0)	84.9 (20.5)	-0.1	0.94	-0.4	0.72
Loops	74.1 (21.9)	77.2 (26.3)	-0.5	0.64	-1.1	0.29

Perceptions of computing

There was no statistical difference in students' pre-post attitudes and interests in computing in either study. Such a ceiling effect is not uncommon when learners self-select into the intervention, and enter with high levels of interest and motivation. However, students in Study2 performed better than those in Study1 on growth of students' perceptions of computing, tested mainly through pre-post responses to the question “*In your view, what do computer scientists do?*” There were no significant differences between Study1 and Study2 in any of

the key coding categories that were used to code the responses. Since DBR-inspired refinements in the Study2 materials were aimed at tackling learners' awareness and understanding of computer science, some improvements in student performance were to be expected. For this reason, we also tested the *a priori* hypothesis that in assessing the more fine-grained aspects of the responses (*richness* and *length* of answers to "In your view, what do computer scientists do?") students in Study2 would do better than those in Study1. Statistical analyses confirmed this hypothesis: the differences across the two studies (as measured by t-tests and Rank Sum tests) were significant, with students in Study2 performing significantly better. The difference in the post-course responses as measured by the number of meaningful codes (richness) was significant at $p < 0.01$, and the increase in response length from Study1 to Study2 was also significant at $p < 0.001$ (Table 6).

Table 6. "What do computer scientists do?". Comparison of Responses in Study 1 & Study 2

Variables (N = 54)	Study 1 Mean	Study 2 Mean	t	P(T > t)	z	P(Z > z)
Length, Before	80.8 (59.5)	57.0 (38.1)	1.7	0.089	1.4	0.161
Richness, Before	0.4 (0.9)	1.0 (1.2)	-2.2	0.03	-2.4	0.016
Length, After	64.0 (39.1)	181.6 (106.9)	-5.6	< .001	-4.8	< .001
Richness, After	2.0 (0.8)	3.0 (1.2)	-3.4	0.001	-3.0	0.003

Final project, presentation and interview as performance assessments in Study 2

Study1 did not have a formal project at the end of the intervention, as the weeks following FACT were disrupted by end-of-year activities. However, the success of the final open-ended project for the few children that did it informally prompted a decision to add a 7th week to FACT in Study2. After the six weeks of the FACT curriculum ended, students formally worked on a culminating project that involved designing and programming a game of their own choosing in Scratch that they then presented at a whole class 'Project Expo Day'. Sometimes they also projected their code to show how certain aspects of the game were programmed. At other times, they called on their peers to demonstrate games that required two or three players. All the final projects were also uploaded to an online project "showcase" (a Scratch 'studio'), a resource enabling students to see all the projects, and play games created by their classmates. The following week, students continued to fix bugs in their projects, created short 'user manual'-styled instructions to describe game play, and documented the final project experience in a form adapted from the *Starting from Scratch* curriculum (Scott, 2013) that was provided to each student. Most importantly, students spent time playing with each other's games (or their own). In their post-survey responses, students were asked to vote the **three** Scratch projects that were "the most challenging", "the most fun" and "the least fun" among all the Scratch projects that they had completed. The final project was ranked highest on "the most fun" and also the "the most challenging".

In contrast to the de-contextualized posttest with questions that involved comprehending Scratch code and answering related questions, the final project was a more meaningful form of performance assessment (Barron & Darling-Hammond, 2008). It embodied learner agency and students felt a sense of accomplishment and pride as they presented their projects to the class and received cheers of praise from their peers. Most importantly, it seemed to work well even for the students who performed poorly on the posttest, although their projects used fewer complex structures than those of students who performed better on the posttest.

Discussion

As these results reveal, the iterative refinements driven by DBR resulted in the blended version of FACT using OpenEdX worked as well (if not slightly better on some metrics) as the face-to-face version of the FACT curriculum. Based on success metrics for an online *replacement* course (Means, et al., 2010), online/blended FACT was successful since learners did at least as well as those in the face-to-face intervention. This result is even more encouraging when we take into account students' preferences for face-to-face learning over online learning (as revealed in pre and post surveys in Study2).

It should be noted that the improvements in Study2 were observed on the aspects of the FACT learning designs that had been refined after the first iteration of this DBR investigation. They also targeted ways in which students could better appreciate the ideas behind the 'Computing is Everywhere!' unit of the curriculum. The DBR effort thus resulted in an improved curriculum in those aspects where re-design was enacted.

Reflections on future improvements

Despite allotting more time in Study2 for hard-to-learn concepts such as variables and loops, learners still struggled with those ideas, especially those with poor math preparation. This finding suggests that perhaps we need to rethink our pedagogy for introducing those ideas to reach diverse learners. The results also suggested

that for a balanced set of questions with robust construct validity, the PFL test should assess learners on the individual concepts taught—serial execution, variables, conditionals and loops—rather than only loops and variables as was the case. Additionally, many items on the pre-post surveys, pre-posttests, quizzes, and the PFL test required students to read non-trivial amounts of text—a real challenge for ELL students. Such students often require the help of aides in core subject classrooms—help that this classroom did not have as this was an elective course. Clearly the curricular materials would have to become less dependent on English, perhaps through images or simpler scenarios. Lastly, although the online FACT course designed for blended learning could be considered a success per the criteria guiding this research, classroom observations and student feedback suggest that the FACT could have incorporated more (guided) exploration and learner agency. Lastly, some students clearly would have benefited from more time; having to stick to the schedule of the research study meant that some students lagged behind in the course. This needs to be improved. Congruent with methodologies of DBR, the current version of FACT is still a work-in-progress. It is hoped that teachers who use this course can address some of these limitations in ways that work for them and their students.

DBR lessons and future work

What takeaways does this DBR effort provide to the field? Edelson (2002) contends that while the lessons of individual design effort are often restricted to the particular design and the individuals involved in it, DBR has the additional goal of developing generalizable domain theories, design frameworks and/or methodologies. Piloting a new curriculum face-to-face (with only one unit online) first to focus on examining PCK, activities and assessments without the online/blended modality, was a useful way of modularizing the problem space. It serves as a design methodology that has applicability in course design using MOOC platforms for blended or all-online settings. FACT design also provides strategies for incorporating active learning in MOOC courseware. The iterative design of FACT as an online course on a MOOC platform for blended in-class learning provides an example of incorporating lessons from the learning sciences in designing an effective learning environment that relies on video-based instruction, while balancing online and offline work by learners. Using inquiry to activate a richer web of mental connections, contextual discussions preceding and following videos, and mechanisms for learners to program and respond to thought questions as they watched videos made for a more active learning experience. Most importantly, FACT provides evidence for a balanced pedagogy for K-12 CS classrooms—one that incorporates guided instruction with active learning, and directed with open-ended programming projects for deeper learning of conceptual ideas underpinning algorithmic thinking and programming. Incorporating a wide range of mechanisms to learn, hone and demonstrate CT through various kinds of assessments including PFL transfer assessments (Grover et al., 2015) makes FACT a unique and innovative introductory CS curriculum. This research also exemplifies the value of DBR as a methodology to iterate on the design of such a first-of-its-kind course. In their role as “design partners”, students provided extensive feedback on many aspects of the course. This was immensely valuable in tailoring worked examples and assignments so that they are more in tune with students’ contexts and interests.

Future iterations involve improving on the curriculum design based on results and student feedback in Study2 and using FACT with broader audiences of middle school students and teachers across the US. Alternate pedagogical strategies will be tested for teaching topics (such as loops and variables) that learners found to be challenging, and in ways that reach all children that have varying levels of prior math preparation.

Any curricular innovation is a continuous process, and any particular version of it is simply a point along the way. This is true of FACT as well. Our learning and findings from these studies present promising directions for future research involving further improvements in FACT. FACT currently embodies a well-conceived, pedagogically robust design of a curriculum for computing in middle school that has been empirically tested for CT learning outcomes in school settings. Nonetheless, it still has room for improvements. This research represents the first two iterations of what should be seen as an ongoing systematic design-based research effort in diverse settings and with broader audiences of middle school students and teachers.

References

- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences, 13*(1), 1-14.
- Barron, B. (2004). Learning ecologies for technological fluency: Gender and experience differences. *Journal of Educational Computing Research, 31*(1), 1-36.
- Barron B., & Darling-Hammond, L. (2008). How can we teach for meaningful learning.? In Darling-Hammond, et al. *Powerful learning: What we know about teaching for understanding*. San Francisco: Jossey-Bass.
- Bell, P., Hoadley, C. M., & Linn, M. C. (2004). Design-based research in education. *Internet environments for science education, 73-85*.

- Brown, J. S., Collins, A., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, 487.
- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences*, 13(1), 15-42.
- Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences*, 11(1), 105-121.
- Engeström, Y. (2009). The future of activity theory: A rough draft. *Learning and expanding with activity theory*.
- Engle, R. A. et al. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist*, 47(3), 215-231.
- Ericson, B., & McKlin, T. (2012). Effective and sustainable computing summer camps. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education* (pp. 289-294). ACM.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95(2), 393-408.
- Grover, S., & Pea, R. (2013). Computational thinking in K-12: a review of the state of the field. *Educational Researcher*, 42(1), 38-43.
- Grover, S., Pea, R. & Cooper, S. (2014a). Remediating misperceptions of computer science among middle school students. In *Proceedings of the 45th ACM SIGCSE (2014)*. New York, NY: ACM.
- Grover, S., Pea, R. and Cooper, S. (2014b). Expansive framing and preparation for future learning in middle-school computer science. In *Proceedings of the 11th ICLS (2014)*, Boulder, CO.
- Grover, S., Pea, R., Cooper, S. (2015). Designing for deeper learning in a blended computer science course for middle school students. *Computer Science Education*, 25(2), 199-237.
- Grover, S., Pea, R. & Cooper, S. (2016). Factors influencing computer science learning in middle school. In *Proceedings of the 47th ACM SIGCSE (2016)*. Memphis, TN. ACM.
- Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. US Dept of Education.
- Meerbaum-Salant, O., Armoni, M., & Ben-Ari, M., (2010). Learning computer science concepts with Scratch. In *Proceedings of ICER '10*. New York: ACM, 69-76.
- Pea, R. D. (2010). *Augmenting educational designs with social learning*. In *NSF SLC PI Meeting*.
- Pellegrino, J. W., & Hilton, M. L. (Eds.). (2013). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.
- Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, 39(4), 199-201.
- Schwartz, D. & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, 22(2), 129-184.
- Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. Mestre. (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1-51). Greenwich, CT, Information Age Publishing.
- Scott, J. (2013). The royal society of Edinburgh/British computer society computer science exemplification project. *Proceedings of ITiCSE'13*, 313-315.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Ed. review*, 57(1), 1-23.
- Soloway, E. (1986). Learning to program=learning to construct mechanisms and explanations. *Communications of the ACM*, 29(9), 850-858.
- Spohrer, J. C. & Soloway, E. (1986) Novice mistakes: are the folk wisdoms correct? *Communications of the ACM*, 29(7), 624-632.
- Tai, R., QiLiu, C., Maltese, A.V., & Fan, X. (2006). Planning early for careers in science. *Science*. 312(5777) 1143-1144
- Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational technology research and development*, 53(4), 5-23.
- Wing, J. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33-36.
- Von Hippel, E., & Tyre, M. J. (1995). How learning by doing is done: problem identification in novel process equipment. *Research Policy*, 24(1), 1-12.
- Zur Bargury, I., Pärvi, B. & Lanzberg, D. (2013). A nationwide exam as a tool for improving a new curriculum. *Proceedings of ITiCSE'13*, 267-272. Canterbury, England, UK.

Acknowledgments

We gratefully acknowledge funding from the National Science Foundation [NSF-0835854, NSF-1343227], and the intellectual contributions of Profs. Stephen Cooper, Daniel Schwartz and Brigid Barron in this work.

Indexes

Author Index

Pages 1-702: Volume 1
Pages 703-1407: Volume 2

- Abdul Waduth, Fauziah Kani, 1306
Abraham, Abigail, 1306
Abrahamson, Dor, 122, 466, 1346, 1353
Acosta, Alisa, 1063
Agesilaou, Andria, 1245
Ahn, June, 1098
Ahonon, Arto K., 839
Ainsworth, Shaaron, 775, 1082
Alameh, Sahar, 1006
Aleven, Vincent, 338
Allen, Carrie D., 442
Allen, Laura, 194
Alqassab, Maryam, 160
Anderson, Craig G., 974
Anderson, Emma, 106
Anderson, Janice, 966, 978
Andrade, Alejandro, 434
Ang, Sian Hong Edith, 1306
Anjewierden, Anjo, 823
Anzai, Yuki, 1207
Applebaum, Lauren, 799
Arastoopour, Golnaz, 631
Arvaja, Maarit, 234
Arvidsson, Toi Sin, 82
Ashley, Robert W., 942
Askari, Emilia, 1138
Asterhan, Christa, 1136
Aw, Yangming, 1306
Awwal, Nafisa, 839
Azevedo, Flávio S., 585, 1098
Bakker, Arthur, 466
Bal, Aydin, 1048
Bamadhaj, Muhammad Helmi Bin Ahmad, 1306
Barber-Lester, Kelly J., 966, 978
Barcellos, Marcília E., 962
Barth-Cohen, Lauren
Barth-Cohen, Lauren A., 386, 1179
Basu, Satabdi, 554
Baxter, Frances, 1163
Becker, Sandra, 1018
Beheshti, Elham, 705
Bell, Adam, 1128
Bell, Philip, 1098, 1128, 1342
Kolikant, Yifat Ben-David, 1055
Ben-Zvi, Dani, 1063, 1335
Bereiter, Carl, 9, 578
Berland, Matthew, 974, 1120
Berner, Annelie, 1239
Bernstein, Debra, 886
Bertram, Charles, 930
Betz, Anica, 815
Bhattarai, Shivaraj, 1338
Bielaczyc, Katerine, 1063, 1090, 1185, 1335
Bilkstein, Paulo, 1346
Binzak, John V., 974
Biswas, Gautam, 554
Blikstein, Paulo, 1277
Blum-Smith, Sarah, 330
Bodemer, Daniel, 274, 458, 906
Borge, Marcela, 266, 878
Boroujeni, Mina Shirvani, 1120
Boston, Carol, 851
Botički, Ivica, 1193
Bouton, Edith, 1136
Brami, Uzi Zevik, 1395
Brandes, Ofra, 1055
Brasel, Jason, 671
Brekelmans, Mieke, 24
Brennan, Karen, 330, 1215
Breuleux, Alain, 1342
Bridges, Susan, 1255, 1356
Brown, David E., 1014, 1267
Buis, Stan C. A., 354
Bumbacher, Engin, 1277
Cacciamani, Stefano, 1195
Cai, Huiying, 990
Cai, Qijie, 795
Cain, Ryan, 1098
Calabrese Barton, Angela, 290, 418
Cannady, Matthew, 1098
Care, Esther, 839
Carsten Conner, Laura, 1269
Castro Superfine, Alison, 1149
Catrambone, Richard, 98
Cavera, Veronica L., 791
Chagab, Noha, 514
Chai, Boon Yen, 258
Champion, Dionne N., 1025, 1033
Chan, Carol, 1255
Chan, Carol K. K., 819, 1273, 1322
Chan, Jody R., 146
Chan, Melvin, 1310
Chan, Puay San, 1302
Chan, Tak-Wai, 1350
Chandrasekharan, Sanjay, 242
Chang, Stephanie, 1041
Charles, Elizabeth S., 1063, 1342
Charoenying, Timothy, 1257

Chartrand, Grant, 1209
 Chase, Catherine C., 186
 Chee, Joon Kit Kelvin, 90
 Chee, Shirleen, 9
 Chen, Bodong, 615, 787, 795
 Chen, De Qi, 1326
 Chen, Guanhua, 1179
 Chen, Jingping, 647
 Chen, Mei-Hwa, 1063
 Chen, Shu, 402
 Chen, Ziyang, 1306
 Cheng, Jiaming, 1235
 Chiam, Kim Yeow, 1326
 Chikunda, Charles, 1048
 Ching, Cynthia Carter, 322
 Chisholm, Cindi, 1286
 Chiu, Thomas K.F., 154
 Choi, Gi Woong, 882, 1243
 Choi, Hyungshin, 1181
 Choi, Koun, 1063
 Christmann, Nina, 918
 Chu, Shien, 178
 Clarke, Sherice, 994, 1106, 1362
 Clegg, Tamara L., 779, 851
 Cocco, Felipe, 1120
 Collier, Wesley, 426, 631
 Collins, Jamie M., 1010
 Cordeiro, Mary, 1249, 1251, 1290
 Cournoyer, Nathan, 737
 Cress, Ulrike, 16, 1333
 Crook, Charles, 1350
 Crouse, Kevin, 791
 Csanadi, Andras, 50, 938
 Cuendet, Sebastien, 522
 Cukurova, Mutlu, 1239
 Cunningham, Jahneille, 226
 Curnow, Joe, 146, 218
 Cutshall, Jessica, 859
 da Silva, Glauco S. F., 962
 Dahn, Maggie, 450
 Dalal, Medha, 1201
 Dalsen, Jennifer, 974
 D'Angelo, Cynthia M., 1211
 D'Angelo, Sarah, 202
 Danish, Joshua A., 130, 450
 Dascălu, Mihai, 914
 Davis, Katie, 1128
 Daxenberger, Johannes, 938
 de Jong, Ton, 823
 de León, Vanessa, 1247
 de Souza, Elisabeth G., 962
 Deiglmayr, Anne, 926
 Deitrick, Elise, 370
 DeLiema, David, 450, 1364
 Derry, Sharon, 966, 978
 DesPortes, Kayla, 890
 DeSutter, Dane, 1082, 1233
 DeVane, Benjamin, 1263
 Dickes, Amanda Catherine, 282
 Dieckmeyer, Emily, 498
 DiGiacomo, Daniela, 482
 Dillenbourg, Pierre, 24, 202, 522, 1114, 1120
 DiSalvo, Betsy, 890
 Do, Ellen Yi-Luen, 1147
 Dobbie, Karen, 1249, 1251, 1290
 Dohn, Niels Bonderup, 1159
 Dohn, Nina Bonderup, 1159
 Domančić, Mirna, 1193
 Dominguez, Kristen, 1189
 Dornfeld, Catherine L. 562, 950
 Dorph, Rena, 1098
 Dragomir, Mihaela, 378
 Drayton, Brian, 886
 Drljević, Neven, 1193
 Duckles, Joyce M., 250
 Dugdale, Michael, 1063
 Duvivier, Robbert, 1405
 Dyer, Elizabeth B., 1033
 Dziallas, Sebastian, 1055
 Easterday, Matthew W., 35, 1157
 Eberle, Julia, 1361
 Echevarria, Ruben, 1025
 Edgar, Ross, 1290
 Eisenhart, Margaret, 442
 Elson, Malte, 918
 Eltoukhy, Moataz, 1179
 Engeström, Yrjö, 1048
 Enyedy, Noel, 450
 Erdmann, Julia, 918, 1217, 1396
 Erkens, Gijsbert, 24
 Erkens, Melanie, 458
 Evans, Chad, 106
 Farris, Amy Voss, 282
 Fathema, Nafsaniath, 950
 Feng, Mingyu, 210
 Ferschke, Oliver, 24
 Fick, Sarah J., 1366
 Fields, Deborah A., 1041
 Fincher, Sally, 1055
 Fischer, Frank, 50, 599, 938, 1227, 1335
 Fishman, Barry J., 1342
 Flake, Sabrina, 815
 Flekova, Lucie, 938
 Fleming, Heather, 1286
 Flood, Virginia J., 122
 Flores Lesama, Manoel, 1048
 Flynn, Paul, 1397
 Forbes, Cory, 639
 Forkosh Baruch, Alona, 1136
 Franklin, Scott V., 930
 Frenza, Jason, 1290
 Froehlich, Jon E., 779
 Fujita, Wataru, 1191
 Furtak, Erin Marie, 1247

Fushikida, Wakako, 1207
 Fynes-Clinton, Elizabeth, 1106
 Fysaraki, Maria, 1227
 Gaesevic, Dragan, 24
 Garner, Brette, 671
 Gaydos, Matthew, 729, 1090
 Ge, Xun, 1259
 Georgiou, Yiannis, 1245
 Gerard, Libby F. 394, 1120, 1152
 Gerber, Elizabeth M., 35
 Gergle, Darren, 202
 Gerritsen, David, 994
 Ghanem, Christian, 938
 Gijlers, Hannie, 823
 Gnesdilow, Dana, 950
 Gobert, Janice, 1002, 1404
 Goh, Wendy, 771, 847
 Goldberg, Tsafir, 1106
 Gomez, Kimberley, 226
 Gomez, Louis M., 226
 Gomoll, Andrea S., 130
 Gonzalez, Carmen, 362
 Grainger, Rebecca, 994
 Gravel, Brian, 370
 Graville-Smith, Cynthia, 954
 Grawemeyer, Beate, 753
 Green, Judith L., 1255
 Greenhow, Chrisine, 1136
 Grimes, Sara M., 1041
 Grover, Shuchi, 695, 1346
 Gu, Xiaoqing, 990, 1350
 Gurevych, Iryna, 938
 Guthrie, Mareca, 1269
 Gutiérrez, José Francisco, 687
 Guzdial, Mark, 98
 Gwee, Susan, 1294
 Habron, Geoffrey, 870
 Hajny, Nik, 1120
 Häkkinen, Päivi, 839
 Håklev, Stian, 114
 Hall, Allison H., 1199
 Halls, Jonathan, 775, 1398
 Halverson, Erica R., 1041
 Hansen, Alice, 58, 753
 Hardy, Lisa, 713, 1399
 Harrer, Benedikt W., 122
 Harris, Christopher J., 1211
 Hartmann, Christian, 530
 Haslir, Mohd Noor Hishamuddin, 1318
 Hecking, Tobias, 918
 Heijnes, Dewi, 1082
 Heimbuch, Sven, 906
 Hensley, Lauren, 859
 Herrmann, Thomas, 918
 Hershkovitz, Arnon, 1136
 Hew, Khe Foon, 679, 1282
 Hickey, Daniel T., 538, 546, 1114
 Hillman, Thomas, 1106
 Hilppö, Jaakko, 1025, 1033
 Hinojosa, Leighanna, 954, 1177, 1231
 Hjorth, Arthur, 894
 Hmelo-Silver, Cindy E., 16, 114, 130, 1255
 Hoadley, Christopher, 1071, 1338
 Hod, Yotam, 807, 1063, 1335, 1368
 Hofer, Sarah, 1167
 Holmes, Wayne, 58
 Hong, Helen, 874
 Hong, Huang-Yao, 787, 1203, 1213
 Honwad, Sameer, 1338
 Hoppe, H. Ulrich, 918
 Hori, Naoto, 1074
 Horn, Ilana Seidel, 671
 Horn, Michael, 705
 Horton, Emily, 226
 Hossain, Zahid, 1277
 Hsu, Pi-Sui, 1183
 Huang, Hai, 1261
 Huang, Jun Song, 998, 1082
 Huang, Kun, 1259
 Humburg, Megan, 450
 Hung, David, 5, 474
 Hußmann, Heinrich, 1227
 Hyvönen, Pirkko, 1169
 Iikubo, Shinya, 1074
 Ikejiri, Ryohei, 1207
 Illum, Randy, 450
 Imelda, Caleon, 24
 Iordanou, Kalypso, 847
 Istance, David, 9
 Jacobson, Michael J., 410, 1071, 1405
 Janssen, Jeroen, 24
 Jeong, Heisawn, 16, 1350
 Jiang, Shiyang, 346, 1179
 Jimenez, Raquel, 1215
 Jin, Yinglian, 402
 Jo, Yohan, 24
 Johnson, Raymond, 1128
 Jona, Kemi, 705, 1025
 Jonathan, Christin, 24
 Joosa, Esther, 1314
 Jordan, Michelle E., 1010
 Jordan-Douglass, Anna, 974
 Jornet, Alfredo, 138
 Joyce, Jeanette, 791
 Judson, Darlene, 986, 1400
 Jung, Yong Ju, 878
 Junokas, Michael, 1271
 Juzar, Masharrat, 1261
 Kabayadondo, Zaza, 827
 Kafai, Yasmin B., 498, 1041
 Kajkara, Manuela, 1193
 Kali, Yael, 4, 1063
 Kang, Seokbin, 779
 Kang, Seokmin, 950

Kapon, Shulamit, 843
 Kapur, Manu, 16, 998, 1063, 1090, 1361
 Kares, Faith, 1229
 Katz, Vikki, 362
 Kaul, Chittaranjan, 506
 Kawasaki, Jarod, 737
 Ke, Li, 803
 Keifert, Danielle T., 1033
 Kern, Anne, 1338
 Keune, Anna, 1041
 Khamis, Nur Azarina, 258
 Khanlari, Ahmad, 1195, 1237, 1275
 Kici, Derya, 1275
 Kiemer, Kathy, 1247
 Kim, Hee-jeong, 1370
 Kim, Hyojin, 1235
 Kim, Mi Song, 1298
 Kimball, Nathan, 1014
 Kirschner, Paul A., 1356, 1361
 Kittinger, Liza, 946
 Knight, Simon, 194
 Koehler-Yom, Jessica, 106
 Koh, Elizabeth, 24, 874
 Koh, Jaime, 1281
 Kollar, Ingo, 50, 599, 811, 938
 Kolodner, Janet L., 1071
 Kong, Siu Cheung, 1350
 Kothiyal, Aditi, 242
 Krajcik, Joseph S., 1211
 Krämer, Nicole C., 918
 Krist, Christina, 306, 894, 1033, 1223
 Kuhn, Deanna, 82, 771
 Kumar, Vishesh, 974
 Kustusich, Mary Bridget, 930
 Kwok, Amy, 679
 Kyewski, Elias, 918
 Kyza, Eleni A., 1245
 Lacelle, Lizanne, 1286
 Laferrière, Therese, 9, 1342
 Lahav, Orly, 514
 Laherto, Antti, 843,
 Lai, Polly K., 410
 Laina, Vasiliki, 934
 Lakhani, Heena, 1128
 Lam, Rachel, 902, 998, 1090, 1372
 Lam, Wan Shun Eva, 1157
 Land, Susan M., 882, 942
 Lanouette, Kathryn A., 570
 Law, Nancy, 16, 679, 761
 Law, Victor, 859, 1259
 Lawton-Sticklor, Nastasia, 1185
 Lay, Hoon Seah, 1382
 Leak, Anne, 930
 Lee, Alison, 186
 Lee, Bing-fai, 1322
 Lee, Christine, 450
 Lee, Chun Kit, 1147
 Lee, Eric Monsu, 1183
 Lee, Jason Wen Yau, 1163
 Lee, Mandi Jieying, 1147
 Lee, Victor R., 170, 1098, 1353
 Lee, Vwen Yen Alwyn, 90
 Lee, Yeung, 679, 761
 Lee, Yew Jin, 1048
 Lee, Yu Ling, 1310, 1318
 Leenaars, Frank, 1082
 Lefstein, Adam, 783
 Lei, Chunlin, 1273
 Lenton, Kevin, 1063
 Lerner, Patrick, 938
 Levrini, Olivia, 843
 Levy, Sharona T., 514
 Li, Pei-Jung, 1203, 1213
 Li, Yi, 970
 Liang, Leming, 761, 1401
 Liles, Melissa, 1163
 Lin, Chiu Pin, 787
 Lin, Feng, 950
 Lin, Hsien-Ta, 1203, 1213
 Lin, Karina, 663
 Linares, Nicholas, 1271
 Lindgren, Robb, 1006, 1014, 1267, 1271, 1353
 Linn, Marcia C., 394, 591, 721, 799, 1120, 1152, 1335
 Littleton, Karen, 194
 Litts, Breanne K., 498, 1041
 Liu, Shiyu, 490, 910
 Liu, Xinyang, 402
 Liu, Yao, 1175
 Liuzzi, Richard, 663
 Lo, Chung Kwan, 1282
 Lockyer, Lori, 1165
 Loibl, Katharina, 58, 530
 Loo, Wei Seng Andy, 1306
 Lotz-Sisitka, Heila, 1048, 1338
 Low, Michelle, 902
 Lu, Lin, 66, 859
 Lu, Weiquan, 1147
 Lucero, Margaret M., 1265, 1374
 Luckin, Rose, 1239
 Ludvigsen, Sten, 16
 Lund, Kristine, 1071, 1335, 1393
 Lyons, Leila, 1120
 Lyu, Xiaoxin, 970
 Ma, Leanne, 615
 Madhok, Jacqueline, 394, 591
 Maggiore, Chrystal, 942
 Mahoney, Charlie, 450
 Mäkitalo, Åsa, 1106
 Malkiewich, Laura J., 186
 Manai, Ouajdi, 1153
 Manches, Andrew, 378, 1353
 Mann, Michele J., 585, 1098
 Margulieux, Lauren E., 98

Markauskaite, Lina, 410, 1405
 Martin, Nicole D., 298, 1402
 Massey, Linda, 9, 1249, 1251, 1290
 Mathayas, Nitasha, 1006, 1014, 1267
 Matsuzawa, Yoshiaki, 615
 Matuk, Camillia, 721, 1120
 Mavrikis, Manolis, 58, 1239
 Mazziotti, Claudia, 58, 753, 926
 McBride, Elizabeth A., 799
 McDonald, Scott, 1063, 1205
 McElhaney, Kevin W., 210, 1211
 McGowan, Veronica C., 866
 McKenney, Susan, 506, 639, 886, 1342
 Megan, Bang, 1128
 Meier, Elisabeth, 1018
 Meixi, 1161
 Meyerhoff, Peter, 1033
 Michalchik, Vera, 1253
 Mierwald, Marcel, 815
 Millet, Chris, 942
 Minor, Rachele, 960
 Minshew, Lana M., 966, 978
 Modir, Bahar, 1225
 Mohd, Shahizha Bte, 1318
 Moher, Tom, 1063
 Mommandi, Wagma, 1128
 Montrieux, Hannelore, 1403
 Morphew, Jason, 1006
 Morrison, Briana B., 98
 Morrison, Deb, 1247
 Moses, George, 250
 Moussavi, Raha, 1002, 1404
 Murthy, Sahana, 242
 Najafi, Hedieh, 1114
 Ng, Kher Hui, 1261
 Nguyen, Linh-Chi, 1147
 Nichols, Emma, 1290
 Nistor, Nicolae, 914
 Norooz, Leyla, 779
 Noventa, Francis, 1290
 Odemwingie, O. Carlette, 960
 Ogan, Amy, 994
 Ogawa, Sonoko, 1074
 Oguamanam, Vanessa, 779
 Oliver, Mary, 775
 Olsen, Al, 1120
 Olsen, Jennifer K., 338
 O'Malley, Claire, 1261
 Ong, Yann Shiou, 266
 Orton, Kai, 705
 Oshima, Jun, 831, 1191
 Oshima, Ritsuko, 831, 1191
 Oura, Hiroki, 1207
 Öztok, Murat, 234
 Pachman, Mariya, 1165
 Panagiotopoulos, Dimitrios, 1082
 Pande, Prajakt, 898
 Parikh, Tapan S., 570
 Park, Juyeon, 1181
 Pasupathy, Sumitra, 1314
 Pauw, Daniel, 851
 Pea, Roy, 522, 695
 Pei, Xinning, 402
 Pekrun, Reinhard, 1018
 Peleg, Ran, 514
 Pellegrino, James W., 1211
 Penney, Lauren, 1025
 Penuel, William R., 482, 855, 1025, 1128, 1342
 Peppler, Kylie, 1041
 Peragine, Wilhelmina, 1215
 Pereira Querol, Marco, 1048
 Perrucci, Vittore, 1195
 Pesianayi, Tichaona, 1048
 Picinatto, Antonio Carlos, 1048
 Plane, Angelisa C., 779
 Plass, Jan L., 1120
 Polman, Joseph L., 954, 1136, 1177, 1231
 Popov, Vitaliy, 354
 Portolese, Alisha, 410, 1405
 Pöysä-Tarhonen, Johanna, 839
 Prain, Vaughan, 1082
 Preece, Jennifer, 851
 Price, C. Aaron, 1229
 Price, Sara, 1353
 Prinsen, Fleur R., 74
 Puhl, Thomas, 1136
 Puntambekar, Sadhana, 298, 562, 950, 1082, 1393
 Pursel, Barton K., 1243
 Quek, Choon Lang, 1281
 Quick, Joshua D., 546, 1114
 Quigley, David, 855, 1128
 Quintana, Rebecca M., 1063
 Rajala, Antti, 1106, 1376
 Ramey, Kay E., 1025, 1033
 Rasi, Päivi, 1169
 Rau, Martina A., 1173, 1378
 Raval, Harini, 506
 Reeve, Richard, 623
 Regalla, Lisa, 1041
 Rehak, Andi M., 130, 1406
 Reimann, Peter, 16, 1071
 Reiser, Brian J., 1219
 Reiss, Kristina, 599
 Ren, Youqun, 1235
 Renshaw, Peter, 1106
 Resendes, Monica, 1249, 1251, 1286, 1290
 Richard, Gabriela T., 1380
 Riedel-Kruse, Ingmar, 1277
 Rienties, Bart, 194
 Ritchie, Allison, 1171
 Rivet, Ann E., 970
 Rodela, Katherine, 226
 Rodriguez, Lilia, 737
 Rojas-Perilla, Diego, 970

Rook, Michael M., 1063, 1205
 Roque, Ricarose, 663
 Rosé, Carolyn P., 24, 607, 1074, 1106, 1114, 1333
 Rosenberg, Joshua, 1223
 Rotsaert, Tijs, 1187
 Rubin, Herbert, 1167
 Ruis, Andrew R. 426, 631
 Rummel, Nikol, 16, 58, 338, 530, 918, 926, 1217, 1361
 Sabri, Nurul Assyifa Ahmad, 1261
 Sagy, Ornit, 807
 Sahed, Suhailin Mohamed, 1306
 Saito, Moegi, 1074
 Saiyed, Gulnaz, 1157
 Sakamoto, Miki, 1155
 Saleh, Asmalina, 450
 Sánchez-García, Raúl, 314
 Sandoval, William
 Sandoval, William A.
 Sandoval, William A., 737, 835, 1090
 Sankaranarayanan, Sreecharan, 607
 Sannino, Annalisa, 1048
 Santo, Rafi, 434
 Sao Pedro, Michael, 1002
 Saucerman, Jenny, 974
 Sayre, Eleanor C., 930, 1225
 Scardamalia, Marlene, 9, 615, 1074, 1237, 1249, 1251, 1275, 1286, 1290
 Schellens, Tammy, 1187, 1403
 Schlottbom, Patrick, 458
 Schnaubert, Lenka, 274
 Schneider, Bertrand, 522, 1346
 Schubert, Jamie, 1173
 Schumacher, Ralph, 1167
 Schunn, Christian, 886
 Schwarz, Baruch, 1106
 Schwarz, Christina V., 803
 Schwendimann, Beat, 1120
 Seah, Lay Hoon, 1090, 1382
 Searle, Kristin A., 498
 Segal, Aliza, 783
 Segovia, Gloria, 1229
 Selvamurthy, Shanker, 1261
 Sengupta, Pratim, 282
 Serviss, Patricia, 1265
 Severance, Sam, 855, 1128
 Sevian, Hannah, 898
 Sezen-Barrie, Asli, 1221
 Shapiro, R. Benjamin, 370, 1055
 Sharma, Kshitij, 202, 522
 Shaw, Bruce W., 9
 Shayan, Shakila, 466
 Shen, Ji, 346, 1179
 Shibani, Antonette, 874
 Shin, Myunghwan, 290
 Shirouzu, Hajime, 1074, 1350
 Siebert-Evenstone, Amanda L., 631
 Sirois, Lindsay, 1249, 1251
 Slater, Stefan, 186
 Slotta, Jame D. 1063, 1114, 1120
 Smirnov, Natalia, 35
 Smith, Blaine Elizabeth, 346
 Smith, Thomas J., 1183
 Smyth, Cliff, 314
 So, Hyo-Jeong, 745
 Solli, Anne, 1106
 Sommer, Stephen, 945, 1177, 1231,
 Sommerhoff, Daniel, 811, 1335
 Souropetsis, Markos, 1245
 Spells, Monet, 890
 Splichal, Jin Michael, 831
 Squire, Kurt, 974
 Stegmann, Karsten, 1227
 Stein, Aliza, 1219
 Steinkuehler, Constance, 974
 Stern, Elsbeth, 3, 1167
 Stevens, Reed, 1025
 Stieff, Mike, 1082, 1149, 1233
 Strachota, Susanne M., 982
 Strijbos, Jan-Willem, 862
 Stromholt, Shelley, 1128
 Stroud, Brandon, 250
 Sullivan, Sarah, 1383
 Sumner, Tamara, 855, 1128
 Sun, Jingjing, 1385
 Sun, Zhiru, 66
 Svihla, Vanessa, 623, 946
 Svoboda Gouvea, Julia, 870
 Swanson, Hillary, 162
 Swanson, Rebecca, 1247
 Swiecki, Zachari, 631
 Tabak, Iris, 1395
 Talis, Vadim, 514
 Tammi, Tuure, 1106
 Tan, Edna, 290, 418
 Tan, Enhui Grace, 1306
 Tan, Esther, 745
 Tan, Hoe Teck, 1298
 Tan, Jennifer, 24
 Tan, Michael, 920, 1041, 1090
 Tan, Samuel, 258
 Tan, Seng Chee, 9, 90, 178
 Tansomboon, Charissa, 1152, 1407
 Tate, Erika D., 210
 Tay, Wan Ying, 258
 Telhan, Orkan, 1041
 Tempelaar, Dirk, 194
 Teo, Chew Lee, 9, 178, 258, 1310, 1318
 Tharuma, Suresh, 1275
 Thomas, Jakita O., 960
 Thompson, John D., 1225
 Thompson, Kate, 870
 Thorn, Christopher, 1153
 Tietjen, Phil, 1063, 1205

Tissenbaum, Mike, 1041, 1120, 1346, 1353
 Tomar, Gaurav Singh, 24, 607
 Trăușan-Matu, Ștefan, 914
 Trninic, Dragan, 1257
 Tsai, Chin-Chung, 787
 Tsovaltzi, Dimitra, 1136
 Tsurusaki, Blakely K., 1269
 Turker, Aybuke, 974
 Tytler, Russell, 1082
 Tzou, Carrie, 1269
 Ufer, Stefan, 599, 811, 862, 1018
 Uk, Irina, 720
 Uttamchandani, Suraj, 546, 1114
 Vaishampayan, Gauri, 1211
 Valkanova, Nina, 1239
 van Aalst, Jan, 639, 819, 1322, 1393
 Van Alstyne, Ryan, 250
 van der Schaaf, Marieke F., 466
 Van Dyke, Margot, 1183
 Van Horne, Katie, 482, 855, 1128
 van Joolingen, Wouter, 1082
 van Leeuwen, Anouschka, 24, 354
 van Riesen, Siswa A. N., 823
 Van Steenis, Erica, 482
 Van Wart, Sarah, 570
 Vanderbeke, Marie, 815
 Vatrapu, Ravi, 1071
 Vedder-Weiss, Dana, 43, 783
 Villanosa, Krystal, 1033
 Visintainer, Tammie, 655
 Vitale, Jonathan M., 591, 799, 1120
 Vogel, Freydis, 599, 1335, 1386
 Walker, Erin, 1201
 Wallon, Robert C., 1014
 Walsh, Elizabeth M., 866, 1189
 Wan, Jarrad, 679
 Wang, Ching-Hua, 1213, 1203
 Wang, Mei, 1235
 Wang, Minhong, 1356
 Wang, Xu, 607
 Warrick, Elizabeth, 851
 Wegerif, Rupert, 1106
 Weidler-Lewis, Joanna, 1177
 Weinberger, Armin, 1136
 Weintrop, David, 705
 Weiser, Gary, 970
 Weiss, Patrice, 1063
 White, Tobin, 713
 Whittaker, Chris, 1063
 Wichmann, Astrid, 918
 Wiedmann, Michael, 58
 Wigmore, Peter, 1082
 Wilensky, Uri, 705
 Wilkerson, Michelle, 934
 Williamson Shaffer, David, 426, 631, 1335
 Williams-Pierce, Caroline, 1241, 1388
 Wilson, Nicholas C., 1253
 Wittmann, Michael C., 386
 Wong, Anita MY, 1255
 Wong, Liyun, 1326
 Wong, Lung-Hsiang, 1090, 1350
 Worker, Steven M., 322
 Worsley, Marcelo, 1346
 Wu, Bian, 1197
 Wu, Han, 795
 Wu, Longkai, 474
 Wu, Sally P., 1173
 Wu, Ying-Tien, 787
 Wylie, Ruth, 1201
 Xiao, Sihan, 402, 835
 Xie, Kui, 66, 859
 Xing, Chenmu, 186
 Yamada, Hiroyuki, 1153
 Yamaguchi, Etsuji, 1155
 Yamauchi, Yuhei, 1207
 Yan, Shulong, 878
 Yang, I-Ting, 1203, 1213
 Yang, Simon, 24
 Yang, Xiaozhe, 1235
 Yang, Yuqin, 819
 Yap, Chong Chieh, 1306
 Yildirim, Hazar, 1163
 Yip, Jason C., 362
 Yoon, Susan A., 106, 114
 Yurkofsky, Maxwell, 330
 Zhang, Jianwei, 1063, 1175
 Zhang, Yibing, 1175
 Zhao, Jian, 1235
 Zheng, Tainian, 402
 Zimmerman, Heather Toomey, 822, 942
 Zufferey, Guillaume, 522

Keyword Index

Pages 1-702: Volume 1
Pages 703-1407: Volume 2

21st century skills	874
accountability	138
action research	1048
active interaction.....	753
active learning	1302
activity theory	130
actor network theory	1205
adaptive scaffolding.....	554
adult learning.....	851
affinity spaces	851
agency.....	946, 1025
agent-based models.....	282
amateur astronomy	585
analogical reasoning	998
analysis	1033
analytic approaches.....	546
anonymity	1187
architecture for learning.....	761
argument skills.....	847, 1294
argumentation	599, 771, 811, 835, 866, 910, 1183
art/science	1267
arts	1263
arts craft	1181
assessment	426, 546, 1179
assessment for learning.....	862
assessment rubric.....	266
attentional anchor	466
audio/video analysis	378
augmented reality	1193, 1245, 1261
augmented reality learning experiences.....	1193
authentic	74
authenticity	815, 843
automated coding.....	938
Big Data.....	974
blended learning	695
boundaries.....	250
boundary objects.....	890
broadening participation	705, 890
calculus	982
case study.....	234, 803, 934, 1048
cell biology	721
citizenship.....	962
civics.....	1157
classroom discourse.....	994
classroom discussions.....	1199
classroom norms.....	1219
classroom orchestration	978, 1120
classroom talk.....	1106
climate literacy	1221

cluster analysis.....	66
coaching.....	506
co-design.....	1128
coding scheme	1195
cognition.....	811, 1241
cognitive and non-cognitive factors.....	1153
cognitive development.....	990
cognitive group awareness.....	458
coherence.....	855
collaboration.....	234, 791, 807, 870, 918, 978, 1010, 1041, 1173, 1191
collaboration scripts.....	599, 1227
collaborative formative assessment.....	647
collaborative imagining.....	122
Collaborative Innovation Networks.....	615, 1237
collaborative knowledge building.....	906, 1219
collaborative learning.....	130, 266, 338, 346, 370, 410, 522, 713, 859, 902, 926, 1207, 1217
collaborative practice.....	1251, , 1286, 1290
collaborative problem solving.....	50, 839
collaborative reflection.....	607
collective knowledge.....	647
collective responsibility.....	615, 1275
communication.....	1010
communication patterns.....	530
communities of practice.....	146, 306
community.....	114, 1055
community college developmental mathematics.....	1153
community transformation.....	250
community-based research.....	655
competency-based education.....	538
complex problem solving.....	902
complex systems.....	106, 514
computational algorithmic thinking.....	960
computational methods.....	1223
computational participation.....	1263
computational thinking.....	705, 1179, 1197
computer models.....	799
computer programing.....	186
computer science.....	370, 1055
computer science education.....	695, 890
computer simulations.....	950, 1014
computer support for collaborative inquiry.....	954
computer supported collaborative learning.....	16, 458, 906, 1231
computer supported curricula.....	106
computing education.....	1215
conceptual change.....	162, 170, 386, 989, 1167
conceptual convergence.....	713
conceptual fluency.....	989
conceptual knowledge.....	58, 530, 1217
conditions of practice.....	482
conduction.....	950
confidence and motivation.....	1209
confidence ratings.....	274
confusion.....	1165
connected learning.....	1041
constructive learning.....	98
constructivism.....	538
conversational functions.....	1195

coordination.....	466
covariation.....	982
creative cognition.....	920
creative computing.....	663
creative learning.....	663
creativity.....	1267
crosscutting concepts.....	970
CSCL.....	354, 562, 607, 787, 874, 878, 990, 1273, 1322
CSCL journal and community.....	16
cultural heritage.....	827
Cultural-Historical Activity Theory (CHAT).....	434, 1221
curriculum.....	886
curriculum development.....	1149
data analysis.....	934
data interpretation skills.....	1002
data management.....	966
data modeling.....	570
data visualization.....	934, 966
DBIR.....	1074
DBR.....	639
deeper school.....	695
delayed instruction.....	410
deliberative democracy.....	1106
deliberative discourse.....	1106
deliberative inquiry.....	1298
democratic education.....	1106
design.....	35, 74, 210, 226, 258, 729, 870, 886, 1010, 1063, 1120, 1211, 1241, 1257, 1263, 1267, 1310
design mode.....	1175
design pedagogy.....	946
design research.....	1155
design thinking.....	827, 878
design-based implementation research.....	1128
design-based research.....	394, 695, 966, 1128, 1161, 1277
design-based science learning.....	322
detail.....	1147
device-mediated collaboration.....	882
dialog.....	771
dialogue analysis.....	914
digital badges.....	538
digital literacy.....	362
digital media.....	1243
disability.....	514
disaster preparedness.....	1209
disciplinary engagement.....	835
disciplinary experts.....	866
disciplinary literacy.....	1199
discourse.....	687, 771, 1173, 1191, 1221
discourse analysis.....	90, 631, 745
discourse analytics.....	24
distributed leadership.....	882
drawing.....	1082
dual eye-trackers.....	522
early science literacy.....	1326
ecological dynamics.....	314
ecological systems.....	570
education.....	815
education.....	1169

educational affordances	882
educational games.....	974, 1006
educational technologies.....	1255
educational technology	591
effective learning teams	874
e-learning.....	918
electronic textiles.....	498
elementary school.....	1179
elementary science education	570
embodied cognition	170, 378, 450, 585, 920, 1233
embodied learning	779, 1271
embodied modeling	386
embodiment	1205
enculturation.....	807
engineering	890, 1010
engineering estimation.....	242
English.....	1294
English composition	1203
English language learning.....	362
enjoyment	1147
environmental education.....	851, 962
environmental science	870
epistemic activities	938
epistemic beliefs	490, 1259
epistemic challenge.....	831
epistemic cognition.....	202, 1223
epistemic emotions	1018
epistemic learning.....	474
epistemic network analysis.....	631
epistemic practice	835
epistemic understanding	1155
epistemic view	787
epistemological cognition.....	426
epistemological framing	1225
epistemology.....	282, 306, 578, 803, 930, 1253
e-Portfolio.....	1163
equity	290, 370, 418
ESL.....	506
e-textiles	1181
ethnomethodology–conversation analysis	122
evaluation of claims.....	835
expansive science learning	1128
experiential learning	1163, 1326
experiment design.....	823
experimental or quasi-experimental design	378
experimental study.....	274
expert study	242
expertise.....	930
explanations.....	1014
exploratory learning environments	58
eye-tracking	202, 466
face	783
families	663
family learning.....	43, 362
feedback.....	910, 1259
feedback design	795
feedback training	862

field trips.....	1326
first person perspective.....	803
flipped classroom.....	66, 1282
formative assessment.....	1211, 1247, 1306
fractions.....	753
framings.....	671
future learning spaces.....	1063
game design.....	960
games.....	186, 450, 729
gender.....	146, 218, 1229
gender differences.....	1167
generative themes.....	962
genetic epistemology.....	466
geography.....	729
gestures.....	1014
gesture-speech mismatch.....	1006
girls.....	960
GIS.....	570
graph-based computer-assisted program.....	1183
graphing.....	591
graphs.....	721
group awareness.....	274
group composition.....	50
group learning.....	386
group products.....	354
group understanding.....	713
guidance.....	823
heat transfer.....	950
heteroglossia.....	1157
high school.....	1294
high school mathematics and science.....	705
high school students.....	1203
higher education.....	791, 807, 1187, 1255
historical concept.....	1310
history or social sciences.....	1048
hobbies.....	585
holistic education.....	1326
idea identification and analysis.....	90
Idea Thread Mapper.....	647
identity.....	234, 655, 1177, 1189, 1267
identity work.....	418
impact.....	639, 1071
implicit guidance.....	458
in situ classroom studies.....	803
infographics.....	954, 1177, 1231
informal education.....	1229
informal environments.....	878
informal learning.....	851, 942, 1025, 1114
informal learning environments.....	1189
informal learning settings.....	954
informal science learning.....	43
Information search.....	362
information visualization.....	1120
innovation.....	9, 615
innovation diffusion.....	474
innovation networks.....	1251
inquiry.....	591, 1265, 1322

inquiry based learning	74
inquiry learning	823, 1201
inquiry stance.....	1185
inquiry-based learning.....	1277
instructional design.....	1298
instructional resources	402
instructional videos.....	1282
integration.....	1326
intelligent tutoring systems.....	58, 338, 1002, 1173
interaction analyses.....	346, 530, 623
interactional ethnography	1255
intercultural learning.....	1163
interdisciplinary	815
Interdisciplinary learning.....	623
interdisciplinary problem solving	870
interest-driven learning	878
interests.....	482, 1098
interests; STEM learning; multiple methods; theory	1098
interleaving	1257
interlocutor	1314
interpretive case study	1177
intersubjectivity	562
intervention.....	811
investigative skills	1298
in-vivo data	1245
joint idea-building.....	266
joint visual attention	522
journalism.....	1157
K-12 education	1197
Kinetic Learning.....	1147
Knowledge Analysis.....	162
knowledge building	9, 90, 178, 306, 578, 615, 647, 787, 819, 986, 1074, 1175, 1195, 1201, 1203, 1213, 1237, 1249, 1251, 1273, 1275, 1286, 1290, 1310, 1318, 1322
knowledge community	1237
knowledge construction	234
knowledge constructive jigsaw	1074
knowledge creation.....	9, 615, 787
knowledge forum.....	1213
knowledge in pieces.....	162, 170
knowledge integration	799
knowledge resources.....	745
language.....	1189
language demands of science.....	1090
language game	43
learning	729, 946, 1136, 1169
learning across settings.....	482
learning analytics.....	24, 202, 426, 546, 1153, 1239
learning by modeling.....	554
learning communities	1063, 1114
learning community	807
learning design.....	498, 1114
learning environment.....	258, 847, 1209, 1233
learning experiences	258
learning gestures	1271
learning innovation	761
learning opportunities.....	671
learning outcomes.....	562

learning perception	1203
learning progression	1247
learning sciences.....	114
learning settings.....	815
learning spaces.....	1205, 1243
learning task.....	990
learning technologies.....	1277
learning through making.....	498
learning visualization.....	434
learning with representations.....	753
legitimate participation	1161
life-long learning	1055
linguistic ethnography	43, 783
literacy education.....	202
literary argumentation.....	1199
low-achieving populations.....	82
LPSV tools.....	779
Macro- and Micro-scripting.....	978
makerspaces.....	920
making	290, 1041
making and tinkering.....	322
marginalization	827
martial arts	314
Massive Open Online Courses (MOOCs)	202, 330, 607, 918, 1207, 1217
math education.....	1165
math learning.....	66
mathematical cognition.....	687
mathematical discourse.....	1302
mathematics.....	154, 314, 338, 599, 862, 1018, 1225, 1241, 1282, 1286
mathematics education	378, 811
mechanistic reasoning.....	894
media commons.....	1243
media literacy	1169
medical education.....	1257
members	114
mental simulation	242
metacognition	1165, 1231, 1273
metadata.....	538
metaphor	314, 970
methodology.....	803, 1245
methods.....	1033
middle school.....	402, 721
mixed methods.....	974
mobile computing.....	942
mobile learning	745, 1193, 1245, 1261
mobile response technology	1187
mobilities of learning.....	290
model dissemination	1215
model-based learning.....	514
modeling.....	282, 803, 1082
motivation.....	1018, 1411
multilevel alignment.....	761
multimedia learning	154
multimodality	346
multiple case study	1227
multiple methods	1098
multivariable reasoning	82

museum learning.....	1261
narrative methods	1055
natural language processing.....	914
natural phenomena.....	775
nature of science.....	578
need for closure	1259
negative feedback	1235
networked communities.....	1249
networks	35
neural networks	989
new creativity	1181
NGSS.....	106, 737, 970
objectification	122, 687
on line learning	1195
online communities.....	1041
online deixis.....	202
online discussions	266
online engagement.....	1153
online forum	1273
online knowledge communities	914
online learning	330, 859, 1114, 1217
online workshop	1207
open ended learning environments	554
open-inquiry	1253
orchestration	35
order thinking skill.....	154
organizational learning	791
out of school	1229
outdoor learning.....	745
outdoor learning environment.....	882
out-of-school learning.....	585
outreach	1071
participatory learning.....	1215
partner modeling.....	458
pathways	250
pedagogical design	942
pedagogical reasoning	994
pedagogical sequence	998
pedagogy.....	1322
peer assessment	910, 1187
peer feedback.....	862
peer review	1155
performance.....	1235
performance trajectory.....	66
persistence	186
personalized learning.....	546
perspectival understandings.....	713
physical and virtual representations.....	1173
physical experimentation.....	950
physics	1225
physiological sensing.....	779
pivotal moments	410
planning	823
play	450
policy	1071
positioning	146
power dynamic	687

p-prims.....	170
practical measurement	855
practice	74, 1253
practice-based learning	1239
pre-calculus.....	982
preparation for collaboration	902, 1090
preservice teachers.....	130, 1185
primary school	338, 679
principle-based approach	1310
prior knowledge	154, 162, 490
problem framing	623
problem-based learning	1255
procedural knowledge.....	58
process tracing methods.....	839
productive failure.....	498, 530, 926, 998, 1090
professional development	226, 330, 737, 1136, 1149, 1247, 1306
professional learning.....	1161, 1249, 1286
professional learning community	178
professional learning networks	1290
professional vision	130
program evaluation	1314
programming	1181
project-based learning.....	855, 946, 1183, 1259
qualitative	1033
quantitative growth	1006
quantitative reasoning.....	1271
quantum mechanics	1225
race	218
reading capacity.....	490
real-time data	1120
reasoning.....	1018
reflection.....	138
reflective assessment	819
reflexive relations	1314
regulation of collaborative learning.....	831
regulatory focus	795
relative expertise.....	1025
relevance.....	843
representation	1082
representational flexibility	753
representational tools.....	954, 1048, 1231
representations of practice	671
research.....	1071
research fields.....	114
research methods	370
research valorisation	639
research-practice partnerships	226
responsive teaching.....	986
revisions.....	1411
rhythms	250
robotics	1179, 1237
robotics programming.....	1197
roles	663
rotating leadership	615
satisfaction	1147
scaffolding	847, 1002, 1197
scale	886

scaling-up	1074
SCAT	960
science	402, 855, 886, 986, 1082, 1211, 1229, 1298, 1411
science and engineering	799
science and math.....	1247
science education.....	282, 410, 514, 554, 562, 679, 775, 843, 894, 942, 994, 1223, 1253
science inquiry.....	450, 1002
science learning	950, 970, 1167, 1261, 1277
science literacy	346
science practices	210, 298, 306, 591, 655, 1219, 1223
science reasoning.....	1014
science teaching.....	298, 737
scientific argument	1155
scientific argumentation	1265
scientific literacy.....	1265
scientific method.....	578
scientific practices	82, 803, 866
scientific reasoning	50, 490, 938
scientific thinking	82
scripts.....	831
secondary school.....	679
segmentation.....	631
self-directed learning	679, 1025
self-efficacy	186
self-explanation	98
self-organization	615
self-regulated learning	274
self-regulation.....	859, 1165
semantic network analysis	1191
senior citizens	1169
shared epistemic agency	1191
simulation	1271
simulations.....	426
situated cognition.....	827
situated learning.....	146, 218
Situated-Action Networks	434
sliding window	631
smooth function	982
social analytics.....	24
social emotional learning.....	1314
social network analysis	90, 434
social networking sites (SNS).....	1136
social practice	442, 482
social science education.....	894
socially shared regulation	859
sociocultural	866, 1063
socio-emotional challenge	831
sociomaterial.....	1205
spacing.....	1257
spatial thinking	1233
standards	1211
standpoint epistemologies.....	218
statistics learning	795
STEM	290, 418, 442, 779, 974, 1098, 1149, 1177, 1233
story.....	623
student engagement	962
student explanations	950

student gestures	1006
student learning	210, 394, 506, 920
student motivation	1306
student reflection	1239
student understanding	930
student voice	1161
student-generated ideas	1090
students with low achievement	819
students with low achievement	1322
students' solutions	926
studio-based learning	35
study abroad	1163
subgoal learning	98
survival analysis	607
sustainability	761
system design	1193
system development	1207
systematic review	1201
tacit knowledge	1263
teacher community	639
teacher education	839, 1221
teacher learning	178, 226, 330, 737, 783
teacher noticing	986
teacher professional development	474, 1215, 1235
teacher workgroups	671
teacher's adaptive expertise	106
teachers	258, 506, 1136
teachers as learners	298
teachers' design capability	1318
teaching	994
teaching and learning	843
teaching strategy	402
team awareness	1227
team-based learning	1302
teamwork	874
technology	9, 210, 394, 721, 771, 799, 934, 978, 1411
Technology Acceptance Model	1243
technology-enabled embodied learning environments	122
technology-enhanced formative assessment	839
technology-enhanced learning environment	242
teleological reasoning	775
temporal synchronicity	354
textual complexity	914
theory	1098
theory improvement	1275
theory of intelligence (TOI)	1235
theory-building inquiries	1275
third space	1157
threshold concepts	989
training	1209
transactivity	354
transfer	998
underachiever	1167
undergraduate education	795, 1149
underrepresented youth	418
unit analysis	138
urban education	442

user interface	938
video	1033
video production styles	1282
video recording glasses	803
video-based professional learning	783
videogames	1241
virtual experimentation	950
visualisations	1239
visualization tool.....	990
voice	1189
Vygotsky.....	138
Web 2.0 tools.....	906, 1294
wikis	906
worked examples	98
workflow.....	966
writing	394, 1294
writing analytics	202
writing performance	1213
writing-to-learn	1265
young children	775
youth identities	442
youth of color	655