Making the invisible visible

Doctoral thesis

MAKING THE INVISIBLE VISIBLE: IMPROVING COLLABORATIVE PROBLEM-SOLVING THROUGH SOCIAL COMPUTING

Steffi Kohl

2023

Cover Steffi Kohl, 2023
Production Off Page Amsterdam, The Netherlands
ISBN 978-94-93278-63-9

MAKING THE INVISIBLE VISIBLE: IMPROVING COLLABORATIVE PROBLEM-SOLVING THROUGH SOCIAL COMPUTING

Dissertation

To obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. Dr. Pamela Habibović,
in accordance with the decision of the Board of Deans,
to be defended in public
on Tuesday 24th of October 2023, at 13.00 hours

by

Steffi Kohl

Supervisor
Prof. dr. Jos G. A. M. Lemmink

Co-supervisor
Dr. Mark Graus

Assessment Committee
Prof. dr. Dominik Mahr (Chair)
Prof. dr. Vanessa Evers, University of Twente
Prof. dr. Alexander Grigoriev
Dr. Sean Rintel, Senior Principal Researcher, Microsoft Research, Cambridge, UK

Everything Not Saved Will Be Lost
- Nintendo "Quit Screen" Message

Acknowledgments

Embarking on this PhD journey has been an incredible adventure, a six-year trek through the labyrinth of academia filled with plot twists and surprises. Numerous individuals have been instrumental in my success and truly deserve acknowledgment. Yet, this acknowledgments section can only contain so much, and I must be content with leaving much unsaid and saying too little about too few people. However, rest assured that every quirk of this journey has left its mark, and I carry the collective wisdom and camaraderie of this incredible expedition with me. I want to thank everybody who has played a role in my life over the past years—a period largely defined by this project. Even if I do not list you here, I am grateful to you. Let's delve into the acknowledgments and celebrate the incredible people who have accompanied me on this journey together.

At the forefront of those deserving my heartfelt appreciation is my promoter Jos, who has been a pillar of support throughout my sometimes unconventional PhD journey. Your guidance was instrumental in shaping my research path. Despite my decision to explore a topic outside your expertise, you stood by me, ready to learn alongside me. Your openness and dedication, even in unfamiliar territory, have unmistakably influenced my development as an academic. I am immensely grateful for your guidance, adaptability, and support, which have played a key role in reaching this academic milestone.

I would also like to acknowledge the contributions of my initial supervisor, Benjamin Lucas, who guided me through the first leg of my journey before embarking on his own academic adventure as an assistant professor in the UK. Your departure, unexpected as it was, marked a turning point in my story, introducing another remarkable character: Mark Graus, my second supervisor and an unexpected hero who rose to the challenge. Mark, you embraced an ongoing thesis, and your expertise and invaluable feedback shaped the narrative of my work in ways I couldn't have imagined. I will always cherish our coffee chats, and teaching alongside you was an absolute pleasure. But the story doesn't end there. In academia, as in life, characters come and go. Mark eventually left academia to move to France, finding the elusive good life - "Audentes Fortuna Iuvat". The final stretch of my PhD was shaped by Kay Schröder, who stepped into the role of my unofficial third supervisor. Kay, your invaluable support extended beyond securing new funding for my second PhD project; you also ignited my passion for data visualization. Thanks to you, I overcame my academic apprehensions and made the life-changing choice to join the HDI Lab at Zuyd Hogeschool, a decision I'll forever cherish. The lab and my now well-acquainted colleagues have seamlessly woven themselves into the fabric of my academic journey.

The final supervisory guidance for this PhD thesis was provided by my thesis committee, consisting of Professor Dr. Dominik Mahr, Professor Dr. Vanessa Evers, Professor Dr. Alexander Grigoriev, and Dr. Sean Rintel. I extend my sincere gratitude to these esteemed members for dedicating their time and effort to review and evaluate my dissertation.

I also want to convey my heartfelt appreciation and gratitude to the following individuals who have made meaningful contributions to or inspired the work presented in this thesis: Emir Efendi, Elena Francu, André Calero Valdez, and Chris de Vries. Further, I want to give a nod to the trusty coffee machine, for being the unsung hero of my research – brewing up inspiration one cup at a time.

However, this acknowledgment would remain incomplete without recognizing the exceptional group of graduate students I've had the privilege to collaborate with. They've been a source of strength, unwavering support, and endless thought-provoking conversations. I extend my heartfelt thanks to Jacky, Hannah, Timna, Gitta, Marcia, Tim, Jenna, Kars, and Ruud, both as my office mates and cherished friends during my adventure of moving from one room to another during my PhD voyage. A resounding applause also goes to Pascalle

and Nicole, our secretaries, whose patience never wavered, even with my one-too-many requests. Furthermore, I want to extend my gratitude to the entire team of professors at the department, whether you're an assistant, associate, or full professor. Each one of you has contributed significantly by providing valuable feedback during seminars, presentations, and our casual conversations by the coffee machine. Last but not least, I want to express my thanks to all those I had the privilege of collaborating with during my four years on the SBE PhD Committee and the two years on the CPCP (the Central PhD Council). Our collective investment of time and dedication to causes I deeply cherish, along with the sense of community we've built, has infused my Ph.D. experience with meaning well beyond my research endeavors.

I'd also like to give a big shout-out to my awesome Paranympfs, Kimberly and Lisi. You both did an incredible job helping me arrange this defense and reception, truly being Steffi's angels in the process. I feel incredibly honored to have you standing beside me today as I defend my PhD. I wouldn't have made it here without your support. Kimberly, we began our journey as colleagues at UM, and though I couldn't believe it when you left me for a while, fate has brought us back together at Zuyd (even though we don't work in the same faculty or building). But regardless of the distance, you've always been more than just a colleague; you're a true friend. Thank you for your unwavering dedication to planning our trips and activities together. No one else could have introduced me to the Dutch culture of "terrasje pakken" and people-watching better than you. And one day, you'll surely develop a taste for spicy food! Lisi, there's no need to convince you of the joys of spicy cuisine; I couldn't imagine sharing amazing Chinese food with anyone else. Thank you for not only providing me with much-needed authentic cuisine but also for being my partner in culinary exploration. Moreover, I appreciate how you've always been there for our coffee breaks, even though I was the only one sipping on the caffeinated brew. Those moments by the coffee machine were more than just caffeine fixes; they were our daily rendezvous for lively conversations and unwavering support. Your presence added a sprinkle of joy to every "early" moring at the office, and I cherish those moments just as much as our spicy food adventures.

Finally, I would like to thank my family and friends for their love, support, and understanding during the long hours spent working on this thesis. Their unwavering encouragement and belief in me have been a constant source of motivation.

Liebe Eltern, ich möchte euch von Herzen danken für eure immerwährende Unterstützung meiner akademischen Ambitionen. Mein PhD wäre nicht möglich gewesen, wenn ihr nicht immer hinter mir gestanden hättet, selbst bei meinen verrückten Ideen. Danke, dass ihr mich sein lasst, wer ich bin, und daran glaubt, dass ich den besten Weg für mich kenne, auch wenn er manchmal viele Wendungen hatte und mich dazu brachte, euch um Hilfe beim Umziehen meiner Habseligkeiten mehr als einmal zu bitten. Ein besonderer Dank geht auch an meinen Bruder, der sich stundenlang bemühte, mir beizubringen, wie man eine Gleichung für x löst, kurz bevor meine Prüfung anstand. Auch wenn ich immer noch nicht verstehe, wie das geht, habe ich es trotzdem geschafft, meinen Doktortitel zu erlangen. Es war großartig, dich als Bruder an meiner Seite zu haben. Vielen Dank, dass du immer in den kniffligen Momenten des Lebens für mich da warst und dass wir unsere Erfolge in der Vergangenheit und auch in der Zukunft gemeinsam feiern werden.

I also want to express my immense gratitude to my Dutch family: Gilles, Elmo, Zoë, Mariet, Jo, Ilse, Jord, Max, and Feye. When I first met Gilles many years ago, I had no idea that our friendship would bring with it an entire extra family to celebrate birthdays and Sinterklaas with. Gilles, it's hard to put into words how much our friendship means to me, and I can never repay the debt I owe you for your unwavering support. Amidst all the chaos, you have always been a consistent source of calm for me. And let's not forget about the emotional support puppy you got me during my PhD, which was a lifeline when I needed it most. Elmo will forever be tiny! I'm incredibly grateful to your entire family for welcoming me with open arms. En een speciaal woord aan Faye: blijf geloven in eenhoorns. Ze bestaan echt!

And last but certainly not least, I want to express my gratitude to my current little family. Vishal, thank you for embarking on this journey with me. Balancing the intricacies of academia and personal life hasn't always been a smooth ride, but your unwavering support during the most challenging moments means the world to me. Your belief in me and our shared dreams have been a source of strength, even when our paths felt uncertain. I cherish the fact that, together with Nala, we've created our own little family here in Liege, and who knows, one day we might finally finish furnishing it! Both of you bring endless smiles to my face and fill my heart with love. Thank you for being an integral part of this important chapter in my life.

In the grand narrative of my PhD, these remarkable individuals have been the characters who shaped the plot, offered comic relief in times of stress, and provided unwavering support through the many plot twists and turns. Their contributions have been nothing short of epic, and I couldn't have asked for a more dedicated and inspiring cast of mentors and colleagues, friends and family to guide me through this scientific adventure.

Steffi Kohl
Liege, September 2023

Contents

List of Figures

List of Tables

1
Introduction

Part of this chapter was presented at the European Conference on Computer-Supported Cooperative Work, 18th Doctoral Colloquium, June 2020

> *The analysis of human behavior in general, and social behavior in particular, is an inherently multidisciplinary problem.*
>
> — Alessandro Vinciarelli et al.
> (Vinciarelli et al., 2009)

1.1 Introduction

The recognition that groups tend to outperform individuals has promoted a shift in organizational focus from individual tasks to team-based tasks (Sawyer, 2011). To support this shift, organizations in diverse industries have increasingly implemented team structures, making team meetings a ubiquitous aspect of organizational life. In knowledge-intensive domains, team meetings have a particularly critical function: Work performance often depends on effective communication between team members to access and utilize relevant information for solving problems, a process referred to as collaborative problem-solving (CPS) (Cross & Cummings, 2004). Furthermore, CPS in these domains often requires high levels of creativity to enable teams to develop new and innovative solutions to novel problems. In these domains, teams often have to work on complex and unfamiliar tasks, with the ability to generate creative ideas critical for finding effective solutions (Kratzer et al., 2004). However, research shows that between 25% and 50% of all meetings are perceived to be of "poor" quality (Allen et al., 2008), with 71% of surveyed managers considering their meetings unproductive (Perlow et al., 2017), often due to low meeting engagement and bad communicative behavior. This has made improving team communication a long-standing area of focus for practitioners (Riedl & Wooley, 2020; Yu, 2005).

Mirroring this trend, research has expanded to investigate the link between team communication, team performance, and creativity (Francu, 2019; Marlow et al., 2018; Salas et al., 2008). Studies have revealed that effective team communication is essential for a team to perform well, enabling team members to share information, collaborate on tasks, and coordinate efforts, and that good communication can lead to increased trust, better decision-making, and improved problem-solving (Parker et al., 2018; Woolley et al., 2015). Elsewhere, research has identified effective communication as critical for accessing and utilizing relevant information and solving complex and novel problems (Stempfle & Badke-Schaub, 2002).

Creativity is also considered an important aspect of team performance. Although most people consider creativity an individual trait, research has proven that creativity can be understood as a social phenomenon, meaning that it is often the result of collaboration and communication between team members (Amabile et al., 1988; Perry-Smith & Shalley, 2003; Woodman et al., 1993). This can be because the diversity of a team's collected perspectives and skills can lead to more creative solutions and new ideas, with effective communication in this context increasing the likelihood that team members share their ideas and collaborate on creative problem-solving (Barczak et al., 2010; Kratzer et al., 2004). This suggests that team performance, team communication, and creativity are interrelated concepts: Effective team communication is crucial for team performance, enabling team members to share information, collaborate on tasks, and coordinate their efforts, which in turn helps foster an environment beneficial for creativity, producing more innovative solutions and improved problem-solving. Consequently, when investigating means of enhancing team performance, team communication represents an important avenue of inquiry.

At its core, team communication relies on the mutual exchange of information across linguistic, paralinguistic, and proxemic modalities. Each modality consists of different cues: Linguistic cues, such as vocabulary and syntax, deliver explicit information; Paralinguistic cues, such as tone of voice and body language, add implicit meaning to verbal communication; Proxemic cues, such as physical distance and posture, help shape the dynamics of a conversation and create social boundaries within a team (Harrigan, 2005; Poyatos, 1983). There is empirical evidence demonstrating that the paralinguistic and proxemic modalities contribute more to the success of a conversation than linguistic modality. Explicit information (i.e., the linguistic modality) in a conversation only contributes 7% to its success. Paralinguistic and proxemic modalities have a much greater impact, with vocal signals (e.g., tone of voice) and nonverbal signals (e.g., body language) contributing 38% and 55% (Mehrabian, 1971).

Team communication has traditionally been studied using observations or (video/audio) recordings. However, this process generates only limited data sets and requires substantial labor expenditure from the researcher, who is limited to attending (or coding, in the case of recordings) only one meeting at a time. Moreover, these means of observation generally limit researchers' access to the paralinguistic and proxemic modalities, which are often difficult to code and sometimes even impossible to assess manually. Furthermore, team communication is dynamic and constantly evolving, making it challenging to capture a snapshot of the communication process at a specific moment in time (Dávila-Montero et al., 2021). In addition to these challenges, studying team communication involves various ethical considerations, including issues of privacy and informed consent, and demands that the researcher ensure that their research does not disrupt the normal functioning of the study subject (i.e., the team). In addition, video or audio recordings can be problematic: Real-world companies often prevent meetings from being recorded due to confidentiality concerns (Pentland, 2010).

Emerging technologies provide opportunities to overcome these challenges by automatically harnessing high-resolution, quantitative, time-series data using, for example, sociometric badges (Solutions, 2014), mobile phones (Eagle & Pentland, 2006), or video annotation tools (Baur et al., 2013). These solutions overcome some of the existing challenges by focusing on collecting paralinguistic and proxemic clues over linguistic clues, enabling the measurement of a variety of factors, including speech volume, tone of voice, laughter, movement, and proximity to others. The rate at which data is collected can vary depending on the specific device and its capabilities. Some devices collect data in real-time, recording every movement and sound made by the wearer. Other devices collect data at a slower rate, for example, recording data only at set intervals or when certain conditions are met.

Apart from new technology being developed to observe teams, technology is increasingly becoming an integral part of how people work and collaborate, including team communication. Teams can now hold virtual meetings using videoconferencing tools that enable real-time interfacing from remote locations, and instant messaging apps and platforms provide a convenient alternative for quick and informal communication. Notably, such technologies also provide the opportunity to seamlessly integrate various existing means to capture information about how teams interact, including, for example, during online meetings (Samrose et al., 2021; Samrose et al., 2018). Using these available online and offline solutions enables researchers to better understand and explore the interplay of communication and team performance (Dávila-Montero et al., 2021; Parker et al., 2018).

Interestingly, computer scientists have been the primary contributors to the literature concerning the development and assessment of the tools and technologies capable of harnessing this high-resolution data on team communication. They have linked interaction efficiency to both the outcomes of meetings and the satisfaction of meeting participants, with lower levels of engagement associated with lower perceived meeting effectiveness, diminished decision quality, and lower collective intelligence (Wang et al., 2021; Woolley et al., 2015; Yoerger et al., 2015), with balanced, active, and equal participation associated with improved team performance (Dong et al., 2012; Leavitt, 1951). However, many of these studies would benefit from incorporating insights from other disciplines engaged in team research. These include the fields of organizational psychology, group studies or management, and business research. Including such insights can enable a more comprehensive understanding of the complex processes and factors that impact team communication and collaboration.

An example of this lack of interdisciplinarity is the failure to incorporate team cognition into the analyses of team communication in the aforementioned studies. Team cognition refers to shared cognitive processes or activities that occur at the team level. Similar to the cognitive processes of individuals, the cognitive processes of teams include reasoning, decision-making, and problem-solving.

According to Cooke et al. (2004), *"Team cognition is more than the sum of the cognition of the individual team members. Instead, team cognition emerges from the interplay of the individual cognition of each team member and team process behaviors."* Team cognition has been conceptualized as both shared mental models (mental representations of a team's task and domain that are collectively shared) and transactive memory systems (group-level knowledge-sharing and memory systems related to a team's task and demands). Recent research supports the possibility that both of these conceptualizations are useful tools for investigations into team performance (DeChurch & Mesmer-Magnus, 2010).

Notably, the relationship between team communication and team cognition is bidirectional. Team communication helps to shape and influence team cognition by providing the means for team members to share information, coordinate actions, and collectively solve problems. Effective team communication can help to build shared mental models and transactive memory systems, which, as mentioned, can improve team performance. Meanwhile, team cognition influences team communication by shaping how team members process and interpret information and how team members make decisions and solve problems. Similarly, it has been argued that team communication is an indicator of team cognition, with communication revealing cognitive processing at the team level (Cooke et al., 2012; Tchupo et al., 2020).

To date, few studies have investigated the association between team cognition and team communication in data-driven modeling of collaborative behavior patterns, including research on the development and assessment of these tools and technologies on team communication during CPS (Gloor et al., 2014; Kidane & Gloor, 2007; Murray &

Oertel, 2018; Parker et al., 2018; Woolley et al., 2010). While the management and psychology literature on CPS has recognized that team cognition can vary throughout the CPS process with different stages requiring unique skills and ways of thinking (Bales & Strodtbeck, 1951; Graesser et al., 2018; Stempfle & Badke-Schaub, 2002; Wiltshire et al., 2018), the aforementioned studies do not include this. Research indicates that teams may need to shift their cognitive state to align with the needs of each stage, such as becoming more creative in the ideation stage. Team dynamics and interaction patterns can also change as teams move through each stage, impacting the overall cognitive state and problem-solving effectiveness of the team (Fiore et al., 2010; Fischer et al., 2007). This thesis adopts the term "mode" to differentiate between the different team cognitive states in these distinct stages.

Some interaction patterns, such as the number of interruptions or the frequency of taking turns, can be beneficial for one mode but detrimental to another (Perry-Smith & Mannucci, 2017). This has recently come into play on a large scale due to the drastic shift from in-person meetings to virtual communication associated with the COVID-19 pandemic, something that has had very different effects on different types of CPS. A study by Brucks and Levav (2022) observed that although changes in communication behavior negatively affected idea generation, positive effects could be observed for decision-making activities. Nonetheless, despite the evidence for this relationship in the theoretical literature, survey-based research, and observational studies, the relationship between team communication behavior and team cognition has not yet been sufficiently investigated via quantitative analysis based on high-resolution time-series data about the social interactions of teams (Graesser et al., 2018).

This thesis aims to make a scientific contribution by adopting an interdisciplinary approach to investigating the relationship between team communication and team performance in the context of CPS. The research explores the composition of modalities for different CPS tasks (Chapters 2 and 3), investigates the impact of these modalities on performance (Chapter 4), and examines how researchers and designers can benefit from these insights when researching CPS or designing support tools for teams working on CPS (Chapter 5).

The outcomes of this thesis have the potential to inform the design of human–computer interaction systems and contribute to knowledge across multiple fields, including social and computer sciences. The interdisciplinary nature of this research should promote the flow of ideas between these fields and provide a holistic understanding of the relationship between team communication and team performance in the CPS context.

1.2 Motivation and Background

This thesis introduces a strand of multidisciplinary research that is substantially informed and motivated by the extant research into human behavior and interactions in four separate disciplines:

1. social psychology,
2. social network analysis,
3. computer-supported collaborative work, and
4. social signal processing.

A key contribution of this thesis is building upon these somewhat disparate research foundations and establishing connections between them. In the following paragraphs, I provide some insight below into these fields of inquiry and highlight the important connections to the work done in this thesis.

Social Psychology Social psychology focuses on the interplay of humans in their social context, investigating the cognitive, affective, and motivational processes underlying social behavior. In the context of this thesis, two subtopics are of relevance: social cognition and social communication. My work primarily draws from social cognition research, an important area to investigate because group cognition features its own ontology, and insights from individually oriented cognition are generally deemed less useful for understanding groups (Cook & Yanow, 1993). However, assessing the social cognition of groups is often difficult and invasive because it involves interrupting the group process using mechanisms such as surveys. Cooke et al. (2004) first proposed that analyzing communication data could enable the assessment of team cognition. Team communication is considered a valid proxy for team cognition, with communication providing insight into cognitive processing at the team level (Cooke et al., 2012; Tchupo et al., 2020). Assessing team communication depends on results from social communication research, which investigates how and why we use language to interact with other people. This research strand investigates the functional role of separate channels (verbal, prosodic, paralinguistic, and kinesic) that together constitute the behaviors characterized as communication (Beattie & Ellis, 2017).

Social Network Analysis Social network analysis (SNA) describes a set of methodological tools concerned with the relationships between social entities, the patterns of these relationships, and the implications of these patterns (Wasserman, Faust, et al., 1994). This field of inquiry builds upon the belief that internal structures can impact the success or failure of societies and organizations and gathers data about networks, primarily from surveys, self-reports, and online behavior, to not only learn the topology but also understand the roles and positions of network members (Freeman, 2004). For this thesis, I use SNA specifically to obtain views of teams as micro-social systems. This involves focusing on the relationships between individuals in a given context rather than viewing individuals as independent and autonomous units (Lusher et al., 2010). Notably, SNA represents an established tool for studying communication patterns that I apply in this context for feature development.

Computer-Supported Collaborative Work Computer-supported collaborative work (CSCW) is the study of people working together while using computer-based tools. Although research in the domain considers various contexts, this thesis specifically draws on applications developed for and insights generated through systems that support communication activities in groups or teams, such as Meeting Meter (Kim et al., 2008), Breakout (Calacci et al., 2016), and MeetingCoach (Samrose et al., 2021). The first project of this thesis uses the sociometric badge, a well-established CSCW tool (Solutions, 2014). The second project utilizes ViCon (Schröder & Kohl, 2022), a custom tool developed based on insights generated from my previous research. That tool is presented in Project One of this thesis.

Social Signal Processing Social signal processing (SSP) is a computing domain aimed at modeling and analyzing social signals present in human–human, and human–machine interactions. The core assumption of this field of inquiry is the belief that social signals are machine-detectable traces of social and psychological phenomena that cannot be accessed via direct observation. This means that social signals are temporal patterns that typically last for short periods of time (e.g., turn-taking is measured in milliseconds) (Vinciarelli et al., 2009; Vinciarelli et al., 2011). Although the domain addresses the modeling, analysis, and synthesis of social signals, this thesis does not engage with synthesis, that is, the generation of artificial social signals. Instead, the focus is on modeling signals to identify principles that govern the use of social signals in teams and advance knowledge more generally. Furthermore, the final chapter gives some consideration to analysis, which, in this context, refers to the automatic detection and interpretation of social signals.

1.3 Dissertation Overview

The title of this dissertation is *Making the Invisible Visible*, a phrase that represents the guiding principle for all the work included in this thesis. Communication, especially the presence of social signals in communication, is often described as unconscious to participants and invisible to the researchers who want to collect data on it. However, if participants cannot report on it and researchers cannot observe it, investigation remains impossible. I hope to contribute to the literature concerning how technology can be used to collect these signals and generate data that can be observed and analyzed by researchers. However, I also want to take a step further: By visualizing this data for participants, I want to demonstrate how such technology can be used to improve collaboration.

This dissertation contains four distinct manuscripts. The first project *(Chapters 2, 3, & 4)* focuses on offline collaboration, with the second project *(Chapter 5)* extending my findings to online collaborations. Because this is a publication-based thesis, there may be some repetition between chapters. Each manuscript has been published with all relevant material for completeness, avoiding readers having to refer to several papers to understand the results presented.

In the following, I briefly introduce each manuscript and detail how each draws from the disciplines mentioned. Furthermore, because the chapters corresponding to my first project utilize the same data set, I subsequently highlight the unique contribution of each chapter relative to the other paper(s). Please refer to the Venn diagram in Figure 1.1 for a visual representation of the interdisciplinary nature of the research presented in this Ph.D. thesis. As shown in Figure 1.1, Chapter 2 connects knowledge from SSP and SNA to generate new insights into the behavioral patterns of interactions. Chapter 3 expands these results by integrating insights from psychology to explain the observed patterns. The chapter further explores how the insights generated can be used for context detection to improve the automated support systems utilized in the CSCW domain. Chapter 4 connects the signals observed during different problem-solving activities to team performance, integrating SSP, SNA, and psychology. Finally, Chapter 5 explores how the insights generated can be practically integrated into a CSCW system.

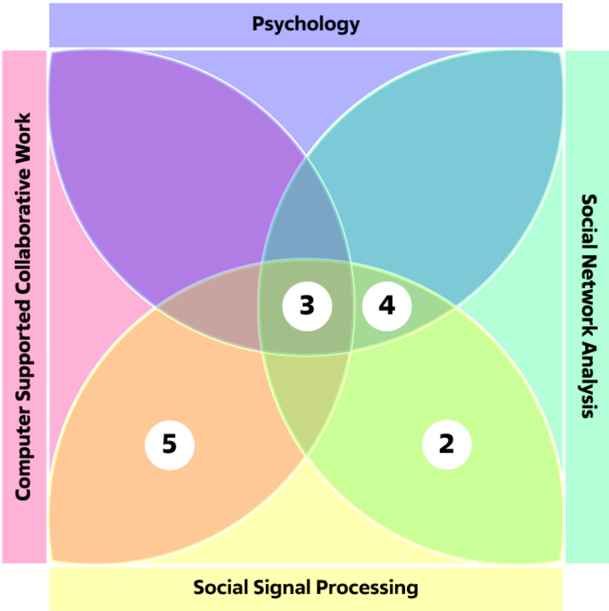

Figure 1.1: The diagram showcases the interdisciplinary nature of the research and its connections, highlighting the scope and focus of the thesis and the interconnections between the research domains. The four main circles represent the domains of computer-supported cooperative work (CSCW), psychology, social signal processing (SSP), and social network analysis (SNA). The overlapping areas between these domains highlight the connections between the research presented in each chapter. Chapter 2 is located in the overlap between SNA and SSP; Chapter 3 is located in the overlap between all four domains; Chapter 4 is located in the overlap between SNA, SSP, and psychology; and Chapter 5 occupies space in the overlap between CSCW and SSP.

Project One
Despite the potential benefits to practitioners and researchers, research into team dynamics has remained elusive because of its complexity and the lack of quantitative measures available. However, wearable electronic devices, such as sociometric badges (Solutions, 2014), have made it affordable to collect detailed information on team communication. Initial studies indicate that these data can help to predict team outcomes with a high level of accuracy (Parker et al., 2018). This deluge of detailed data both unlocks new challenges and provides new insights into collaboration. One of these challenges is the need to move beyond simply collecting these data streams and start questioning the context in which the data are collected. Specifically, most existing studies that rely on social signals to investigate CPS do not differentiate between types of problem-solving activities (Olguın et al., 2009), or they assume that all problem-solving activities will benefit from the same signals. For this project, I collected data from four New Product Development teams over the course of four weeks using sociometric badges.

Chapter 2 explores the integration of social cognition perspectives into SSP research, opening up a novel avenue of inquiry. It presents a proof-of-concept study to demonstrate the potential of this approach. The chapter poses questions that function as the catalyst for my further investigation. It inquires whether we should utilize modeling approaches that stand in contrast to what we know about cognitive processes at the team level. My co-authors and I provide evidence for this claim by demonstrating that modeling the collected signals does indeed reveal significant pattern differences between the different problem-solving modes. In doing so, we reveal what we term the sociometric DNA of an interaction, that is, specific interactions have specific properties of social signals that differ distinctly from other interactions.

The difference between these properties is further investigated in *Chapter 3* via a focus on advanced modeling approaches. We move beyond single-class analysis to provide evidence that different modes of CPS can also be predicted beyond chance in a multiclass prediction task. This chapter further investigates the properties of the social signals using subclass and cluster analysis. This enables us to demonstrate the feasibility of using social signals for context detection. This chapter's main contribution is an expansion of our understanding of how to build context-aware systems that can provide more relevant recommendations to teams working with computer-based support tools.

Chapter 4 finishes the work on the data set by investigating the impact of the detected signals on team performance. This analysis is especially relevant for practitioners, working toward a level of understanding that enables them to provide targeted interventions to teams. We demonstrate differences in the impact of social signals on team performance for different types of problem-solving activities.

Project Two
As a logical continuation of the insights generated by Project One, the last part of my thesis aimed to develop a tool to test whether relevant recommendations could be provided to teams working with computer-based support tools. Although these insights were generated based on a face-to-face data set, the COVID-19 pandemic led to substantial restrictions on in-person research. Furthermore, it dramatically accelerated the implementation of videoconferencing meetings for work. Due to these changes, I joined forces with the Human Data Interaction Lab at Zuyd University of Applied Sciences to develop a videoconferencing tool that could use SSP to generate automated visualization of real-time feedback on team communication behavior. The resulting tool, ViCon (Schröder & Kohl, 2022), aims to balance participation in teams by visualizing the individual contributions of all team members, responding to findings indicating that equal contributions are vital to the success of CPS tasks (Dong et al., 2012). *Chapter 5* reports on a user study deploying ViCon during divergent-thinking tasks. We find that our system positively impacts creativity performance.

2
Deciphering the Code: Evidence for a Sociometric DNA in Design Thinking Meetings

Kohl, S., Graus, M., & Lemmink, J. (2020). Deciphering the code: Evidence for a sociometric dna in design thinking meetings, In *Hci international 2020*, Switzerland, Springer, Cham. https://doi.org/10.1007/978-3-030-60700-5_7.

An early version of this paper was presented at 6th International Conference on Computational Social Science, July 2020 at MIT Media Lab, Boston, MA, USA

Abstract

Despite the increased popularity of virtual teams, in-person teamwork remains the dominant way of working. This paper investigates to what extent social signals can be used to infer the work domain of team meetings. It reveals insights into the complex nature of team dynamics, that are not often quantified in literature, during the design thinking process. This was done by using sociometric badges to measure the social interactions of four teams over a three-week development cycle. From these interactions we were able to discriminate different modes in the design thinking process used by the teams, indicating that different design thinking modes have different dynamics. Through supervised learning, we could predict the modes of Need Finding, Ideation, and Prototyping with F1 scores of 0.76, 0.71, and 0.6 respectively. These performance scores significantly outperformed random baseline models, corresponding to a doubling of the F1 score of predicting the positive class, indicating that the models did indeed succeed in predicting design thinking mode. This indicates that wearable social sensors provide useful information in understanding and identifying design thinking modes. These initial findings will serve as a first step towards the development of automated coaches for design thinking teams.

2.1 Introduction

DNA is a sequence in every cell that defines the characteristics of the organism made up of these cells. In a similar way, social signals such as voice, body motion, and relative location define the characteristics of social interaction; Every interaction is composed of these basic elements. Like the sequence of nucleotides in DNA, we posit that sequences of social signals determine the properties of the interaction. These signals can be read using sensor technology but the scientific community is only beginning to explore their structure. These interactions may show different patterns of signals for different tasks and thus have a different Sociometric DNA. As such, researchers could start relying on these patterns when analyzing work in team interactions.

The majority of studies relying on social signals have not included differentiation of modes (Olguın et al., 2009) or assume that all modes will benefit from the same signals. One example of this is the study by Woolley et al. (Woolley et al., 2010) on understanding collective intelligence in teams. The study makes no attempt to differentiate between different modes in teams when determining collective intelligence, even though participants are asked to solve a wide range of tasks. The study does not attempt to identify how the relationship between different signals and collective intelligence changes for different types of tasks. This shortcoming might explain why turn-taking is the only social signal found to correlate with collective intelligence. The same lack of differentiation between modes can be observed in approaches to understand collaborative problem-solving within the cognitive science literature (D'Mello et al., 2019; Eloy et al., 2019). While the study by Eloy et al. (Eloy et al., 2019) differentiates between goal types it makes no distinction between tasks.

But can we reasonably believe that the same non-verbal dynamics exist between all modes? Is it helpful to identify social signals that predict performance for all modes if we have reason to assume that interaction dynamics are not the same for all modes?

Existing research suggests that distinct phases exist in group problem-solving. Phases are identified with distinct primary needs and while relational and structural elements (such as the number of interruptions or frequency of taking turns talking) are beneficial for one phase, they are detrimental to another. In a similar fashion. Stempfle and Badke-Schaub (Stempfle & Badke-Schaub, 2002) identify four cognitive operations that design thinking teams utilize at different stages and for different tasks. Throughout this paper, the term 'mode' will be used to differentiate interactions for different tasks within teams.

Different modes require different relational and structural elements (Perry-Smith & Mannucci, 2017). Although extensive research has been carried out on sociometric signals in teams, the difference between modes is not often taken into consideration. A study that did use social signals and take into account different modes is failed to draw any conclusions on how patterns of social signals differ over modes (Jayagopi et al., 2010).

For this paper, we rely on robust cues that have been studied in non-verbal communication. Human proxemic and paralinguistic behavior have been studied extensively since the 1920s, in a diverse set of domains such as sociology, psychology (Harrigan, 2005) and, most recently, human-robot interaction (Mumm & Mutlu, 2011). Proxemic behavior relates to how people use space when communicating, while paralinguistic behavior relates to all aspects of spoken communication (such as when someone is speaking, their tone of voice, etc.), except for the semantic content.

Research on organizational science has identified areas in which measuring proxemics and paralinguistics provides important insights. Among others, research shows that face-to-face social interactions play a significant role in the workplace, ranging from job perception (Ibarra & Andrews, 1993) to organizational commitment (Hartman & Johnson, 1989). Furthermore, postural markers have been used in human activity recognition to differentiate team members group functions (Dietzel et al., 2018).

This paper demonstrates that social signals can be used to identify modes in design thinking teams. The specific research questions addressed in this paper are: *1) Can the design thinking mode of an individual in a team be predicted by features extracted from social signals? 2) How do the features extracted from social signals help in predicting individual modes?*

2.2 Meeting Dataset

To investigate the research questions, we collected social signals from a group of young professionals engaging in a new product development (NPD) sprint exercise at a large consultancy. The dataset was collected during the entire four-week sprint from start to prototype development. In total 4 groups with either 5 or 4 members were observed on 13 days in an open space office floor. All teams worked without supervision, so each team structured their work days and scheduled team meetings or requested support from consultants as necessary.

Participants wore sociometric badges during working hours whenever engaged in potentially work-related activities but not during lunch breaks. A sociometric badge is a wearable electronic device capable of automatically measuring social signals, derived from vocal features, body motion, and relative location. In addition, printed questionnaires were used to collect daily information on design thinking activities. The participants were asked to fill in the questionnaire once a day after finishing all work-related activities.

In order to distinguish different modes, a variety of definitions of formal methods that underlie a design thinking approach have been suggested. However, three modes are commonly identified within a design thinking approach: Need Finding are the activities related to problem definition. Ideation is the process of generating ideas and solutions. Prototyping encompasses building models to facilitate the development and selection of concepts (Liedtka, 2015; Seidel & Fixson, 2013). The questionnaire assessed how much time in percentages participants spent each day in these three different design modes. In total, for the 16 days, each team member completed this survey, resulting in a total of 222 responses.

2.2.1 Outcome Variable

The outcome variables, ground truth, or labels to be predicted are defined by the reported percentages for the three design thinking modes by participants. Each observation corresponds to a participant's answer on a given day. For each observation, three variables were created each describing if the participant indicated that the day was spent more than 50% on the corresponding design thinking mode.

This cut-off point was chosen to ensure single labeling. For example, the variable corresponding to Need Finding is positive when participants reported working on Need Finding at least 50% on a given day. Whereas the negative class is those participants who reported less than 50% Need Finding. The threshold of 50% was used for all modes.

2.2.2 Feature Extraction

Many different features are used to assess nonverbal signals, such as turn-taking, activity levels, or proximity within the network (Pentland, 2010). Previous studies have shown that even with a small sample size, patterns measured by sociometric badges were significant and revealing. These patterns can successfully be used to predict a variety of outcomes such as creativity and team performance (Parker et al., 2018).

The features included in this paper have been selected to capture the rotation of non-verbal behavior in teams, specifically rotating leadership (RL), Rotating contributions (RC), turn taking (TTK) and successful and unsuccessful interruptions(SI & UI). The descriptions of these are covered in the following sections. All features are derived from the speech and proximity measurements provided by the Sociometric Solutions software (Solutions, 2014).

Rotating Leadership. Rotating Leadership (Davis & Eisenhardt, 2011; Gloor et al., 2014; Kidane & Gloor, 2007) measures how frequently people change their centrality in the team when represented as a graph. Betweenness centrality is a measure of how central a node in a graph is, and it is calculated by dividing the times the node is located on the shortest path by the total number of paths. In our case, the nodes in the graph are participants, and edges are formed when people are within close proximity as measured by both infrared and Bluetooth sensors. Rotating Leadership counts the number of local maxima and minima in the betweenness curve over time of a person. These peaks and valleys indicate how often people change from a central position to a peripheral one.

The proxemics domain has explored the theme of Rotating Leadership. RL, revealed by changing network structures where people oscillate between peripheral and more central network positions, has a relevant impact on the knowledge-sharing dynamic, affects individual and group creativity (Gloor et al., 2014; Kidane & Gloor, 2007), and is a predictor of innovative performance. The power of RL in predicting innovation performance has not only been observed between individuals but between organizations as well (Davis & Eisenhardt, 2011).

Rotating Contribution. Rotating Contribution measures the oscillation of the Contribution Index (CI) and thus represents how frequently people change the amount of time they spend listening vs speaking (Parker et al., 2018). We calculate the CI of each person ($speaking - listening/speaking + listening$) over time and count the number of local maxima and minima in the CI curve of a person. Just like RL, RC has been shown to be a consistent indicator of creativity. For both signals, more rotation is positively related to performance during creative tasks, while less rotation is preferable for non-creative tasks (Gloor et al., 2014).

Turn Taking. Turns (Chuy et al., 2011; Woolley et al., 2010) are speaking segments that occur after and within 10 seconds of another speaking segment. By default, a speaking segment must be made within 10 seconds after the previous one has ended in order to be considered a turn. Note that the two speaking segments need not be from two different people to count as a turn—a person can pause and then start speaking again. This would count as two speaking segments, and one "self-turn."

Turn-taking in groups has been associated with collective intelligence (Woolley et al., 2010). Chuy et al. (Chuy et al., 2011) argue that taking turns while engaging all team members is crucial for team success. Longer mean speaking segment length, and hence lower number of turns taken has been shown to be correlated with diminished perceptions of individual and group creativity, and imagination, as well as lower levels of involvement and enjoyment (Parker et al., 2018).

Interruptions. An interruption is considered [un]successful (Parker et al., 2018) if the following scenario takes place. Person A is talking, Person B starts talking over A. If Person A talks for [more]less than 5 out of the next 10 seconds, then Person B [un]successfully interrupted Person A. Similar to turn taking, interruptions have been related to creativity and work enjoyment. Successful interruptions show that different members take over and guide the discussion (Parker et al., 2018). By the same token, teams show lower creativity if a few people dominate the discussion (Woolley et al., 2010).

2.3 Predictive Modeling

2.3.1 Data Preparation

The final dataset contained a class imbalance for each of the 3 design thinking modes. This led to the decision to use over-sampling. Need Finding, Ideation, and Prototyping have positive labels for 49 (22%), 57 (26%), and 74 (33%) samples to classify respectively. There are a total of 222 examples to classify. In order to avoid bias, and to compensate for class imbalance and the number of examples, the training sample is over-sampled using SMOTE (Chawla et al., 2002). SMOTE creates new synthetic training examples by creating new examples in between two existing examples in the minority class. This results in a training data set that is more balanced and larger by creating synthetic observations based on the existing observations.

After features were extracted from the raw signals collected by the sociometric badges, they were further transformed by oversampling and feature scaling. The features have different ranges of values. All features were normalized between 0 and 1 using min-max scaling where the smallest value for the feature becomes 0 and the largest 1, and all other values are mapped linearly in between. This prevents features with much larger numeric values from dominating the predictions of the model.

2.3.2 Models and Model Evaluation

Eight different models were trained on the 222 examples, with the aim of comparing linear and nonlinear models as well as instance-based and model-based models. The scikit learn package (Pedregosa et al., 2011) was used to train and test all models. Results are only reported for the three best models. All models are evaluated using leave-one-out cross-validation which is known to be an almost unbiased estimator of model generalization performance on unseen examples (Cawley & Talbot, 2010). The results of the three best classifiers can be seen in Table 2.1. Nonlinear models had higher performance suggesting that there is a more complex relationship between the features and the design thinking mode.

As this first is the approach to tackling design thinking prediction there are no established baselines. Therefore, a stratified random baseline was created where the probability of a label occurring is proportional to the occurrence in the data set. Unpaired sample t-tests were performed between the baseline and trained model. All t-tests were significant ($p < 0.0001$) indicating that the trained models statistically significantly outperform the random baseline models according to the F1

Table 2.1: Performance of the features on predicting design thinking mode

Model	Mode	$F1$	$F1_{pos}$	$F1_{neg}$	Precision	Recall
SVM Linear	NF	0.75	0.41	0.84	0.75	0.75
Random Forest	NF	0.73	0.39	0.82	0.73	0.72
SVM RBF	NF	**0.76**	**0.42**	**0.86**	0.76	0.77
Random Baseline	NF	0.65	0.23	0.78	0.66	0.66
SVM Linear	ID	0.69	0.29	0.83	0.68	0.72
Random Forest	ID	**0.71**	**0.44**	**0.81**	0.71	0.71
SVM RBF	ID	0.67	0.42	0.76	0.69	0.66
Random Baseline	ID	0.61	0.24	0.74	0.61	0.61
SVM Linear	PR	0.66	0.62	0.68	0.74	0.65
Random Forest	PR	**0.68**	**0.54**	**0.76**	0.69	0.68
SVM RBF	PR	0.65	0.60	0.68	0.73	0.64
Random Baseline	PR	0.55	0.34	0.67	0.56	0.56

score. This indicates that the proposed approach has predictive power for design thinking modes.

2.3.3 Results

The results can be seen in Table 2.1. The model performance is reported as $F1$ score, Precision and Recall. In addition, $F1_{pos}$ and $F1_{neg}$ for the positive and negative classes are reported. A random baseline model is reported. The trained models always outperform the random baseline. The best performing classifiers are Random Forest and SVM RBF Kernel with F1 of 0.76 for Need Finding, 0.71 for Ideation and 0.68 for Prototyping. The F1 for the positive class is typically 0.2 higher than that of the random baseline, indicating the positive class is predicted correctly at a rate often approaching double that of the random baseline. This indicates that the models are clearly performing better than expected by random chance and have indeed learned something from the data. These models also show high F1 for predicting the negative class with 0.86, 0.81, and 0.76 respectively, indicating low false positive predictions.

The Random Forest and SVM RBF Kernel are nonlinear models which indicate that there exist predictive performance gains in nonlinear combinations of the features. The results indicate that the features have power of discrimination for predicting design thinking modes. The feature importance scores indicate which signals are important for predicting each of the different design thinking modes. Figure 2.1 shows the feature importance for Random Forests, where UI, TTK, and SI were the top performing features for Need Finding contributing 21.4%, 20.9%, and 20.5% respectively. For Prototyping, similar trends are observed with these features scoring 25.6%, 24.2%, and 24.1%. In contrast, the top performing cues for Ideation are UI, RC, and TTK with 23.9%, 22.0%, and 20.1%. In all three classification models, UI is the top-performing feature. However, it should be noted that the contribution is fairly equal for all features except RL.

While Random Forests show feature importance for design thinking mode, they do not show a positive or negative association. The coefficients of the separating hyperplane from the linear SVM in Figure 2.1 show which features help predict the positive and negative classes. Where the Random Forest feature importance scores are quite similar across modes, the SVM feature scores show different correlations between features and modes. In contrast, the most important feature for all modes in Random Forests is UI, whereas for SVMs this is only the case for Need Finding and Prototyping. For Ideation, it is the third most important feature. It only has a positive association with Need Finding but has a negative association with Ideation and Prototyping. UI is negatively associated with Ideation and Prototyping. It is very clear that features are differently positively and negatively associated with different design thinking modes. This answers our research question that different modes have different dynamics in rotation behavior.

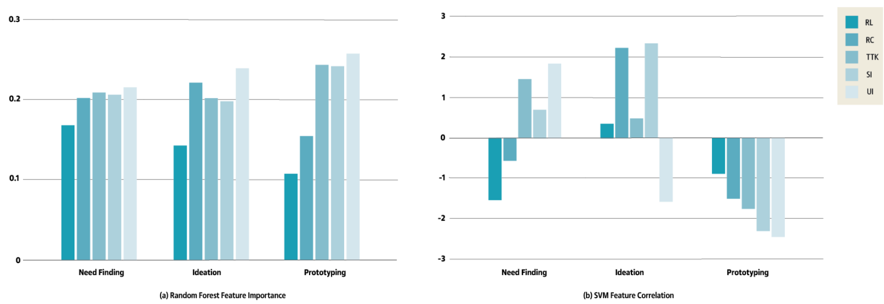

Figure 2.1: Model Features

In conclusion, rotation features can be used to infer design thinking modes. Each design thinking mode has different importance of rotation features. The direction of certain features changes for different modes.

2.4 Discussion and Conclusion

This paper described evidence for a Sociometric DNA in design thinking meetings. This study indicates that differences exist between design thinking modes and can be captured using automatically extracted nonverbal behavior. The sequence of social signals forms the DNA of the interaction within a team and different modes have different DNA which is reflected in different social signal patterns.

To the best of our knowledge, this is the first study to predict design thinking modes for nonverbal behavior despite the advantage of utilizing mode-specific data to understand team dynamics. The signals used in this study were measured using sociometric badges capturing temporal aggregation of nonverbal behavior. The proposed models predict design thinking modes with an F1 of 0.76, 0.71, and 0.68 for Need Finding, Ideation, and Prototyping respectively. The findings of this investigation complement those of earlier studies. These results advance prior research (Jayagopi et al., 2010) by increasing the number of modes predicted and demonstrating that they can be predicted outside of a laboratory setting.

A potential application of these findings is for use in building an automated coach. Relevant interventions can be presented to meeting participants based on their automatically classified design thinking mode. While coaching teams was previously only possible using in-person coaches, this new technology can provide a coach on a much larger scale.

3

Context is Key: Mining Social Signals for Automatic Task Detection in Design Thinking Meetings

Kohl, S., Graus, M., & Lemmink, J. G. A. M. (2022a). Context is key: Mining social signals for automatic task detection in design thinking meetings, In *Design, user experience, and usability: Ux research, design, and assessment*, Cham, Springer International Publishing. https://doi.org/10.1007/978-3-031-05897-4_2.

Abstract

Despite the importance of team communication for successful collaborative problem-solving, automated solutions for teams are notably absent from the literature. One promising avenue of research has been the development and integration of speech-based technology for team meetings. However, these technologies often fall short of meeting the needs of the teams as they do not take meeting context into consideration. In this paper, we demonstrate the efficacy of context detection with data collected during real team meetings. By capturing and analyzing social signals of rotation in team dynamics, we can demonstrate that different stages of collaborative problem-solving using the design thinking methodology differ in their dynamics. Using supervised machine learning, we successfully predict design thinking mode with an overall F1 score of 0.68 and a best-performing sub-class model of 0.94. We believe this to be an essential step towards improving speech-based technology that aims to assist teams during meetings. Making these automated systems context-aware will enable them to provide teams with relevant information, such as resources or guidance.

3.1 Introduction

Collaborative problem-solving (CPS) has become an essential skill critical to efficiency, effectiveness and innovation in the modern workforce. This has placed an increasing amount of attention on teams. Understanding team needs and supporting them while creating minimal disruption to their workflow has been a prominent focal point of team researchers in the last decade (Stewart et al., 2019). One promising avenue of research has been the development and integration of speech-based technology during team meetings (DiMicco et al., 2004; Huber et al., 2019; Kim et al., 2008; Schröder & Kohl, 2022).

While the idea of integrating speech-based support tools into team meetings satisfies the requirement of minimal disruption to the workflow, these technologies often fall short of meeting the needs of the teams (McGregor & Tang, 2017). Past studies indicate that to generate value from these tools, future research needs to focus on understanding meeting context. Making these automated systems context-aware will enable them to provide teams with relevant information, such as resources or guidance (Greenberg, 2001).

This paper contributes to the literature on speech-based support tools by demonstrating the efficacy of context detection during CPS episodes within teams based on social signals alone. We use predictive modeling of proxemic and paralinguistic signals in face-to-face conversational interactions captured through sociometric badges. Research demonstrates the feasibility and utility of leveraging these signals to capture behavioral patterns (Parker et al., 2018).

We aim to demonstrate the value of using these multi-modal signals to detect high-level collaborative patterns such as design thinking contexts. Specifically, we use the proxemic feature, rotating leadership, and the paralinguistic features, rotating contribution, turn-taking, successful interruption, and unsuccessful interruption.

3.2 Related Work & Hypothesis Development

We rely on existing literature when investigating the feasibility of context detection for design thinking modes. The first section highlights theoretical work on convergent and divergent thinking in design thinking that is relevant for the development of the hypotheses. The second section introduces multimodal features collected from sociometric badges, followed by a review of existing speech-based technology approaches that aim to support teams and how context detection poses a challenge to them.

3.2.1 Collaborative problem-solving through Design Thinking

CPS is a critical skill in modern teamwork. It involves two or more people engaged in a coordinated attempt to find a joint solution to a problem by establishing common ground that pertains to the problem space and jointly developing a solution that accommodates multiple perspectives (Roschelle & Teasley, 1995; Stewart et al., 2019).

While there are many methods used to facilitate CPS, Design Thinking (DT) is one of the most popular methods used today, with applications ranging from New Product Development (NPD) to education (Dorst, 2011). This human-centered problem-solving method relies heavily on communication and is used by many companies aiming to foster innovation and generate a competitive advantage (Elsbach & Stigliani, 2018). While there are different governable elements of DT depending on the underlying school of thought, three macro stages can be identified, namely need finding, ideation and prototyping (Brown et al., 2008; Efeoglu et al., 2013). Need finding is the definition of a problem. Ideation is the process of generating ideas and solutions. Prototyping encompasses building models to facilitate the development and selection of concepts (Liedtka, 2015; Seidel & Fixson, 2013). As argued by Kohl et al. (Kohl et al., 2020), research should differentiate between tasks when evaluating communication in teams during CPS to account for the underlying cognitive operations utilized by the team members. As such, this paper will use the term "mode" to differentiate interactions for different tasks within teams during the individual DT stages.

In all DT modes, there are two main types of thinking: divergent and convergent. While divergent thinking aims to find many possible answers or options to a particular problem, convergent thinking narrows down multiple ideas into a single solution (Guilford, 1950). During the DT process, teams use convergent and divergent thinking to explore the problem and the solution space in order to successfully apply CPS. Need finding inhabits the problem space, and ideation and prototyping share the solution space. This process can be visualized as a double diamond as shown in Figure 3.1. Teams working on CPS with the DT methodology will go through this double diamond process iteratively, meaning they will move forwards and backward (Efeoglu et al., 2013).

In a similar manner, existing research on CPS suggests that distinct phases exist in group problem-solving. Each of these phases requires different relational and structural elements (Perry-Smith & Mannucci, 2017). Phases are identified with distinct primary needs. Some relational and structural elements, such as the number of interruptions or frequency of taking turns, are beneficial for one phase but detrimental to another. Similarly, Stempfle and Badke-Schaub (Stempfle & Badke-Schaub, 2002) identify four cognitive operations that DT teams utilize at different stages and for different tasks. Prior studies have not included differentiation of interaction modes (Olguın et al., 2009) or have assumed that all modes will benefit from the same signals.

Given these elemental properties, this paper will quantitatively measure the theoretically proposed elements of convergent and divergent thinking with respect to the impact on context detection for the DT mode.

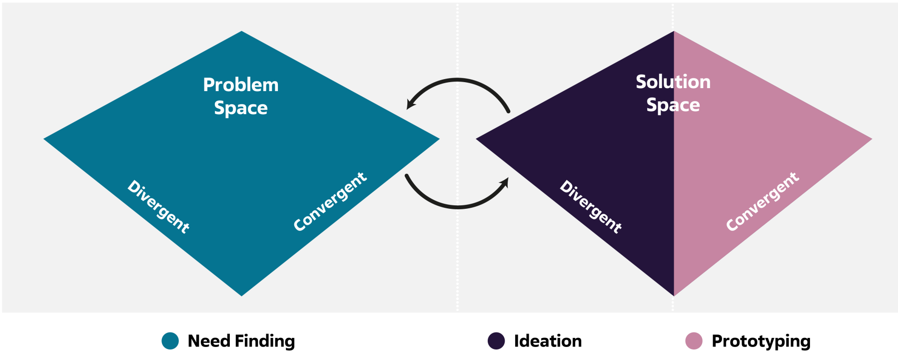

Figure 3.1: The double diamond maps the divergent and convergent stages of a design process to the team modes. Need finding consists of both convergent and divergent thinking, while ideation consists only of divergent and prototyping of convergent thinking.

3.2.2 Modeling Behavioral Patterns During Collaborations

While early work for data-driven modeling on collaboration behavior patterns has mainly aimed to model lower-level behavioral dimensions, such as turn-taking (Pentland, 2010), recent efforts go beyond low-level signals to model high-level collaborative behavioral patterns. For example, postural markers have been used in human activity recognition to differentiate team member group functions (Dietzel et al., 2018), and proxemic features have been shown to be indicators of knowledge-sharing dynamics and affect group creativity (Gloor et al., 2014; Kidane & Gloor, 2007).

As unimodal features often cannot richly capture complex social interactions, multimodal signals have been increasingly used in modeling high-level collaborative patterns. An example of this is the study by Murray and Oertel (Murray & Oertel, 2018) that modeled task performance on a team problem-solving task. They trained a random forest classifier to predict task performance from vocal and linguistic features. The multimodal feature set outperformed the unimodal feature set, demonstrating the added value of multimodal data collection.

Utilizing these signals in successful machine learning models requires theoretical foundations drawn from human–computer interaction (HCI) and organizational science literature with regards to team dynamics. Until recently, these dynamics have remained elusive due to their complexity and lack of quantitative measures. However, wearable electronic devices have made it affordable to collect detailed information on team communication. Research indicates that there is predictive power in social signals collected with these devices (Parker et al., 2018). This is particularly applicable to colocated collaboration settings because face-to-face teamwork remains the dominant mode for solving complex problems despite the increase of virtual teams. Furthermore, colocated collaboration provides unique benefits that are not easy to achieve in digitally mediated forms of teamwork (Olson et al., 2002), such as increasing creativity (Gloor et al., 2014) and performance (Olguın et al., 2009). While preliminary work has demonstrated the feasibility and utility of leveraging multimodal signals to predict behavioral patterns during collaboration activities, more research is needed to understand which data sources work best to predict certain activities.

This paper addresses this gap in the literature by combining proxemic and paralinguistic features. These features have proven to be robust in previous studies in nonverbal communication. They stem from a diverse set of domains such as sociology, psychology (Harrigan, 2005) and, most recently, human–robot interaction (Mumm & Mutlu, 2011). Scholars in these domains have investigated human proxemic and paralinguistic behavior since the 1920s. Proxemic behavior relates to how people use space when communicating. Paralinguistic behavior relates to all aspects of spoken communication, except for the semantic content.

We are collecting these quantitative, non-verbal features using a wearable sensor called the sociometric badge that has been successfully deployed in a variety of organizational contexts with a variety of predictor variables. The badges have been used to predict organizationally relevant outcomes, such as job attitudes and performance (Olguın, 2011), job satisfaction (Olguın et al., 2008), network cohesion (Wu et al., 2008), creativity (Tripathi & Burleson, 2012), group performance (Olguın et al., 2009), and group collaboration (Kim et al., 2008). This study introduces the novel context of predicting DT as a context for the sociometric badge.

The current paper is one of the first investigations using proxemics and paralinguistics multimodal models to understand DT collaboration dynamics. We aim thereby to show the value of using these multimodal signals when modeling high-level collaborative patterns.

3.2.3 Context Detection for Speech-based Technology

Past research has produced many forms of speech-based technology to support meetings for online (Calacci et al., 2016; Faucett et al., 2017; Schröder & Kohl, 2022) and colocated collaboration (Chandrasegaran et al., 2019; DiMicco et al., 2004; Huber et al., 2019; Kim et al., 2008; Nathan et al., 2012). However, these approaches have exclusively conducted data collection in lab settings with short term study setups. In the case of the Meeter study (Huber et al., 2019), data were only captured for a duration of 10 minutes per meeting. While these studies demonstrate the potential of speech-based technology, their limitations also negatively affect ecological validity and limit the extent to which the findings can be generalized. Studies investigating speech-based technology in real team meetings, similar to the data presented in this paper, are still uncommon. Further, these studies do not consider context and are, therefore, unable to provide differentiated guidance or feedback.

A severe shortcoming when it comes to the perceived usefulness of such tools is noted by McGregor and Tang (2017). Their study aimed to investigate if speech-based agent systems could support teams by proactively detecting useful actions that could be presented to a team for improved performance. Their results highlight difficulties in applying automated technology to team meetings, concluding that future research needs to focus on detecting meeting context to produce helpful recommendations from automated systems.

Automated context detection in meetings can be challenging. Research within the domain of context detection focuses on (1) the difficulties associated with defining the set of potential contexts and (2) the feature selection to accurately determine the correct state (Coutaz et al., 2005; Greenberg, 2001), among other elements. As Greenberg states, "Determining an appropriate set of canonical contextual states may be difficult or impossible" (Greenberg, 2001). However, if designers can a priori determine a limited set of likely contexts and what describes them, application building becomes significantly easier. Within the specific context of DT, the contextual states are known. This research focuses on determining features for context detection for this limited set of known contexts.

3.2.4 Hypothesis Development

The literature review outlines the gap of knowledge about the dynamics of team communication during CPS. It further indicates a need to understand the various patterns of communication dynamics that exist among teams working in different modes to provide useful guidance to them. It follows that context detection is essential for the development of speech-based technology. Given this, we will evaluate the research questions: Can the DT mode of an individual in a team be predicted using team rotation features? To answer this research question, this study tests four hypotheses.

Hypothesis 1 *Design thinking modes can be predicted from the selected features with an above-baseline performance of the F1 score.*

Meetings can have more or less focus on a single DT mode. Certain meetings may be dedicated to only ideation, while other meetings may go back and forth between ideation and prototyping. Meetings with more focus on a single DT mode are expected to have more distinct interaction characteristics and thus are expected to be easier to predict.

Hypothesis 2 *Meetings with a higher percentage of a single design thinking mode have better predictability.*

We expect to find unique interaction characteristics for each mode, given the elemental properties of the modes. The predictive models developed to answer H1 are often black-box and do not show the properties of the features in the feature space. We expect a class analysis to better illustrate the properties of the features, demonstrating how each mode has a unique dynamic between the selected features.

Hypothesis 3 *Design thinking modes have different dynamics.*

As highlighted in Section 3.2.1, we expect to find differences between the predictability of the different modes given the theoretical literature as need finding shares divergent and convergent properties that are also found in ideation and prototyping.

Hypothesis 4 *Need finding is harder to distinguish than the other two modes.*

3.3 Methods

3.3.1 Study set up

To investigate the research question, we collected social signals from a group of young professionals (N = 18) engaging in an NPD sprint exercise at a large consultancy. The data set was collected during the 4-week sprint lasting from idea generation to prototype development. In total, for the duration of the sprint, four groups with either five or four members were observed on 13 days in an open-space office floor. All teams worked autonomously without direct supervision by consultants, so each team structured their workdays and scheduled team meetings or requested support from consultants as necessary.

Participants wore sociometric badges during working hours whenever engaged in potentially work-related activities but not during lunch breaks. The badge is an unobtrusive device originally developed by the MIT Media Laboratory and later commercialized by Sociometric Solutions (now Humanyze). This sensor was specifically selected due to its ability to collect multimodal data streams with minimal disruption to the workflow. The sociometric badge is worn around the neck and is the approximate size and shape of an ID tag. It records data via a microphone, an infrared sensor, a Bluetooth detector, and an accelerometer. These four sensors are used to capture individual data about the wearer's voice, body motion, dyadic data of face-to-face interactions, and proximity to other wearers. After undergoing a series of computations, the raw data from these sensors are used to create measures of lower-level behavioral dimensions, such as body movement, colocation, and verbal activity. While the collected data can be considered raw, they are generated using the badge firmware. As such, the data do not necessarily reflect the true values for the external stimuli observed (Solutions, 2014). In addition, the Sociometric Solutions software can enrich the data set by generating additional variables.

Printed questionnaires were also used to collect daily information on DT activities. The participants were asked to fill in the questionnaire

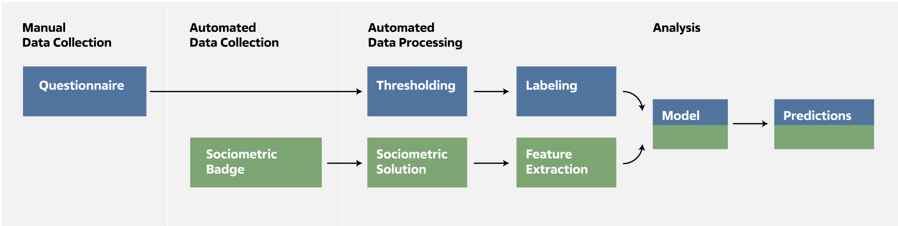

Figure 3.2: Pipeline of data

once a day after finishing all work-related activities. The questionnaire assessed how much time in percentages participants spent on need finding, ideation or prototyping on a given day. The response rate to this question was 100% (222 responses).

Figure 3.2 illustrates the data flow. The manual data collection consists of questionnaires capturing the percentages of time spent in different modes on a given day. These data are used to generate thresholds that create labels. The automated data collection consists of the raw data from the sociometric badges, processed by the Sociometric Solution software, which provides low-level measurements of team activity. The mode prediction model calculates highly predictive features from these low-level measurements. The labels and features are combined with machine learning algorithms to predict the DT mode.

3.3.2 Thresholding the Outcome Variable

The outcome variables, ground truth, or labels are determined by thresholding the percentages reported by participants. For example, at a threshold of 90%, the positive class for need finding is participants who reported need finding at least 90% on a given day. The negative class is those participants who reported less than 90% need finding.

Multiple thresholds have been used for comparison to understand the impact the focus on a single mode has on predictability. While most meetings will use a single mode, that is, a meeting will focus on prototyping only, some meetings will be more diverse and include a mix of, for example, ideation and need finding. Thresholds of 50%, 60%, 70%, 80%, and 90% have been used. The relevant threshold is indicated in the respective tables.

3.3.3 Features

This study uses rotation features to predict the DT mode as described in Section 3.2.2. All signals are collected using the sociometric badges. The conversation characteristics used as features in this study are derived from the speech conditions as provided by the Sociometric Solutions software (Solutions, 2014). The features are defined in the following sections.

Rotating Leadership. Rotating leadership measures how frequently people change their network position in the team. In order to calculate how often people change from a central position to a peripheral one, we count the number of local maxima and minima in the betweenness centrality curve of a person. This is described in Equation 1 and visualized in Figure 3.3.

Higher numbers indicate more rotation of leadership. The figure shows a hypothetical betweenness centrality curve of a person over time. The red Xs mark instances of local minima and maxima as described in Equation 1, where the superscript BC indicates that they are for the betweenness centrality curve, and i indicates the person.

$$RL_i = \#localMinima_i^{BC} + \#localMaxima_i^{BC} \tag{3.1}$$

Rotating Contribution. Rotating contribution measures how frequently people change the amount of time they spend listening vs. speaking. We calculate the contribution index (CI) of each person ($speaking - listening/speaking + listening$) and count the number

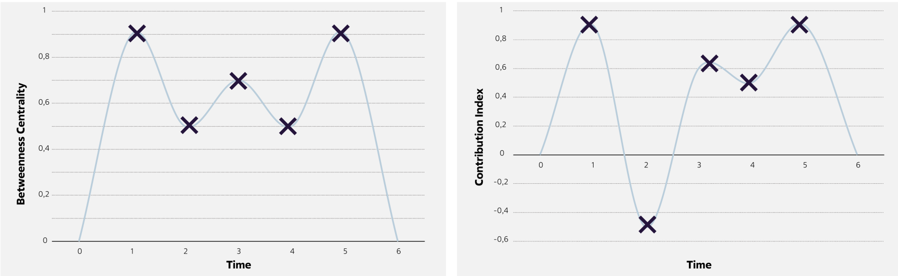

Figure 3.3: Rotating Leadership (left) and Rotating Contribution (right)

of local maxima and minima in the CI curve of a person. This is described in Equation 2 and visualized in Figure 3.3.

Lower numbers indicate fewer changes in the level of contribution for an individual person. The figure shows a hypothetical CI curve of a person over time. The red Xs mark instances of local minima and maxima as described in Equation 2, where the superscript CI indicates that they are for the CI curve, and i indicates the person.

$$RC_i = \#localMinima_i^{CI} + \#localMaxima_i^{CI} \tag{3.2}$$

Turns. Turns are speaking segments that occur after and within 10 seconds of another speaking segment. By default, a speech segment must be made within 10 seconds after the previous one ended in order to be considered a turn. Note that the two speech segments need not be from two different people to count as a turn—a person can pause and then start speaking again. This would count as two speech segments and one "self-turn."

Interruptions. Interruptions in speech can be successful and unsuccessful. Interruptions are defined as a situation during which Person A is talking, and Person B starts talking over Person A. If Person A talks for less than 5 of the next 10 seconds, then Person B successfully interrupted Person A. If Person A instead talks for more than 5 of the next 10 seconds, then Person B unsuccessfully interrupted Person A.

3.4 Results

Multiple models were trained, and DT modes were analyzed. The results support our hypotheses. First, single and multiclass models were trained. These models outperformed the F1 score of the random baseline models, demonstrating that the DT mode can be predicted as stated in hypothesis **H1**. Throughout all of the models and analyses, the threshold for class membership was adjusted to test hypothesis **H2**. This led to the finding that days with more focused DT mode activity can be predicted more accurately.

A class analysis revealed evidence to support the notion that different DT modes have different dynamics as per hypothesis **H3**. Evidence to support hypothesis **H4** was discovered by building models on subsets of the modes, indicating that need finding shares properties with both ideation and prototyping. The following sections will present the results and explain their relationship to the hypotheses.

3.4.1 Predicting Design Thinking Mode

DT mode was predicted for different values of the threshold for class membership. Single class and multiclass models were trained in sci-kit learn using a non-linear Support Vector Machine (SVM) with a radial basis function kernel (Pedregosa et al., 2011). This model showed a greater increase over the random baseline than other models tested, including random forest, logistic regression, AdaBoost, KNN, and naive Bayes when comparing F1 scores to a stratified random baseline F1 score where the probability of a label occurring is proportional to the occurrence in the original data set.

While a single-class model can be used to demonstrate the feasibility of the approach, a multiclass model is more practical for use in a context-aware system. In a single class model, for each input observation, the output describes whether a single DT mode is relevant. A single-class model would have to be constructed for each of the DT modes of interest. Consequently, for an observation, more than one model might indicate their corresponding DT mode to be active, requiring additional logic to make the final decision about which DT mode is active.

The results of the trained models are shown in Tables 3.1 and 3.2. All scores are evaluated using leave-one-out cross-validation, which is known to be an almost unbiased estimator of model generalization performance on unseen examples (Cawley & Talbot, 2010). One hundred experiments were conducted to measure the variability of the models and oversampling. All models showed low variation to initialization with standard deviations between 0.01 and 0.02.

Single Class Model. Table 3.1 shows the results of 1 versus all models at the different thresholds for the three modes. When comparing the baseline to the model performance, different thresholds have higher F1 scores for the different DT modes, showing different underlying dynamics supporting **H3**. For need finding, the F1 score peaks at a threshold of 50% with an improvement over the baseline of 0.12. In contrast to this, ideation peaks at 70% with an improvement of 0.15, whereas prototyping peaks at 90% with an improvement of 0.3.

Multiclass Model. The performance of a multiclass model in predicting DT mode can be seen in Table 3.2. The table also displays the predictive performance for different threshold values. The F1 score of the model peaks at 0.68 when the threshold is at 90%. The best performing model has a 0.25 increase over the baseline model and an F1 score of 0.68. A model with a 70% threshold provides an F1 increase of 0.23 over the baseline model, whereas models with a threshold of 60%, 80% or 50% lead to improvements of 0.2, 0.18 and 0.18, respectively.

The 90% threshold model also shows the highest F1 score for the ideation and the prototyping class with 0.63 and 0.84, respectively, indicating that these classes are identified well by the model. However, the model does not perform as well for the need-finding class, as it only shows an F1 score of 0.27. What stands out in the table is that the need-finding class overall performs worst among all cut-off

Table 3.1: Single class model, where the subscripts of b, p and n correspond to the baseline, the positive and the negative class. The metrics along which the models are compared are F1, accuracy (Acc), precision (P), and recall (R). In addition, F1 and support (S) for the negative and positive class are reported.

Cut Off	Mode	$F1_b$	Acc	F1	P	R	$F1_n$	$F1_p$	S_p	S_n
50	NF	0.64	0.78	0.76	0.76	0.78	0.86	0.49	50	172
60	NF	0.61	0.76	0.74	0.74	0.76	0.85	0.46	49	173
70	NF	0.72	0.76	0.74	0.74	0.76	0.85	0.43	40	182
80	NF	0.66	0.53	0.57	0.71	0.53	0.62	0.39	27	195
90	NF	0.68	0.67	0.69	0.71	0.67	0.78	0.37	16	206
50	ID	0.58	0.67	0.67	0.67	0.67	0.76	0.48	79	143
60	ID	0.53	0.67	0.67	0.67	0.67	0.76	0.49	56	166
70	ID	0.53	0.68	0.68	0.7	0.68	0.75	0.53	47	175
80	ID	0.66	0.76	0.77	0.78	0.76	0.83	0.61	33	189
90	ID	0.77	0.88	0.87	0.87	0.88	0.93	0.61	12	210
50	PR	0.48	0.68	0.68	0.72	0.68	0.67	0.68	82	140
60	PR	0.54	0.69	0.69	0.74	0.69	0.68	0.69	74	148
70	PR	0.45	0.7	0.7	0.74	0.7	0.69	0.71	68	149
80	PR	0.62	0.75	0.74	0.77	0.75	0.71	0.78	59	163
90	PR	0.46	0.77	0.76	0.77	0.77	0.65	0.82	45	177

values. This provides partial support for hypothesis **H4**, indicating that need finding can not be separated as well as the other classes as suggested by the theoretical framework of convergent and divergent thinking. The following section investigates this problem in more detail by conducting a sub-class analysis.

Model and Threshold Conclusions. In conclusion, we are able to predict DT modes with F1 scores above the baseline models, supporting **H1**. Further, increasing the threshold has a positive effect on the ideation and prototyping classes in single-class models. When used in the multiclass model, the overall predictability increases even though

Table 3.2: Multiclass analysis, where the subscripts of b, nf, id and pr correspond to baseline, need finding, ideation and prototyping. S indicates the support for a class.

Cut Off	$F1_b$	Acc	F1	P	R	$F1_{nf}$	$F1_{id}$	$F1_{pr}$	S_{nf}	S_{id}	S_{pr}
50	0.33	0.54	0.51	0.55	0.54	0.32	0.51	0.64	50	79	82
60	0.35	0.58	0.55	0.6	0.58	**0.43**	0.5	0.67	49	56	74
70	0.34	0.6	0.57	0.61	0.6	0.33	0.56	0.71	40	47	68
80	0.38	0.62	0.60	0.59	0.62	0.3	0.58	0.75	27	33	59
90	0.43	0.73	**0.68**	0.7	0.73	0.27	**0.63**	**0.84**	16	12	45

the need-finding predictability suffers. These observations provide partial support for **H2** and indicate that further work is needed to better differentiate need finding.

3.4.2 Understanding Design Thinking Modes

In the previous section, different DT modes could be successfully predicted, outperforming random baseline models. This suggests that the modes are distinguishably different from each other with regard to the underlying feature dynamics. This led us to investigate hypothesis **H3** by analyzing the relationship between the features and the DT modes. Specifically, we conducted a class analysis using three different approaches to understand the structure of the features in the feature space.

For approach 1, a mean was produced for every mode by taking the average of all features for all examples in the class: a class centroid. A mean was created to reduce the effect of outliers. The Euclidean distance between these class centroids was measured as shown in Table 3.3. A smaller number indicates the modes are more similar. As the threshold is increased, the modes become more separated where the effects between need finding and prototyping and between need finding and ideation do not always increase; that is, they are non-monotonic.

This indicates that the modes have different dynamics with respect to the rotation features. If they were similar, the numbers would be close to zero as there would be a minimal distance between the features in the feature space.

In approach 2, we calculated the average Euclidean distance between all pairs of points in two modes. In contrast to approach 1, this approach is more sensitive to outliers as the mean of the features was not taken first. The outliers are directly compared and added to the average distance. This is reflected in the results. Similar to approach 1, a higher number indicates a larger distance between the modes. As shown in Table 3.3 the between-class similarity scores indicate that all modes have different underlying dynamics. Unlike approach 1, the between-class similarity shows that at most thresholds, all modes are equally separated.

Approach 3 highlights the density of a given mode in the feature space. The average Euclidean distance between all pairs of points in each mode was calculated. The results in Table 3.3 show that prototyping is the most dense mode with the lowest average Euclidean distance between all pairs of points at any threshold. As the threshold is increased, the modes become more dense by removing less-focused meetings from the data set.

All three approaches indicate that the underlying features for the individual DT modes are different and show different dynamics. Especially, prototyping shows different dynamics of the features compared to the other modes, as the numbers in Table 3.3 indicate that it is better separated and more compact. In contrast, need finding shows more similarity to the other modes overall. We find support for **H3**, as all measures show differences between the modes.

Table 3.3: Results of class analysis with three approaches: distance between class means (Approach 1), average distance metween class observations (Approach 2) and class density (Approach 3)

	Distance Between Class Means			Average Distance Between Classes			Class Density		
Cut off	NF-ID	NF-P	ID-P	NF-ID	NF-P	ID-P	NF	ID	P
50	0.1	0.28	0.25	0.48	0.41	0.41	0.46	0.49	0.23
60	0.11	0.31	0.27	0.51	0.45	0.43	0.5	0.51	0.24
70	0.1	0.26	0.31	0.48	0.41	0.45	0.45	0.5	0.24
80	0.17	0.26	0.4	0.54	0.39	0.53	0.45	0.6	0.22
90	0.29	0.32	0.56	0.78	0.67	0.78	0.77	0.71	0.46

3.4.3 Evidence for Convergent and Divergent Thinking in Design Thinking Modes

The previous section showed that prototyping is better separated, more compact, and is easier to predict. Need finding tends to be more similar to ideation and prototyping. In the theoretical framework of convergent and divergent thinking (Efeoglu et al., 2013), need finding consists of convergent and divergent thinking, whereas ideation is convergent thinking, and prototyping is divergent thinking. While this overlap of elemental properties can be observed in the class analysis, it becomes clearer in a sub-class analysis where all combinations of models were built using only two of the three modes. The results are shown in Table 3.4.

The best ideation and prototyping model outperforms the other two best models with an improvement over the random baseline of 0.3. The improvements for the other two models are 0.16 (NF-I, 90%) and 0.17 (NF-P, 80%). Furthermore, an F1 score of 0.94 is reached for the ideation and prototyping model, which approaches a perfect F1 score of 1.0. In other words, these findings show that while ideation and prototyping can be distinguished almost perfectly, it is more difficult to identify the difference between need finding on the one hand and

Table 3.4: Subclass analysis, where the subscripts of b, nf, id and pr correspond to baseline, need finding, ideation and prototyping.

Cut Off	Model	$F1_b$	F1	P	R	$F1_{nf}$	$F1_{id}$	$F1_{pr}$
50	NF-I	0.51	0.57	0.62	0.63	0.29	0.75	-
60	NF-I	0.5	0.58	0.66	0.62	0.41	0.72	-
70	NF-I	0.5	0.52	0.52	0.53	0.39	0.62	-
80	NF-I	0.51	0.54	0.54	0.55	0.45	0.62	-
90	NF-I	0.54	0.7	0.72	0.71	0.78	0.6	-
50	NF-P	0.5	0.67	0.72	0.69	0.57	-	0.76
60	NF-P	0.51	0.71	0.75	0.73	0.56	-	0.81
70	NF-P	0.53	0.71	0.76	0.74	0.52	-	0.82
80	NF-P	0.55	0.72	0.75	0.76	0.46	-	0.84
90	NF-P	0.63	0.74	0.79	0.79	0.38	-	0.87
50	I-P	0.50	0.67	0.68	0.68	-	0.63	0.71
60	I-P	0.49	0.72	0.72	0.72	-	0.64	0.77
70	I-P	0.51	0.76	0.76	0.77	-	0.69	0.81
80	I-P	0.52	0.84	0.85	0.85	-	0.77	0.89
90	I-P	0.64	0.94	0.95	0.95	-	0.86	0.97

ideation and prototyping on the other. This supports hypothesis **H4**, indicating that need finding shares divergent thinking with ideation and convergent thinking with prototyping.

3.5 Conclusion

Our research set out to answer the question: Can the DT mode of an individual in a team be predicted using team rotation features? Using proxemic and paralinguistic signals in face-to-face conversational interactions captured through sociometric badges, we introduce novel predictive models of the DT mode. We analyzed a data set of NPD teams using design thinking, and we found that the DT mode can be predicted well above random chance by the predictive models, thus reproducing the self-reported labeling of the recordings on previously unseen examples.

We answered our research question to what extent DT modes can be predicted from social signals. With an overall F1 score of 0.68 and a best-performing sub-class model of 0.94, the analysis demonstrated the efficacy of context detection during CPS episodes within teams based on social signals alone. The analysis shows the potential and challenges involved in detecting context during DT meetings and thus supports the idea of further studying context-aware systems which can provide teams with relevant information, such as resources or guidance.

However, our findings should be interpreted in consideration of the small sample size. Future research should focus on increasing the sample size, expanding the set of contexts detected, and revisiting the selected feature to optimize model performance.

4

Zooming in on the Effect of Sociometric Signals on Different Stages of the Design Process

Kohl, S., Schröder, K., Graus, M., Efendic, E., & Lemmink, J. G. (2022b). Zooming in on the effect of sociometric signals on different stages of the design process, In *Proceedings of the 14th conference on creativity and cognition*, Venice, Italy, Association for Computing Machinery. https://doi.org/10.1145/3527927.3532810.

Abstract

Collaborative creativity is an essential part of modern teamwork and is often supported by formal techniques, such as design thinking. Current support tools are often limited in scope as understanding the time-varying nature and structure of team communication is insufficient. We investigate how collaborative creative activities in new product development teams can be digitally supported while maintaining face-to-face communication. This work analyzes to what extent paralinguistic and proxemic features of team interaction relate to performance in new product development teams and if and how this relationship differs for different stages in the design process. This is investigated by applying multilevel modeling on data collected during a four-week new product development cycle. The cycle was completed by four teams, during which data were collected automatically using sociometric badges that capture social signals of team interactions. In addition, the data are combined with survey-based measurements on the team's daily design process and periodic performance evaluations. The current paper provides evidence that social signals are related to team performance and that this relationship varies across the stages in the product design process. Certain social signals contribute positively in one stage but less in other stages, showing the importance of using multimodal signals when modeling high-level collaborative patterns. This research contributes to the literature by providing a better understanding of relevant factors when designing supporting tools or methods for collaborative creative problem-solving.

4.1 Introduction

Organizations increasingly realize that teams can be a highly effective working unit when striving for creativity in the workplace (Anderson et al., 2014). However, not all teamwork is considered good teamwork and, as Salas and Reyes (Salas et al., 2018) write, *"a team of experts does not necessarily make an expert team."* In other words, extensive domain knowledge does not prevent a team from failing if its members do not know how to work together effectively. Indeed, research on collaborative problem-solving has found that more often than not, collaboration results in *"process loss"* instead of *"process gain"*; group interaction fails to yield performance that exceeds that of the individual group members (Hesse et al., 2015). Collaborative problem-solving requires more than simply joining forces. However, the field remains mute on what determines the success of collaborative problem-solving.

Teams need to exchange knowledge and information while coordinating skills to stimulate idea generation. We will refer to this process of *communication, coordination and interpretation* as collaborative creative problem-solving (CCPS) (Hilliges et al., 2007). While prior research shows that communication contributes to effective CCPS, we possess only limited knowledge of how team communication contributes to performance. Team communication relies on mutual information exchange across various modalities, including verbal (i.e., paralinguistic) and nonverbal (i.e., proxemic) modalities. Paralinguistics relates to all aspects of spoken communication except the semantic content. Proxemics relates to how people use space when communicating.

These features have proven to be robust in previous studies of nonverbal communication. They stem from a diverse set of domains such as sociology and psychology (Harrigan, 2005) and, most recently, human–computer interaction (HCI) (Mumm & Mutlu, 2011). Scholars in these domains have investigated human proxemic and paralinguistic behavior since the 1920s. The utilization of paralinguistic and proxemic modalities has shown to vary across tasks, and some studies have found a structural difference between communication patterns in teams during different CCPS tasks (Jayagopi et al., 2010; Kohl et al., 2020).

One reason for the variability across tasks seems to be related to team cognition. Indeed, it has been argued that team communication is an indicator of team cognition as communication reveals cognitive processing at the team level (Cooke et al., 2012; Tchupo et al., 2020). Further, research has identified that team cognition can vary throughout the CCPS process depending on the group's cognitive state. For example, Stempfle and Badke-Schaub (Stempfle & Badke-Schaub, 2002) have identified different cognitive operations teams utilized for different tasks during the design thinking process, a method used for CCPS (Figure 4.1).

Consequentially we propose that when investigating communication in teams during CCPS, research should differentiate between tasks when evaluating team performance to account for the underlying cognitive operations utilized by the team members. Throughout this paper, the term *mode* will be used to differentiate these interactions for different tasks within teams. Few studies have investigated the association between different modes and social signals on performance. The current paper is one of the first investigations using proxemics and paralinguistics multimodal models to understand design thinking collaboration dynamics. We aim to demonstrate the value of using these multi-modal signals when modeling high-level collaborative patterns to understand team performance.

This study investigates the relationship between social signals and overall team performance and how that relationship differs across modes. We use a data set collected during real-world meetings of new product development teams to answer this question. Using multilevel modeling, we observe that the relationship between the proxemic feature *rotating leadership* on performance is significantly moderated by mode. We demonstrate that variation between working individually, in small subgroups and in full teams collaborating in close proximity is more beneficial for performance in the need-finding

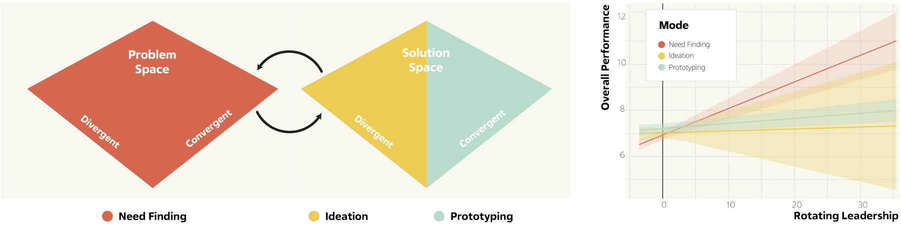

Figure 4.1: Left: Double diamond framework illustrating how the design thinking modes are separated into convergent and divergent cognitive operations. Right: Interaction plot showing the effects of rotating leadership on team performance for different modes.

phase of the product development process than in the prototyping phase.

Our contribution is that we have found a measurable interaction effect of *mode* on the relationship of *social signals* on *team performance*. This finding is important as the literature thus far does not differentiate between CCPS tasks, ignoring the underlying cognitive operations of the teams. This highlights a significant shortcoming of the current literature that could limit understanding: the effects of different social signals during specific modes might go unnoticed when they are combined within the same analysis. We believe that the conclusions drawn from this analysis are applicable to research on team performance and contribute to the foundations for future research aiming to enhance collaborative capabilities, such as through automated support tools for CCPS or novel design thinking process frameworks. Both these research streams need to consider the underlying cognitive operations to help teams perform better.

4.2 Related Work

4.2.1 Collaborative Creative Problem-Solving Through Design Thinking

Design thinking is commonly used as a methodology by teams working with creative or innovative processes within companies. It involves two or more people engaged in a coordinated attempt to find a joint solution to a problem by establishing common ground that pertains to the problem space and jointly developing a solution that accommodates multiple perspectives (Dorst, 2011; Stewart et al., 2019).

Various definitions of formal methods that underline the design thinking methodology exist in the literature. However, three modes commonly cited within a design thinking approach are need finding (NF), ideation (ID), and prototyping PR. NF is the process of defining the problem. ID is the process of generating ideas and solutions. PR encompasses building models to facilitate the development and selection of concepts (Liedtka, 2015; Seidel & Fixson, 2013). We chose design thinking because it is a commonly applied methodology that entails two very distinguished cognitive operations: divergent and convergent thinking. While divergent thinking aims to find many possible answers or options to a particular problem, convergent thinking narrows down multiple ideas into a single solution (Guilford, 1950).

During the design thinking process, teams use convergent and divergent thinking to explore the problem and solution space to apply CCPS successfully. NF inhabits the problem space, utilizing a mix of divergent and convergent thinking. ID and PR share the solution space while utilizing only divergent and convergent thinking, respectively (see Figure 4.1).

4.2.2 Modeling Behavioral Patterns During Collaborations

Research on the impact of communication on team performance is spread across various streams of literature. The most prominent re-

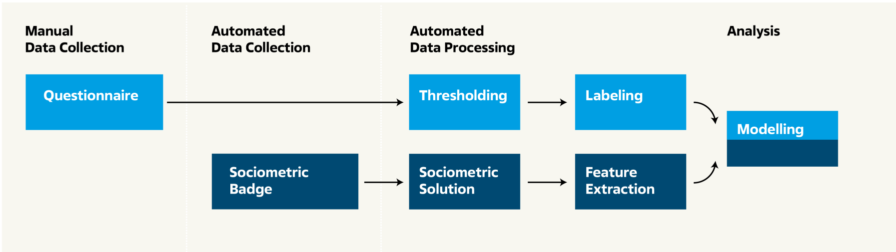

Figure 4.2: Overview of the data collection and analysis process

search streams are HCI, organizational science, and cognitive science literature. This research has identified a variety of factors found to predict team-level task performance (Penzkofer et al., 2021). One prominent research stream focuses on evaluating social signals, such as proxemic and paralinguistic behavior, during team interactions (Kubasova et al., 2019; Murray & Oertel, 2018; Zhong et al., 2020).

However, few studies have investigated the association between different modes and social signals on performance. Prior work is often limited to investigating a single type of mode (D'Mello et al., 2019; Eloy et al., 2019; Kubasova et al., 2019; Schröder & Kohl, 2022), with studies such as Eloy et al. (Eloy et al., 2019) differentiating between goal type but not between modes. Studies that do include multiple modes assume that all will benefit from the same social signals, as seen in the study by Woolley et al. (Woolley et al., 2010) on understanding collective intelligence in teams. While that study asks participants to solve a wide range of tasks, it does not attempt to identify how the relationship between different signals and collective intelligence changes for the different participant modes.

While early work for data-driven modeling on collaboration behavior patterns has mainly aimed to model lower-level behavioral dimensions, such as turn taking (Pentland, 2010), recent efforts go beyond low-level signals to model high-level collaborative behavioral patterns. For example, postural markers, such as forward and backward leaning, have been used in human activity recognition to differentiate team member group functions (Dietzel et al., 2018). Proxemic features, such as time spent in close proximity, are indicators of knowledge-sharing dynamics and affect group creativity (Gloor et al., 2014; Kidane & Gloor, 2007).

As uni-modal features often cannot richly capture complex social interactions, multimodal signals have been increasingly used in modeling high-level collaborative patterns. An example of this is the study by Murray and Oertel (Murray & Oertel, 2018) that modeled task performance on a team problem-solving task. They trained a random forest classifier to predict task performance from vocal and linguistic features. The multimodal feature set outperformed the uni-modal feature set, demonstrating the added value of multimodal data collection.

Utilizing these signals in successful machine learning models requires theoretical foundations drawn from HCI and organizational science literature with regard to team dynamics. Until recently, these dynamics have remained elusive due to their complexity and lack of quantitative measures. However, wearable electronic devices have made collecting detailed information on team communication affordable. Research indicates that there is predictive power in social signals collected with these devices (Parker et al., 2018). This is particularly applicable to collocated collaboration settings because face-to-face teamwork remains the dominant mode for solving complex problems despite the increasing amount of work done by virtual teams.

Furthermore, collocated collaboration provides unique benefits that are not easy to achieve in digitally mediated forms of teamwork (Olson et al., 2002), such as increasing creativity (Gloor et al., 2014) and performance (Olguın et al., 2009). While preliminary work has demonstrated the feasibility and utility of leveraging multimodal signals to predict behavioral patterns during collaboration activities, more research is needed to understand which data sources best predict certain activities. A detailed overview of the current state of the literature on collocated collaboration modeling using multimodal interaction modeling can be found in Praharaj's (Praharaj et al., 2021) comprehensive literature analysis.

4.3 Method

4.3.1 Data Set

We aim to investigate team-level performance within a collaborative product development task using a data set published earlier (Kohl et al., 2020). We chose this data set for several reasons. First, it was collected during a design thinking process and provides information on the participants' mode on a given day. Second, the data set contains team-level performance assessed within the context of a real company. This will help us better understand the interplay of social signals and modes with regard to team performance. Lastly, prior analysis of this data set revealed a "sociometric DNA" within the modes expressed as structural differences in the patterns of the individual modes(Kohl et al., 2020). We believe that understanding how these differences link to performance is the next step toward understanding and improving social signals in team communication.

The data set contains social signals collected from a group of young professionals engaging in a new product development (NPD) sprint exercise at a large consultancy. In total, 18 participants split into four groups of 4–5 members each were observed over 13 days. Team members were unfamiliar with each other before the start of the project. All teams worked without supervision, so each team structured their workdays and scheduled team meetings as necessary. The teams had no formally defined hierarchy, and no formal roles were established. Data on the participants were collected using wearable electronic devices during working hours, excluding lunch breaks. The wearable electronic device used in this data collection is a sociometric badge that automatically measures social signals derived from speech, body motion, and relative location.

Sociometric badges are well established in the literature, having been successfully deployed in a variety of organizational contexts with a variety of predictor variables. The badges have been used to predict organizationally relevant outcomes such as job attitude and performance (Olguın, 2011), job satisfaction (Olguın et al., 2008), network cohesion (Wu et al., 2008), creativity (Tripathi & Burleson, 2012), group performance (Olguın et al., 2009), and group collaboration (Kim et al., 2008).

As not all group members were present on all days, only 222 instances of sociometric signals are available out of 234 potential instances. The features used in this study are derived from these speech and proximity signals as provided by the Sociometric Solutions software (Solutions, 2014). All participants were asked to complete daily questionnaires to provide information on the design thinking activities after finishing all work-related activities. We have only included the answers assessing how much time (in percentages) participants spent each day on the three different modes. The response rate to this question was 100%, providing 222 responses. Eight senior consultants assessed the team performance of each group at the location site.

4.3.2 Variables

Data from several studies suggest that even at small sample sizes, significant patterns can be measured by social signals extracted from the sociometric badges. These signals, hereinafter referred to as features, include turn taking, activity levels, and proximity within the network (Parker et al., 2018). Prior work (Kohl et al., 2020) focuses on features that capture the rotation of nonverbal behavior in teams. That research assessed the predictability of design thinking mode using five different rotation features: rotating leadership (RL), rotating contributions (RC), turn taking (TTK), successful interruptions (SI), and unsuccessful interruptions (UI).

We used variance inflaction factor (VIF) values to assess if multicollinearity is present to avoid inflated regression coefficients. Due to the multicollinearity of the interruption features (SI & UI) with each other and with turn taking, the interruption features will be excluded from the analysis. We made this decision as turn taking is a well-established feature in the literature (Parker et al., 2018). Only the descriptions of RL, RC, and TTK will be covered in the following sections. Furthermore, the dependent variable team performance (Overall Performance) and the moderation variable design thinking mode (MODE) will be explained.

Rotating Leadership. Rotating leadership (RL) in this data set is calculated from the proxemic features measured by the sociometric badges via infrared and Bluetooth signals. It represents the physical location of each participant in relation to the others, and it changes over time. Teams showing high values in RL go through many changes in the number of team members in proximity to each other throughout the day. The groups in this study often split up and worked individually before coming together again to jointly work on the project. The resulting data form a social network evolving through the data collection period.

RL reveals changing network structures where people oscillate between peripheral and more central positions. In social network analysis terms, RL is a measure of the frequency in which people change their betweenness centrality in the team when represented as a graph. Betweenness centrality is calculated by dividing the times a node in a network is located on the shortest path by the total number of paths. RL represents the changes within the centrality by counting the number of local maxima and minima over time for a person (Kohl et al., 2020; Kohl et al., 2022a).

This is described in the equation below, where the superscript BC indicates that the local minima and maxima are for the betweenness centrality curve, and i indicates the person. Higher numbers indicate more rotation of leadership.

$$RL_i = \#localMinima_i^{BC} + \#localMaxima_i^{BC} \tag{4.1}$$

Research has shown that this measure has a relevant impact on knowledge-sharing dynamics, affects individual and group creativity, and can be a predictor of innovative performance (Alberti et al., 2021; Gloor et al., 2014). RL is a day-level variable.

Rotating Contribution. Like RL, rotating contribution (RC) is a consistent indicator of creativity, a key component of CCPS. For both signals, more rotation is positively related to performance during creative tasks, while less rotation is preferable for non-creative tasks (Gloor et al., 2014). In contrast to RL, which is a proxemic measure of rotation, RC measures paralinguistic rotation. Specifically, it measures the oscillation of the contribution index (CI, or $speaking - listening / speaking + listening$) by counting the number of local maxima and minima in the CI curve of a person over time. It thus represents how frequently people change the amount of time they spend listening vs. speaking (Parker et al., 2018). RC is a day-level variable.

Turn Taking. TTK in groups has been associated with collective intelligence and is, in general, considered a measure of the involvement of all team members, which is crucial for team success (Chuy et al., 2011; Woolley et al., 2010). Lower numbers of turns taken and increased mean speaking segment length are correlated to group effects such as diminished perceptions of individual and group creativity, as well as lower levels of involvement (Parker et al., 2018). Within this data set, turns are speaking segments that occur after and within 10 seconds of another speaking segment. Two speaking segments do not need to be from different people to count as a turn; one person can "self-turn" by pausing and then starting to speak again. This would count as two turns. TTK is a day-level variable.

Team Performance. Overall performance was collected via a survey from senior consultants. The consultants were asked to rate the overall quality of each team's performance at the end of every workweek. The performance measure was assessed on a Likert scale ranging from 1 to 10, with 5 as a neutral point. The team performance measure was measured on a weekly basis and is thus a week-level variable.

Design Thinking Mode. MODE is defined by the percentages reported by the participants. Each observation corresponds to a participant's answer on a given day and is a day-level variable. Each observation was assigned one of four classes: NF, ID, PR or MIXED, depending on whether the participant indicated that at least 60% of the day was spent in the corresponding mode. This cut-off point was chosen to ensure single labeling. For example, an observation indicating 40% NF and 60% ID was assigned to the class ID. The MIXED class was assigned if no mode was present for more than 60% on a given day. MODE is a day-level variable.

4.4 Analysis and Results

Multilevel modeling was applied because measurements in the data are not independent. All variables were either: (1) measured per participant (RL, RC, TTK and MODE); or (2) measured per team (overall performance). The participant variables were measured on a daily basis, and the team variable is measured on a weekly basis. To account for our data structure, we conducted mixed effect modeling where individuals were nested in teams, and time was included as a random slope.

In the first step, we tested the assumptions and prepared the data. Data inspection revealed that the data is normally distributed, and homogeneity of variance can be assumed. The features were standardized because they were measured at different scales. With the aim of this analysis in mind, we removed all data points assigned the class *MIXED* in MODE, reducing the number of data points from 222 to 156.

Table 4.1: Log Likelihood Test

Model	AIC	LogL	Chisq	Df	Pr(>Chisq)
Unconditional Means	381.36	−186.68			
Random Slope	215.43	−99.71	173.94	4	0.0000
Random Slope	215.43	−99.71			
Perf - RL*MODE	203.71	−88.86	21.71	5	0.0006
Perf - RC*MODE	212.81	−93.41	12.62	5	0.0273
Perf - TTK*MODE	210.55	−92.27	14.88	5	0.0109

Note: Akaike information criterion (AIC); Log likelihood values (LogL) χ^2 *(Chisq); Degrees of freedom (df)*

4.4.1 Overall Performance

The guiding research question for this analysis is how different features affect overall performance and how this relationship differs for different modes. To evaluate the moderating effect of MODE, effect coding was used on the remaining three levels of MODE, creating two regressors. We used the class *PR* in MODE as a contrasting group, so observations in this class were assigned a -1. Thus, we created one regressor named *NF-PR* for modes of the NF class assigned a 1 and another named *ID-PR* for observations of the ID class assigned a 0.

In line with the recommended procedure for multilevel analysis, we developed a sequence of models from the simplest to the most complex. We compared model fit using AIC, which penalizes for model complexity. If the AIC is lower for a more complex model, then the gain in fit is worth the extra complexity. In addition, we also conducted likelihood ratio tests for all presented models to assess if the model fit improvement is statistically significant. All models use overall performance as the dependent variable.

We start by building the most basic model, an unconditional means model. This model has no predictors and allows the intercept to vary for participants nested in teams. This model lets participants have their own baseline values but assumes that participants respond to time in exactly the same way. The next step in model fitting is to build a random intercept and slope model. In addition to allowing the intercept to vary for participants nested in teams, the random slope model allows each group line to have a different slope, hence allowing the explanatory variable to have a different effect for each group over time. We conduct a likelihood ratio test using maximum likelihood estimates to see if allowing the slope to vary per week improves our model fit.

The results in Table 4.1 show a significant improvement from the unconditional mean model to the random slope model, increasing χ^2 by $173.94, (df = 4, p < 2.2e - 16)$. The improved fit is also indicated by the AIC score, which is lower for the random slope model. As this model not only shows a better model fit but also makes sense for our data structure, we will be using a random slope model structure when adding our level-one predictors. Table 4.1 shows the results of the likelihood ratio test. Six different models were generated, but only models with significant coefficients for the predictors are presented in the tables. Adding the predictors by themselves did not yield significant coefficients. However, all moderation models showed significant coefficients. Evaluating the model fit between all models, we can observe that Model *Perf - RL*MODE* (Table 4.1) outperforms all models with an AIC score of 203.71. Adding the predictors MODE and RL to this model provides the largest χ^2 improvements. Table 4.2 shows the results for the random slope models with the interaction effects between the factors and MODE.

Looking at the models, we can observe a negative direct effect of the regressor NF-PR, which indicates that the NF class of MODE contributes negatively to overall performance compared to the contrasting group PR, which was coded with -1. There seems to be no significant effect of the class ID on performance compared to PR as measured by the regressor ID-PR. Looking at the interaction effect specifically, we can observe that the relationship between RL and performance is significantly moderated by MODE.

The expected change in the overall performance for a one-unit increase in RL when a team is in the NF class of MODE is 0.067 higher than when a team is in the PR class. The expected change in the overall performance for a one-unit increase in RL when a team is in the ID class of MODE is -0.04 lower than a team in the PR class (Table 4.2 – Model *RL*MODE*). This is to say, for NL, an increase in RL is related to an increase in overall performance. This effect is not as strong (PR) or even not detected (ID) for the other modes. These relationships are visualized in the interaction plot Figure 4.1. We do not observe a significant moderation effect for RC or TTK (Table 4.2 – Model *RC*MODE* & *TTK*MODE*)

Table 4.2: Mixed Effect Models for Overall Performance

	Overall Performance		
	RL*MODE	RC*MODE	TTK*MODE
RL	0.049**		
	(0.017)		
RC		−0.001	
		(0.008)	
TTK			0.00004
			(0.00003)
NF-PR	−0.127*	−0.176**	−0.183**
	(0.059)	(0.058)	(0.059)
ID-PR	−0.031	0.018	0.011
	(0.053)	(0.054)	(0.055)
RL:NF-PR	0.067*		
	(0.027)		
RL:ID-PR	−0.040*		
	(0.018)		
RC:NF-PR		−0.004	
		(0.013)	
RC:ID-PR		0.010	
		(0.010)	
TTK:NF-PR			−0.00001
			(0.00003)
TTK:ID-PR			0.00002
			(0.00003)
Constant	7.058***	7.019***	7.027***
	(0.101)	(0.107)	(0.114)
Observations	156	156	156
Log Likelihood	−104.595	−110.350	−126.408
Akaike Inf. Crit.	235.190	246.701	278.816
Bayesian Inf. Crit.	274.839	286.349	318.464
Note:	*p<0.05; **p<0.01; ***p<0.001		

4.5 Discussion and conclusion

Following our analysis, we can answer our research question. Our results show differences in the effect of social signals on team performance for different participant modes. Different features (specifically RL) affect overall team performance, and this relationship differs over different design thinking modes. While all modes benefit from increased RL, NF does so significantly more than ID and PR. More research is necessary to understand why we observe this relationship. A likely explanation is that while high RL values are associated with positive performance in creative tasks, the directionality of the correlation changes for non-creative activities. The wide spread of variance within the ideation mode and the overall lower impact on the performance mode might indicate that these modes include more non-creative activities as they occur later in the design thinking process.

The key contribution of our paper is the investigation into the relationship between proxemic and paralinguistic features and performance for different modes, which until now has not been explored to a large extent. By analyzing a data set of fairly limited rotational features, this research demonstrates that communication patterns have correlations with team performance and that these correlations differ across different modes. These findings can guide future research designs aimed at understanding the relationship between communication and team performance.

Further, we are adding to the limited knowledge pool of understanding the relationship between social signals and team performance. This contribution to the existing knowledge is vital as it can help to analyze team interaction dynamics during CCPS. Most work on team performance within the social signal processing community has focused on single-mode analysis. The impact of mode on performance has been underexplored. We demonstrate the usefulness and necessity of differentiating between modes when understanding proxemic and paralinguistic features in relation to task performance by utilizing those features captured in a real-world work environment.

Future work will look at different features and how to combine them to yield further improvements. In addition, more modes should be explored to understand how paralinguistic and proxemic features relate to different types of CCPS.

Furthermore, the proposed relationship should also be evaluated in a more controlled setting to examine the effects with less noise. Such studies should separately evaluate the modes without the interaction between them that is observed in the current study. Such examinations could also include a more focused performance evaluation that evaluates the performance of the individual mode instead of overall performance. Finally, these findings pave the way for the exploration of automated coaching of teams by creating interventions to guide teams to interact in ways that should result in optimal team performance for their current mode.

5
Using Speech Contribution Visualization to Improve Team Performance of Divergent Thinking Tasks

Kohl, S., Calero Valdez, A., & Schröder, K. (2023). Using speech contribution visualization to improve team performance of divergent thinking tasks, In *Proceedings of the 15th conference on creativity and cognition*, Virtual Event, USA, Association for Computing Machinery. https://doi.org/10.1145/3591196.3596824.

Abstract

The COVID-19 pandemic has led to a surge in video-conferencing software usage in businesses, universities, and schools. However, research indicates that video-conferencing can hinder creative tasks that require social presence, immediacy, and equal contribution. To address these concerns, this paper advocates the use of analytical and feedback tools to minimize the negative effects of video-conferencing. Additionally, the paper examines how computer-based creative thinking training can improve divergent thinking abilities in teams. An exploratory study was conducted using a real-time visual support system that encouraged equal participation by visualizing contribution levels. The results demonstrated that the visual support system had a positive impact on creativity performance in teams during video-conferencing meetings. These findings add to the literature on digital technology's ability to support team creativity.

5.1 Introduction

As the COVID-19 pandemic has forced organizations to shift towards remote working, there has been a significant need for teams to explore options for how to collaborate remotely. However, teams working on creative tasks have shown to suffer the most, as solving creative tasks benefits from social presence, immediacy, and equal contribution. Video-conferencing has been shown to negatively affect idea generation (Brucks & Levav, 2022), which underscores the importance of developing effective communication strategies for remote creative teams.

One potential solution is the use of computer-based communication training, which can enhance divergent thinking abilities. An exploratory study was conducted to investigate how a real-time visual support system impacts the performance of teams working on a divergent thinking task during a video-conferencing meeting. The visual support system was designed to encourage equal participation by visualizing participation contribution. The study aimed to assess if a visual support system not only subjectively reduces over-participation but also has a measurable impact on team performance

This research contributes to a deeper understanding of how communication links to creative performance in teams and extends the literature on computer-based creative thinking training tools. By leveraging visual support systems, remote creative teams can potentially improve their communication and collaboration, which could help mitigate the negative effects of remote work on creativity.

5.2 Literature Review

5.2.1 The Impact of Social Signals on Meeting Success

Previous studies on the effects of social signals on team outcomes are inconsistent, identifying various signals as potentially relevant (Praharaj et al., 2021). In line with this, balanced and equal participation, as measured in speaking time vs listening time, has been shown to improve team performance (Dong et al., 2012) while simultaneously not always being optimal, as different forms of collaboration show different communication patterns (Jayagopi et al., 2010; Kohl et al., 2020; Kohl et al., 2022a). This strongly suggests that different types of communication benefit from different patterns. A study by Kohl et al. (2022b) supports this conclusion, demonstrating that different stages of design thinking benefit from different communication patterns. While not all stages benefit from equal participation, they conclude that participation from all members is beneficial in a collaborative discussion focusing on divergent thinking. These results are in line with earlier work showing that for creative tasks, more rotation of contribution from team members is positively related to performance (Gloor et al., 2014).

Next, to improve performance, balanced and equal participation ensures, at a minimum a sense of 'being heard' and rapport building, ultimately improving team satisfaction (Lawford, 2003). These findings for in-person communication align with research investigating online meeting effectiveness, which correlates with meeting inclusiveness, participation, and comfort in contributing (Cutler et al., 2021). Establishing and maintaining a workplace culture in which everyone feels free to contribute can hold large financial benefits for companies. Technological solutions have a large potential for assisting with this, helping attendees to understand their own and others' meeting dynamics, and improving meeting effectiveness (Samrose et al., 2021).

5.2.2 Improving Creativity Through Real-Time Feedback Systems

Among the strategies of teaching creativity, technology is regarded as fundamental (Tang et al., 2022) and computer-based creative thinking training has been shown to enhance divergent thinking abilities as successfully as traditional training (Benedek et al., 2006). Despite this, effectively integrating technological tools to enhance creativity is challenging (Bereczki & Kárpáti, 2021; Tang et al., 2022). Video conferencing provides a unique opportunity in this context as it enables the use of real-time feedback which is considered to be one of the most critical support strategies for optimal learning and skill development (Shute, 2008; Wisniewski et al., 2020). Up to now, far too little attention has been paid to this opportunity.

While some commercial video conferencing tools provide real-time features to make up for their constraints—such as clapping—to date, no major platform utilizes social signals to enable users with actionable meeting metrics. Studies reveal the need for feedback assistance in video-conferencing meetings to make them more effective and inclusive (Samrose et al., 2021). Research investigating the effect of feedback systems on meeting dynamics and discussion outcomes for video-conferencing meetings have included, among others, differences in timing comparing real-time (Calacci et al., 2016; DiMicco et al., 2004; Faucett et al., 2017; Leshed et al., 2009) with post-meeting (Samrose et al., 2021; Samrose et al., 2018) feedback.

These works lay the foundation for designing effective feedback design for group communication. They highlight the need for any tool aiming at equally distributed communication to focus on reducing the contribution of the over-participator (DiMicco et al., 2004). Prior studies have shown that the system can help reduce subjective over-participation (Schröder & Kohl, 2022). However, too much information in the feedback can be cognitively taxing (Samrose et al., 2021), diminishing the potential benefits of any system. Therefore, selecting the right signals to base the feedback on is a crucial part of the process.

5.3 Methodology

We conducted an experiment with a mixed design of the between factor *exposure to the stimuli* and a within factor as *type of a divergent think-*

ing task. We selected the context of divergent thinking tasks for the experiment as the literature suggests that these types of tasks benefit from equal participation. Divergent thinking tasks are widely used in psychometric and experimental studies of creativity, and address the potential for creative thinking and problem solving (Silvia et al., 2008). In the following, we elaborate on the procedure, tasks, and scoring.

5.3.1 Participants

We recruited 150 undergraduates to take part in a 30-minute study in exchange for course credit. All undergraduates that participated are part of a business degree with a high focus on collaborative, problem-based learning. Participants are all part of the same degree and have limited prior rapport with each other as all classes prior were conducted online due to social distancing guidelines.

Participants were randomly assigned to teams of four. Incomplete teams were excluded from the study. Two teams were excluded from the analysis as they experienced technical difficulties, leaving 72 participants (25.9% male; M

age=18.7$years$,SD $= 1.21$) in 18 teams (N=18). Each team was assigned to a mixed design group which was randomized on task order as well as stimuli. This study measured stimuli as a between-subject factor and task as a within-subject factor, resulting in 4 different group types (Figure 5.1). We counterbalanced task order and stimuli to exclude the possibility of any sequencing effects through a Graeco-Latin square design.

5.3.2 Visual Support System

In video-conferencing, understanding individual speech contributions within a specific task is challenging as the attention is primarily drawn to the task itself, namely the verbal interaction through speech as well as visual feedback via facial expressions. Therefore a visual support

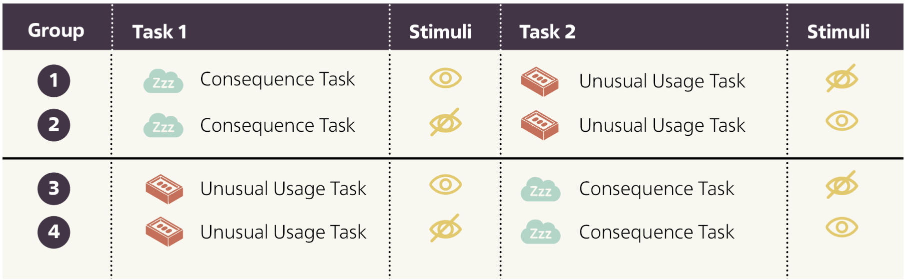

Figure 5.1: Study Design—Group overview showing which groups were exposed to the stimuli during which part of the study. I.e group 1 saw the stimuli during the first task but not during the second.

system needs to be very intuitive so that users can improve collaboration based on real-time feedback without being distracted by it or by the burden of interpreting it.

In general, research has shown that subjective perception is challenging in the given scenario (Fiorella et al., 2012; Kalyuga et al., 1999), but can be enhanced through the use of a complementary, in this case, visual, channel (O'Neil et al., 2010). The two primary attributes for understanding the real-time data in context are the categorical relationship (who is talking) and the quantified individual speech contribution (the quantity of individual speech contribution).

We use a visual-support tool to assess the impact of real-time feedback in video conferencing meetings on team performance (see Fig. 5.2). According to the visualization literature, spatial encoding is the most efficient way to encode the categorical meaning (Munzner, 2014), but it remains unclear which type of marks are effective in the given situation to encode multiple dynamic quantitative data types.

Visualization research focuses on the visual understanding of data as a primary task, while our visual support tool tries to use visualization in an assistive scenario. While past researchers used area encoding and utilized orbs for speech representation (Calacci et al., 2016; Fiorella et al., 2012; Kim et al., 2012), more recent research indicates that the

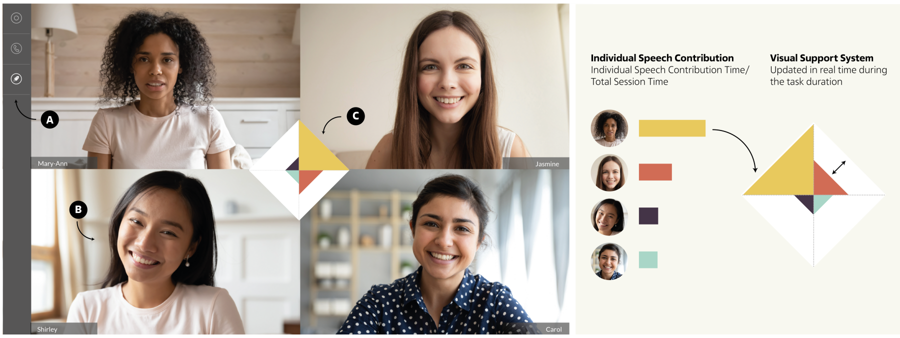

Figure 5.2: Example of the Video conferencing application in use with the visual support system activated: A) shows the general interface with three options: to record the session, to end the session as well as activate or deactivate the visual support system. B) The individual video streams and C) the visual support interface that show the individual speech contribution in real time encoded in the triangle size.

most efficient visual encoding for the given task is centred triangles (Schröder & Kohl, 2022).

Therefore, the visual-support tool used in this study utilizes individual coloured triangles to encode the user's participation visually. The visualisation updates in real-time, based on the individual amounts of speech contribution via the microphone for the current duration of the task. The four triangles are centred between the users' video streams to facilitate the real-time understanding of the user's contribution directly next to the individual video streams.

The resulting tool specifically addresses the primary design challenge for complementary encoding to generate an intuitive feedback system. The current version of the system only utilizes a camera and microphone but no text chat. The experimental work presented here provides one of the first investigations into how teams perform with and without using a visual-support system.

5.3.3 Procedure

Participants signed a consent form at the testing site and were instructed to be creative, which is recommended in the literature on divergent thinking. This helps increase creativity variance and eliminate ambiguity in results (Runco et al., 2005). Each team member was taken to a separate room and given an introduction to the video-conferencing application. After confirming that they understood how to use it, participants received written instructions for the task. The researchers ensured that everyone understood the task before starting. Teams had 3 minutes to complete each task before moving on to the next one.

5.3.4 Divergent Thinking Task

For this study, the participating teams were instructed to solve two divergent thinking tasks: an Unusual Uses Task (UUT), a classic and widely used measure of divergent thinking (Guilford, 1967), and a Consequences Task (CT) (Silvia et al., 2008). Teams generated alternative usages for a brick in the Unusual Uses Task (UUT) and creative consequences for a hypothetical scenario in the Consequences Task (CT). Both task descriptions emphasized the importance of generating creative ideas.

5.3.5 Scoring the Responses

There are various methods for scoring divergent thinking tasks, with uniqueness scoring being the most standard. However, it has been criticized for being an unreliable scoring method as it can confound creativity with fluency, providing participants with high creativity scores simply by virtue of generating a large number of responses. To overcome this issue, we followed the subjective rating approach proposed by Silvia et al. (2008).

Audio recordings of the tasks were digitally transcribed by the researchers blind to conditions. The researchers grouped the digitized responses under keywords for similar answers. The keywords grouped ideas that are closely related but worded differently, e.g., "smash something" and "break something" as usages for a brick, were included as one keyword, "to break something". The teams in the study generated 324 keywords in total, with 93 and 44 unique keywords for UUT and CT, respectively. All keywords, including the duplicates, were sorted alphabetically within each task, ensuring that raters were blind to ordering effects.

Two raters per task were asked to score each keyword on a 1 (not at all creative) to 5 (highly creative) scale. The scoring framework used was *'originality'* composed of the dimensions *uncommon*, *remote*, and *clever*. Raters were instructed to consider all three dimensions, and that strength in one facet can balance weakness in another facet (Silvia et al., 2008; Wilson et al., 1953). Each keyword was assigned an originality score based on the reviewers' average ratings. We calculated two creativity indexes for the analysis:

Average creativity

The first index is the average rating of all keywords per team per task. The ratings per keyword were first summed and then divided by the number of keywords. This index preferences quality of a keyword over quantity, as a team with two creative keywords, will have a higher average than a team with one creative keyword and four uncreative ones.

Overall creativity

The second index used was overall creativity. For this index, the sum of all unique keywords per team per task was calculated. Keywords were included only once, meaning that a team listing “smash something” and “break something”—which were each listed as the keyword “to break something”—will be awarded points only once for this keyword.

5.4 Findings

The results of two creativity indexes are shown per task in table 5.1. For both tasks, teams on average had slightly more unique keywords for tasks with the visual support stimuli present with 1.08 and 1.42 keywords on average more. Differentiating within a task, between teams presented with the stimuli and teams without, shows that, for both tasks, independent of the task order, teams perform better when presented with the stimuli. This holds true for the overall index (yes = 27.3, no = 24.6) as well as the average index (yes = 2.17, no = 2.10). When accounting for task order, most teams show a general increase in creativity over time with the second task presented to them having higher creativity scores in both indexes (except for CT–no).

This result might indicate the effect of rapport building for the teams working together. This effect is strongest when presented with the stimuli on task two with overall scores of 36.5 (UUT) and 19 (CT) and average scores of 2.53 (UUT) and 2.13 (CT). Overall, teams seem not to perform well when presented with the stimuli on task 1, with some of the lowest creativity scores in the overall index (16 for UUT and 21.7 for CT) and average index (1.95 for UUT).

The results suggest that presenting unknown stimuli to a team working together for the first time could be too distracting. Although creativity scores only slightly improved when stimuli were presented, there was a significant difference in creativity indexes when evaluating the second task performed. At this point, all teams had prior work experience together and had built rapport and familiarity with the platform.

Table 5.1: Results of the Divergent Thinking Tasks

Task	Stimuli	N	Overall	Average	SD	#Key	Overall		Average	
							Task 1	Task 2	Task 1	Task 2
UUT	no	10	24.60	2.10	0.52	11.3	24.10	25.10	2.04	2.15
UUT	yes	8	**27.31**	**2.17**	0.41	12.38	21.70	**36.50**	1.95	**2.53**
CT	no	8	13.50	1.89	0.33	6.88	19.00	10.20	2.11	1.76
CT	yes	10	**17.50**	**2.09**	0.23	8.3	16.00	**19.00**	2.05	**2.13**

Note: Overall creativity index (Overall); Average creativity index (Average)
Number of unique keywords generated (#Key)

On the Overall index, teams performed 11.4 points better in the UTT and 8.8 points better in the CT. These results are promising and suggest that more established teams will benefit from using the system. However, further research is needed to rule out the possibility of sequencing effects, such as the impact of presenting stimuli in the first task on the results of working without stimuli in the second task.

These results are a promising indication that more established teams will benefit from using our system. However, more research is necessary to exclude sequencing effects.

5.5 Conclusion

Our study measured creativity in a computer-mediated communication setting using established tests, comparing the performance of groups with and without our visualization system. We found that the system improved creativity-related tasks by making participants aware of their speech contributions in meetings. This led to an increase in the quantity and quality of creative answers, especially for the Unusual Uses Task. While there was a slight learning curve, the system ultimately helped teams be more creative.

In light of the current situation with remote work and study, such a system can be helpful for teams that rely on video-mediated communication. Not everyone may feel comfortable contributing in digital meetings, and our system can address these inequalities by providing feedback. Additionally, the system can ensure that every participant's voice is heard, thereby improving team creativity. Overall, our findings provide a systematic evaluation of the impact of a live-feedback visualization system on team performance in divergent thinking tasks and suggest that such a system can contribute to improved computer-mediated communication for creativity.

5.6 Limitations

Although our study highlights the potential of a visual support system to enhance creativity, we did not measure if teams had more equal speaking times under the with-stimuli condition. Experimenters noted a qualitative impression of more equal speaking times, but we lack quantitative data. Moreover, there might be other factors that explain our results, which we did not measure, such as baseline creativity differences among groups. However, our sample size of 72 participants and random assignment make this unlikely. Team composition could also play a role, and it is possible that teams were not equally composed in terms of complementarity and rapport. Future studies could measure these factors to provide a better understanding of the effects of a visual support system on team creativity. Measuring team rapport and perceived complementarity in group composition could at least point towards real differences and help compare perceived differences.

5.7 Future Work

During this study, teams interacted with the visual support stimuli only once. There could be learning, habituation, and saturation effects over time that impact the long-term effects of the system. Future studies should assess if frequent and repeated usage can provide better insights into the impact of our system on meeting dynamics.

Generally, levels of creativity per keyword were low in this study, which might indicate a lack of engagement within the sample population related to the given task. Developing a task that is better suited to measuring divergent thinking for this specific context might provide a more realistic assessment of the potential benefits of visual support systems and help raise creativity levels to avoid any confounding measures between low levels of creativity in the sample and low levels of engagement. Another way to accomplish this is to choose a setting with higher intrinsic motivation for the participants.

Moreover, we will utilize the data collected on individual contributions to a meeting to measure the impact of the visualization on the equality of contributions. At the moment, this was not measured systematically. To validate the finding that our system improves the equality of contributions, we will compare the data with and without the visualization in a future study.

To reduce the impact of the Hawthorne effect (Jones, 1992) – participants not behaving naturally because of being in an experimental setting – we plan to utilize the system in remote teaching settings where students work regularly with the system. This should reduce the impact of the artificiality of the situation.

Lastly, we want to test the impact of different instructions. Participants were not made aware of the idea that equal contributions of team members should yield more creative results. In follow-up studies, we want to study, if explicit instruction on this phenomenon will impact the equality of contribution even more.

6
Conclusion

> *The most important thing in communication is hearing what isn't said.*
>
> — Peter F. Drucker

6.1 Summary of Findings

The compilation of chapters in this dissertation advances our understanding of how social signals can be utilized to better understand the relationship between team communication and team performance and explores how researchers and designers can benefit from these insights when designing supporting tools for teams.

The first project *(Chapters 2, 3, & 4)* in this thesis explores the sociometric DNA of interactions and the relationship that these unique social signal patterns have with team performance. The project starts by presenting evidence *(Chapter 2)* of the existence of a sociometric DNA in design thinking meetings. The presented study found differences between design thinking modes, which are reflected in the different social signal patterns measured using sociometric badges. *Chapter 3* investigates how the idea of the sociometric DNA of a problem-solving activity can be used for context detection. The chapter provides evidence to prove that different CPS stages can be predicted beyond chance in a multiclass prediction task. *Chapter 4*, this project's final stage, explores the relationship between social signals and team performance for different design thinking modes. It provides evidence that social signals are related to team performance and this relationship varies between modes. It concludes that certain social signals contribute more positively to one mode but less positively to others. This project's findings not only demonstrate that social signals can be used for context detection but also highlight the importance of building context-aware systems to improve team performance.

The second project *(Chapter 5)* builds on the results of the previous studies. It represents a first step toward understanding the impact of context-aware systems on team performance. Specifically, this study tests the impact of a custom-built, real-time visual support system on the performance of teams working on a particular problem-solving mode, producing findings that show that the system positively affects team performance.

This project demonstrates the potential impact of implementing context-specific interventions on team performance. Although more exploration is necessary to determine which interventions to use in other problem-solving modes, the results provide the groundwork for future research.

6.2 Connecting Domains

As the introduction to this thesis articulates, a key contribution of this research is that it builds upon the foundations of somewhat disconnected research domains to develop connections between them. The following discussion highlights how merging several literature streams contributes to the investigation of connections between team communication and team performance by examining how my work integrates the targeted disciplines and thereby generates novel insights. This involves presenting the findings and contributions of each chapter with regard to the fields of inquiry that they connect (as shown in Figure 6.1) before discussing the practical relevance of my work. I conclude with suggestions of potential avenues for future research before providing my final thoughts.

6.2.1 Connecting SNA and SSP

In Chapter 2, my co-authors and I introduce the idea of the sociometric DNA of interactions. We generate multimodal features of interactions—based on the use of network and graph theory measures—as inputs for modeling approaches to analyze the higher-level behavioral dimensions of collaborative behavior patterns.

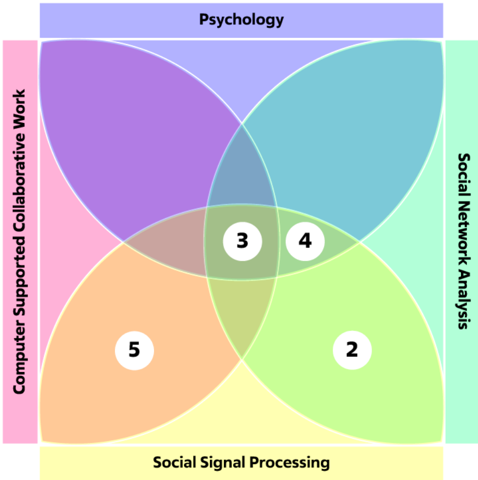

Figure 6.1: Connections between the research domains and chapters

By utilizing SNA metrics to generate Rotating Leadership and Rotating Contribution features, we reveal previously unidentified distinguishable collaborative behavior patterns for problem-solving activities (in this case, Need Finding, Ideation, and Prototyping). The results of the analysis show not only that rotation features can be used to infer problem-solving activities but also that the importance of rotating features differs between activities. The direction of certain features changes for different activities.

Although SNA has generated tremendous insights into group dynamics, including the seminal work by Burt (2004) on structural holes and creative idea generation, SNA metrics are seldom applied during SSP, which often relies on low-level features. This work demonstrates how these two domains can synergize to advance the analysis of team behavior.

6.2.2 Connecting SNA, SSP, Psychology, and CSCW

Chapter 3 expands on the idea of a unique sociometric DNA of interactions by providing evidence that individual problem-solving activities can be differentiated on the basis of observed multimodal features. The analysis results show that the models can reproduce the self-reported labeling of the recordings on previously unseen examples at a level well above chance. This chapter provides a theoretical explanation for the observed differences, linking the signals to the cognitive stages of divergent and convergent thinking. Furthermore, this chapter connects this line of research to the CSCW domain, highlighting the importance of context detection for automated support tools and explaining how the insights generated by the study could be integrated into automated support systems.

This chapter also integrates insights on team cognition from psychology into SSP, overcoming the limitations of earlier studies, which have often struggled to identify significant social signals when combining various problem-solving activities as outlined in my introduction. Finally, this chapter outlines how researchers and practitioners working within CSCW to develop automated support systems can use these new insights.

6.2.3 Connecting SNA, SSP, and Psychology

Chapter 4 sets out to connect the sociometric DNA of individual problem-solving activities with team performance. This study adds to the limited knowledge surrounding the relationship between social signals and team performance. Most work on team performance in the SSP community has either focused on analyzing only one problem-solving activity or else failed to differentiate in cases where different activities are included in a single study. However, this study's results demonstrate differences in the impact of social signals on team performance for different problem-solving activities. Different features affect overall team performance, and this relationship differs for Need Finding, Ideation, and Prototyping.

This contribution to the existing body of knowledge provides critical support to the analysis of team interaction dynamics. By utilizing features captured in a real-world work environment, we demonstrate the usefulness and necessity of differentiating between problem-solving activities when understanding social signals in relation to task performance. This enables my co-authors and I to provide quantitative evidence with high-resolution, quantitative, time-series data to support the insights of psychologists working on the connection between team cognition and team communication. Much of the available literature in this area has produced only theoretical assessments or qualitative evaluations.

6.2.4 Connecting SSP and CSCW

Chapter 5 explores the impact on team performance of using a visual support system—developed based on the insights generated by this thesis—while the team works on divergent-thinking tasks. For this study, teams worked on one of two tasks with a support system that visualizes the speech contributions of individual team members. By comparing the average creativity of answers given for two creativity tasks, my co-authors and I have been able to demonstrate that this type of live-feedback visualization could improve team performance on divergent-thinking tasks.

Although the connection between SSP and CSCW is far more established than connections between the other domains, it remains fairly uncommon to analyze social interactions and provide feedback in real-time, with near–real-time feedback a far more common form of immediate support (Fiorella et al., 2012; Samrose et al., 2021; Samrose et al., 2018; Wang et al., 2021). This is because it is challenging to implement feedback in real-time because it may disrupt the process and distract people from their primary task. The system used in this chapter, developed in collaboration with the Human Data Interaction Lab, represents a feasible mechanism for future researchers in the CSCW and SSP domains, especially given that research on the effectiveness of feedback for skills development shows that real-time feedback is far more effective than near–real-time feedback, which indicates that a deeper understanding of how this form of feedback can be realized is crucial for establishing successful training tools and team interventions.

6.3 Practical Relevance

Whether in-person, hybrid, or remote, meetings will remain relevant in the future. Teams will continue to work together to solve problems, and these interactions will benefit from the support of context-aware systems. The results of this thesis lay the foundation for further exploration of the development of these systems, which can positively impact team performance across a variety of problem-solving modes.

This thesis has significant practical relevance for the design and development of context-aware support systems for teams. Chapters 2 and 3 present clear evidence that context detection is possible, even for noisy data with fairly limited features, and chapter 4 provides evidence that the composition of features is correlated to team performance. The findings demonstrate that social signals, including those captured using sociometric badges, can be used to understand team communication and performance in a non-intrusive way. This importantly allows for the development of support systems that can recognize which cognitive stage a team is in and present relevant and actionable interventions.

One of the key outcomes of this thesis is the demonstration of the potential of visual support systems to use social signals to support teams. The study presented in Chapter 5 provides evidence of the positive impact on team performance of a custom-built, real-time visual support system. This highlights the potential for similar systems to be deployed in various settings to improve the performance of teams working on problem-solving activities.

Using visual support systems that rely on social signals collected during team interactions offers several benefits. First, by relying on paralinguistic and proxemic cues, such as speech contribution, body language, tone of voice, and other social signals, these systems can provide an in-depth understanding of the dynamics of team interactions. This understanding can then be used to design interventions that can improve team performance without violating the privacy and confidentiality of team members.

Second, the visual nature of these support systems makes them accessible and easy to use for teams of all sizes, regardless of their technical proficiency. These support systems can be customized to meet the specific needs of different teams and can provide real-time feedback to help teams improve their performance. Future systems that include these kinds of relevant and actionable interventions could help individuals be more inclusive and productive during meetings, which should ultimately increase their comfort and productivity.

In conclusion, using visual support systems that leverage social signals collected during team interactions has the potential to revolutionize how teams work together. These systems can provide teams with new levels of insight into their interactions, allowing them to work more effectively and efficiently and to achieve better outcomes. Although further exploration into the use of social signals for visual support systems is necessary to fully realize the potential of this approach, the results of this thesis represent a promising first step in this direction.

Additionally, it is important to recognize that although we captured audio of subjects during the project's data collection phase, none of the analysis and results in this thesis depend on the ability to identify the spoken words that constitute the semantic content of the encounter.

Systems based on the technology deployed in this thesis can be utilized in privacy-sensitive settings and provide options for researchers to work within the restrictions of confidentiality agreements in real-world scenarios. For practitioners, this work provides a clear message that it is possible to gather detailed information on team interactions without compromising their privacy. This is particularly relevant in today's increasingly regulated data protection and privacy landscape.

6.4 Future Research and Limitations

The work presented in this thesis provides valuable insights into the use of social signals to understand team communication and performance and design visual support systems for teams. By using new sets of sensing and analysis techniques, this thesis represents the first step toward an attempt to address some of the fundamental issues with team communication analysis and context-aware technology development. In doing so, this research endeavors to investigate modern, unprecedented issues that cannot be managed using conventional methods. However, there remains substantial work to be done in this field to fully realize its potential. This section highlights some of the areas that I find particularly intriguing.

6.4.1 Further exploration of social signals

Although this thesis provides evidence for the existence of what we have called the sociometric DNA of team interactions and for the relationship between social signals and team performance, much remains to be revealed and understood about the various social signals present during team interactions. Further research into the specific social signals that are most predictive of team performance in certain modes will enable a better understanding of this area and lead to the development of more effective visual support systems.

6.4.2 Future research on context-aware systems

The findings of this thesis demonstrate the potential for context-aware systems to improve team performance. However, a primary limitation of this thesis is its use of a fairly limited sample size. It is important that future researchers continuing this investigation expand the number of observed interactions. Context-aware technologies need a larger corpus of labeled interactions to correctly classify interactions. In addition, more features should be tested to assess which features have the greatest predictive power while being the least invasive for users. Future work should focus on developing these systems, including developing new algorithms and methods for collecting, processing, and using social signals to support teams.

6.4.3 Future research on visual support systems

Future research on visual support systems should focus on exploring the long-term impact of the system on team dynamics. The current study only assessed team interactions with the visual support stimuli once; repeated and frequent usage of the system could lead to better insights, with further investigation of learning, habituation, and saturation effects over time necessary to better understand the system's long-term impact. Additionally, exploring the effect of repeated usage of the system would reduce the impact of the Hawthorne effect, which describes participant behavior changing due to the fact that they are being observed. Studying the long-term impact of the system would provide a more accurate assessment of its effectiveness at enhancing team performance.

Additionally, although this Ph.D. project has focused on the importance of providing feedback based on the mode of the team, it was limited by time and budget constraints and was only able to test the visual support system in one mode. Future research should aim to explore the impact of the system in different modes and even consider developing new ways of providing feedback and deploying interventions based on the meeting context. This is especially important because, based on the results of this study, it is hypothesized that the system will need to be adjusted for activities that involve convergent thinking as the primary mode of team cognition.

6.5 Final Thoughts

There is immense potential for interdisciplinary collaboration between the computer science field and other disciplines, and much remains to be discovered about utilizing social signals in team performance. My thesis establishes the foundation for further exploration of how social signals can inform the design and development of context-aware support systems. The integration of psychology and computer science knowledge in this thesis has already produced great results, and I believe that as we continue to push the boundaries of what is possible, we will continue to uncover new and exciting insights.

Moving forward, it will be important to continue to explore the relationship between social signals and team performance and to develop new methods for analyzing and utilizing social signals in the context of team communication. This will require computer science researchers to collaborate with specialists from other disciplines as well as integration of new technologies, tools, and advancements in these other disciplines.

In conclusion, I believe that the future of interdisciplinary collaboration between computer science and other disciplines holds great promise, and I hope that my thesis serves as a small but significant step toward realizing the full potential of this collaboration.

Bibliography

Ajdadilish, B., Kohl, S., & Schröder, K. (2022). Enhancing Evaluation of Room Scale VR Studies to POI Visualizations in Minimaps, In *Eurovis 2022*, The Eurographics Association. https://doi.org/10.2312/evp.20221136

Alberti, F. G., Belfanti, F., & Giusti, J. D. (2021). Knowledge exchange and innovation in clusters: A dynamic social network analysis. *Industry and Innovation*, *28*(7), 880–901.

Allen, J. A., Rogelberg, S. G., & Scott, J. C. (2008). Mind your meetings: Improve your organization's effectiveness one meeting at a time. *Quality Progress*, *41*, 48.

Amabile, T. M. Et al. (1988). A model of creativity and innovation in organizations. *Research in organizational behavior*, *10*(1), 123–167.

Anderson, N., Potočnik, K., & Zhou, J. (2014). Innovation and creativity in organizations: A state-of-the-science review, prospective commentary, and guiding framework. *Journal of management*, *40*(5), 1297–1333.

Bales, R. F., & Strodtbeck, F. L. (1951). Phases in group problem-solving. *The Journal of Abnormal and Social Psychology*, *46*(4), 485.

Barczak, G., Lassk, F., & Mulki, J. (2010). Antecedents of team creativity: An examination of team emotional intelligence, team trust and collaborative culture. *Creativity and innovation management*, *19*(4), 332–345.

Baur, T., Damian, I., Lingenfelser, F., Wagner, J., & André, E. (2013). Nova: Automated analysis of nonverbal signals in social interactions, In *Human behavior understanding: 4th international workshop, hbu 2013, barcelona, spain, october 22, 2013. proceedings 4*. Springer.

Beattie, G., & Ellis, A. (2017). *The psychology of language and communication*. Routledge.

Benedek, M., Fink, A., & Neubauer, A. C. (2006). Enhancement of ideational fluency by means of computer-based training. *Creativity Research Journal*, *18*(3), 317–328.

Bereczki, E. O., & Kárpáti, A. (2021). Technology-enhanced creativity: A multiple case study of digital technology-integration expert teachers' beliefs and practices. *Thinking Skills and Creativity*, *39*, 100791.

Brown, T. Et al. (2008). Design thinking. *Harvard business review*, *86*(6), 84.

Brucks, M. S., & Levav, J. (2022). Virtual communication curbs creative idea generation. *Nature*, *605*(7908), 108–112.

Burt, R. S. (2004). Structural holes and good ideas. *American journal of sociology*, *110*(2), 349–399. https://doi.org/10.1086/421787

Calacci, D., Lederman, O., Shrier, D., & Pentland, A. (2016). Breakout: An open measurement and intervention tool for distributed peer learning groups. *arXiv preprint arXiv:1607.01443*.

Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, *11*(Jul), 2079–2107.

Chandrasegaran, S., Bryan, C., Shidara, H., Chuang, T.-Y., & Ma, K.-L. (2019). Talktraces: Real-time capture and visualization of verbal content in meetings, In *Proceedings of the 2019 chi conference on human factors in computing systems*.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321–357.

Chuy, M., Zhang, J., Resendes, M., Scardamalia, M., & Bereiter, C. (2011). Does contributing to a knowledge building dialogue lead to individual advancement of knowledge, In *Connecting computer-supported collaborative learning to policy and practice: Cscl2011 conference proceedings*. Isls.

Cook, S. D., & Yanow, D. (1993). Culture and organizational learning. *Journal of management inquiry*, *2*(4), 373–390. https://doi.org/10.1177/105649269324010

Cooke, N. J., Duchon, A., Gorman, J. C., Keyton, J., & Miller, A. (2012). Preface to the special section on methods for the analysis of communication. *Human Factors*, *54*(4), 485–488.

Cooke, N. J., Salas, E., Kiekel, P. A., & Bell, B. (2004). Advances in measuring team cognition. *Team cognition: Understanding the factors that drive process and performance.*, 83–106. https://doi.org/10.1037/10690-005

Coutaz, J., Crowley, J. L., Dobson, S., & Garlan, D. (2005). Context is key. *Communications of the ACM*, *48*(3), 49–53.

Cross, R., & Cummings, J. N. (2004). Tie and network correlates of individual performance in knowledge-intensive work. *Academy of management journal*, *47*(6), 928–937.

Cutler, R., Hosseinkashi, Y., Pool, J., Filipi, S., Aichner, R., Tu, Y., & Gehrke, J. (2021). Meeting effectiveness and inclusiveness in remote collaboration. *Proceedings of the ACM on Human-Computer Interaction*, *5*(Cscw1), 1–29.

Dávila-Montero, S., Dana-Lê, J. A., Bente, G., Hall, A. T., & Mason, A. J. (2021). Review and challenges of technologies for real-time human behavior monitoring. *IEEE Transactions on Biomedical Circuits and Systems*, *15*(1), 2–28.

Davis, J. P., & Eisenhardt, K. M. (2011). Rotating leadership and collaborative innovation: Recombination processes in symbiotic relationships. *Administrative Science Quarterly*, *56*(2), 159–201.

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of applied psychology*, *95*(1), 32.

Dietzel, J., Francu, R. E., Lucas, B., & Zaki, M. (2018). Contextually defined postural markers reveal who's in charge: Evidence from small teams collected with wearable sensors.

DiMicco, J. M., Pandolfo, A., & Bender, W. (2004). Influencing group participation with a shared display, In *Proceedings of the 2004 acm conference on computer supported cooperative work*.

D'Mello, S., Stewart, A. E., Amon, M. J., Sun, C., Duran, N., & Shute, V. (2019). Towards dynamic intelligent support for collaborative problem solving, In *Ceur workshop proceedings*. Ceur-ws.

Dong, W., Lepri, B., Kim, T., Pianesi, F., & Pentland, A. S. (2012). Modeling conversational dynamics and performance in a social dilemma task, In *2012 5th international symposium on communications, control and signal processing*. Ieee.

Dorst, K. (2011). The core of 'design thinking'and its application. *Design studies*, *32*(6), 521–532. https://doi.org/https://doi.org/10.1016/j.destud.2011.07.006

Eagle, N., & Pentland, A. (2006). Reality mining: Sensing complex social systems. *Personal and ubiquitous computing*, *10*, 255–268.

Efeoglu, A., Møller, C., Sérié, M., & Boer, H. (2013). Design thinking: Characteristics and promises, In *Proceedings of 14th international cinet conference on business development and co-creation*.

Eloy, L., EB Stewart, A., Jean Amon, M., Reinhardt, C., Michaels, A., Sun, C., Shute, V., Duran, N. D., & D'Mello, S. (2019). Modeling team-level multimodal dynamics during multiparty collaboration, In *2019 international conference on multimodal interaction*.

Elsbach, K. D., & Stigliani, I. (2018). Design thinking and organizational culture: A review and framework for future research. *Journal of Management*, *44*(6), 2274–2306.

Faucett, H. A., Lee, M. L., & Carter, S. (2017). I should listen more: Real-time sensing and feedback of non-verbal communication in video telehealth. *Proceedings of the ACM on Human-Computer Interaction*, *1*(Cscw), 1–19.

Fiore, S. M., Rosen, M. A., Smith-Jentsch, K. A., Salas, E., Letsky, M., & Warner, N. (2010). Toward an understanding of macrocognition in teams: Predicting processes in complex collaborative contexts. *Human Factors*, *52*(2), 203–224.

Fiorella, L., Vogel-Walcutt, J. J., & Schatz, S. (2012). Applying the modality principle to real-time feedback and the acquisition of higher-order cognitive skills. *Educational Technology Research and Development*, *60*(2), 223–238.

Fischer, U., McDonnell, L., & Orasanu, J. (2007). Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation, space, and environmental medicine*, *78*(5), B86–B95.

Francu, R. (2019). *Sending all the right signals: How social signaling drives creativity in b2b new service development* (Doctoral dissertation). Maastricht University. Off Page Amsterdam. https://doi.org/10.26481/dis.20190606rf

Freeman, L. (2004). The development of social network analysis. *A Study in the Sociology of Science*, *1*(687), 159–167.

Gloor, P. A., Almozlino, A., Inbar, O., Lo, W., & Provost, S. (2014). Measuring team creativity through longitudinal social signals. *arXiv preprint arXiv:1407.0440*.

Graesser, A. C., Fiore, S. M., Greiff, S., Andrews-Todd, J., Foltz, P. W., & Hesse, F. W. (2018). Advancing the science of collaborative problem solving. *psychological science in the public interest*, *19*(2), 59–92.

Greenberg, S. (2001). Context as a dynamic construct. *Human–Computer Interaction*, *16*(2-4), 257–268. https://doi.org/10.1207/s15327051hci16234_09

Guilford, J. P. (1950). Creativity. *American Psychologist*, *5*(9), 444–454. https://doi.org/10.1037/h0063487

Guilford, J. P. (1967). *The nature of human intelligence.* McGraw-Hill.

Harrigan, J. A. (2005). Proxemics, kinesics, and gaze. *The new handbook of methods in nonverbal behavior research*, 137–198.

Hartman, R. L., & Johnson, J. D. (1989). Social contagion and multiplexity communication networks as predictors of commitment and role ambiguity. *Human Communication Research*, *15*(4), 523–548.

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*. Springer.

Hilliges, O., Terrenghi, L., Boring, S., Kim, D., Richter, H., & Butz, A. (2007). Designing for collaborative creative problem solving, In *Proceedings of the 6th acm sigchi conference on creativity & cognition*.

Huber, B., Shieber, S., & Gajos, K. Z. (2019). Automatically analyzing brainstorming language behavior with meeter. *Proceedings of the ACM on Human-Computer Interaction*, *3*(Cscw), 1–17.

Ibarra, H., & Andrews, S. B. (1993). Power, social influence, and sense making: Effects of network centrality and proximity on employee perceptions. *Administrative science quarterly*, 277–303.

Jayagopi, D. B., Kim, T., Pentland, A. S., & Gatica-Perez, D. (2010). Recognizing conversational context in group interaction using privacy-sensitive mobile sensors, In *Proceedings of the 9th international conference on mobile and ubiquitous multimedia*. Acm.

Jones, S. R. (1992). Was there a hawthorne effect? *American Journal of sociology*, *98*(3), 451–468.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *13*(4), 351–371.

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.

Kim, T., Chang, A., Holland, L., & Pentland, A. S. (2008). Meeting mediator: Enhancing group collaborationusing sociometric feedback, In *Proceedings of the 2008 acm conference on computer supported cooperative work*.

Kim, T., Hinds, P., & Pentland, A. (2012). Awareness as an antidote to distance: Making distributed groups cooperative and consistent, In *Proceedings of the acm 2012 conference on computer supported cooperative work*.

Kohl, S. (2020). Improving collaboration in design thinking teams through automated coaches, In *Proceedings of 18th european conference on computer-supported cooperative work*. European Society for Socially Embedded Technologies (EUSSET). https://doi.org/10.18420/ecscw2020_dc09

Kohl, S., Calero Valdez, A., & Schröder, K. (2023). Using speech contribution visualization to improve team performance of divergent thinking tasks, In *Proceedings of the 15th conference on creativity and cognition*, Virtual Event, USA, Association for Computing Machinery. https://doi.org/10.1145/3591196.3596824

Kohl, S., Graus, M., & Lemmink, J. (2020). Deciphering the code: Evidence for a sociometric dna in design thinking meetings, In *Hci international 2020*, Switzerland, Springer, Cham. https://doi.org/10.1007/978-3-030-60700-5_7

Kohl, S., Graus, M., & Lemmink, J. G. A. M. (2022a). Context is key: Mining social signals for automatic task detection in design thinking meetings, In *Design, user experience, and usability: Ux research, design, and assessment*, Cham, Springer International Publishing. https://doi.org/10.1007/978-3-031-05897-4_2

Kohl, S., Schröder, K., Graus, M., Efendic, E., & Lemmink, J. G. (2022b). Zooming in on the effect of sociometric signals on different stages of the design process, In *Proceedings of the 14th conference on creativity and cognition*, Venice, Italy, Association for Computing Machinery. https://doi.org/10.1145/3527927.3532810

Kratzer, J., Leenders, O. T. A., & Engelen, J. M. v. (2004). Stimulating the potential: Creative performance and communication in innovation teams. *Creativity and Innovation Management*, *13*(1), 63–71.

Kubasova, U., Murray, G., & Braley, M. (2019). Analyzing verbal and nonverbal features for predicting group performance. *arXiv preprint arXiv:1907.01369*.

Lawford, G. R. (2003). Beyond success: Achieving synergy in teamwork. *The Journal for Quality and Participation*, *26*(3), 23.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *The Journal of Abnormal and Social Psychology*, *46*(1), 38.

Leshed, G., Perez, D., Hancock, J. T., Cosley, D., Birnholtz, J., Lee, S., McLeod, P. L., & Gay, G. (2009). Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups, In *Proceedings of the sigchi conference on human factors in computing systems*.

Liedtka, J. (2015). Perspective: Linking design thinking with innovation outcomes through cognitive bias reduction. *Journal of product innovation management*, *32*(6), 925–938.

Lusher, D., Robins, G., & Kremer, P. (2010). The application of social network analysis to team sports. *Measurement in physical education and exercise science*, *14*(4), 211–224.

Marlow, S. L., Lacerenza, C. N., Paoletti, J., Burke, C. S., & Salas, E. (2018). Does team communication represent a one-size-fits-all approach?: A meta-analysis of team communication and performance. *Organizational behavior and human decision processes*, *144*, 145–170.

McGregor, M., & Tang, J. C. (2017). More to meetings: Challenges in using speech-based technology to support meetings, In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing*.

Mehrabian, A. (1971). *Implicit communication of emotions and attitudes*. California: Wadsworth Publishing Company.

Mumm, J., & Mutlu, B. (2011). Human-robot proxemics: Physical and psychological distancing in human-robot interaction, In *Proceedings of the 6th international conference on human-robot interaction*. Acm.

Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Murray, G., & Oertel, C. (2018). Predicting group performance in task-based interaction, In *Proceedings of the 20th acm international conference on multimodal interaction*.

Nathan, M., Topkara, M., Lai, J., Pan, S., Wood, S., Boston, J., & Terveen, L. (2012). In case you missed it: Benefits of attendee-shared annotations for non-attendees of remote meetings, In *Proceedings of the acm 2012 conference on computer supported cooperative work*.

Olguın, D. O. (2011). *Sensor-based organizational design and engineering* (Doctoral dissertation). Massachusetts Institute of Technology.

Olguın, D. O., Gloor, P. A., & Pentland, A. S. (2009). Capturing individual and group behavior with wearable sensors, In *Proceedings of the 2009 aaai spring symposium on human behavior modeling, sss*.

Olguın, D. O., Waber, B. N., Kim, T., Mohan, A., Ara, K., & Pentland, A. (2008). Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE*

Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), *39*(1), 43–55.
Olson, J. S., Teasley, S., Covi, L., & Olson, G. (2002). The (currently) unique advantages of collocated work. *Distributed work*, 113–135.
O'Neil, H. F., Chuang, S.-h. S., & Baker, E. L. (2010). Computer-based feedback for computer-based collaborative problem solving, In *Computer-based diagnostics and systematic analysis of knowledge*. Springer.
Parker, J. N., Cardenas, E., Dorr, A. N., & Hackett, E. J. (2018). Using sociometers to advance small group research. *Sociological Methods & Research*, 0049124118769091.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Et al. (2011). Scikit-learn: Machine learning in python. *Journal of machine learning research*, *12*(Oct), 2825–2830.
Pentland, A. (2010). *Honest signals: How they shape our world*. MIT press. https://doi.org/10.5555/1450928
Penzkofer, A., Müller, P., Bühler, F., Mayer, S., & Bulling, A. (2021). Conan: A usable tool for multimodal conversation analysis, In *Proceedings of the 2021 international conference on multimodal interaction*.
Perlow, L. A., Hadley, C. N., & Eun, E. (2017). Stop the meeting madness. *Harvard Business Review*, *95*(4), 62–69.
Perry-Smith, J. E., & Mannucci, P. V. (2017). From creativity to innovation: The social network drivers of the four phases of the idea journey. *Academy of Management Review*, *42*(1), 53–79.
Perry-Smith, J. E., & Shalley, C. E. (2003). The social side of creativity: A static and dynamic social network perspective. *Academy of management review*, *28*(1), 89–106.
Poyatos, F. (1983). Language and nonverbal systems in the structure of face-to-face interaction. *Language & Communication*, *3*(2), 129–140.
Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. M. (2021). Literature review on co-located collaboration modeling using multi-

modal learning analyticscan we go the whole nine yards. *IEEE Transactions on Learning Technologies*.

Riedl, C., & Wooley, A. (2020). Successful remote teams communicate in bursts. *Harvard Business Review*.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving, In *Computer supported collaborative learning*. Springer.

Runco, M. A., Illies, J. J., & Eisenman, R. (2005). Creativity, originality, and appropriateness: What do explicit instructions tell us about their relationships? *The Journal of Creative Behavior*, *39*(2), 137–148.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human factors*, *50*(3), 540–547.

Salas, E., Reyes, D. L., & McDaniel, S. H. (2018). The science of teamwork: Progress, reflections, and the road ahead. *American Psychologist*, *73*(4), 593.

Samrose, S., McDuff, D., Sim, R., Suh, J., Rowan, K., Hernandez, J., Rintel, S., Moynihan, K., & Czerwinski, M. (2021). Meetingcoach: An intelligent dashboard for supporting effective & inclusive meetings, In *Proceedings of the 2021 chi conference on human factors in computing systems*.

Samrose, S., Zhao, R., White, J., Li, V., Nova, L., Lu, Y., Ali, M. R., & Hoque, M. E. (2018). Coco: Collaboration coach for understanding team dynamics during video conferencing. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, *1*(4), 1–24.

Sawyer, R. K. (2011). *Explaining creativity: The science of human innovation*. Oxford university press.

Schröder, K., & Kohl, S. (2022). Vicon - towards understanding visual support systems in collaborative video conferencing, In *International conference on digital human modeling and applications in health, safety, ergonomics and risk management*. Springer. https://doi.org/10.1007/978-3-031-05890-5_22

Schröder, K., Kohl, S., & Ajdadilish, B. (2022a). Netimmerse-evaluating user experience in immersive network exploration, In *International conference on digital human modeling and applications in health, safety, ergonomics and risk management*. Springer. https://doi.org/10.1007/978-3-031-06018-2_27

Schröder, K., Kohl, S., de Jongh, F., Putzu, M., Ziefle, M., & Calero Valdez, A. (2022b). Rethinking pension communication–the role of metaphors in information visualization, In *International conference on digital human modeling and applications in health, safety, ergonomics and risk management*. Springer. https://doi.org/10.1007/978-3-031-06018-2_29

Schroeder, K., Kohl, S., & Sinsel, J. (2023). The key to the future–mapping sustainable energy production, In *Proceedings of the 2023 chi conference on human factors in computing systems, workshop: Physicalization from theory to practice: Exploring physicalization design across domains*.

Seidel, V. P., & Fixson, S. K. (2013). Adopting design thinking in novice multidisciplinary teams: The application and limits of design methods and reflexive practices. *Journal of Product Innovation Management*, *30*, 19–33.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153–189.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, *2*, 68.

Solutions, S. (2014). Sociometric badge 03-02: Preliminary user guide. *Boston, MA: Author*.

Stempfle, J., & Badke-Schaub, P. (2002). Thinking in design teams-an analysis of team communication. *Design studies*, *23*(5), 473–496.

Stewart, A. E., Vrzakova, H., Sun, C., Yonehiro, J., Stone, C. A., Duran, N. D., Shute, V., & D'Mello, S. K. (2019). I say, you say, we say: Using spoken language to model socio-cognitive processes during computer-supported collaborative problem solving. *Proceedings of the ACM on Human-Computer Interaction*, *3*(Cscw), 1–19.

Tang, C., Mao, S., Xing, Z., & Naumann, S. (2022). Improving student creativity through digital technology products: A literature review. *Thinking Skills and Creativity*, 101032.

Tchupo, D. E., Kim, J. H., & Macht, G. A. (2020). Fuzzy cognitive maps (fcms) for the analysis of team communication. *Applied ergonomics*, *83*, 102979.

Tripathi, P., & Burleson, W. (2012). Predicting creativity in the wild: Experience sample and sociometric modeling of teams, In *Proceedings of the acm 2012 conference on computer supported cooperative work*.

Vinciarelli, A., Pantic, M., & Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Image and vision computing*, *27*(12), 1743–1759.

Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F., & Schroeder, M. (2011). Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Transactions on Affective Computing*, *3*(1), 69–87.

Wang, T., Keck, M., & Vosough, Z. (2021). Discussion Flows: An Interactive Visualization for Analyzing Engagement in Multi-Party Meetings, In *Eurovis 2021*, The Eurographics Association. https://doi.org/10.2312/evs.20211060

Wasserman, S., Faust, K. Et al. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.

Wilson, R. C., Guilford, J. P., & Christensen, P. R. (1953). The measurement of individual differences in originality. *Psychological bulletin*, *50*(5), 362.

Wiltshire, T. J., Butner, J. E., & Fiore, S. M. (2018). Problem-solving phase transitions during team collaboration. *Cognitive science*, *42*(1), 129–167.

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, 3087.

Woodman, R. W., Sawyer, J. E., & Griffin, R. W. (1993). Toward a theory of organizational creativity. *Academy of management review*, *18*(2), 293–321.

Woolley, A. W., Aggarwal, I., & Malone, T. W. (2015). Collective intelligence and group performance. *Current Directions in Psychological Science*, *24*(6), 420–424.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *science*, *330*(6004), 686–688.

Wu, L., Waber, B. N., Aral, S., Brynjolfsson, E., & Pentland, A. (2008). Mining face-to-face interaction networks using sociometric badges: Predicting productivity in an it configuration task. *Available at SSRN 1130251*.

Yoerger, M., Crowe, J., & Allen, J. A. (2015). Participate or else!: The effect of participation in decision-making in meetings on employee engagement. *Consulting Psychology Journal: Practice and Research*, *67*(1), 65.

Yu, L. (2005). How team communication affects innovation: Good communication is a prerequisite for good teamwork. but how much is enough? *MIT Sloan Management Review*, *46*(4), 7–8.

Zhong, S.-C., Su, B.-H., Huang, W., Liu, Y.-C., & Lee, C.-C. (2020). Predicting collaborative task performance using graph interlocutor acoustic network in small group interaction. *Proc. Interspeech 2020*, 3122–3126.

Impact Paragraph

Most projects in the business world are realized through teamwork (Sawyer, 2011). Teams, by definition, work interdependently, and communication bonds team members. Team communication has many essential roles, enabling information sharing and promoting the exchange of ideas (Cross & Cummings, 2004). The effectiveness of team communication also influences almost every other aspect of cooperation. As such, poor communication has various adverse consequences, disrupting information flow, delaying progress, and, ultimately, causing projects to fail (Salas et al., 2008; Stempfle & Badke-Schaub, 2002).

Recordings and observations have traditionally been used to study team communication, a method that produces only limited data sets and requires significant labor expenditure from the researcher, who can only attend a single meeting at a time. Participant privacy is another issue associated with video and audio documentation, especially in the case of real-world business meetings, which frequently cannot be recorded because of confidentiality concerns. Emerging technology provides promising opportunities to automatically harness high-resolution, quantitative, time-series data about social interactions, enabling researchers to investigate the links between communication and team performance in more detail (Dávila-Montero et al., 2021; Parker et al., 2018).

This dissertation's main objective is to advance understanding of how these emerging technologies can be used to capture information about team interactions during face-to-face and virtual meetings and how researchers can utilize this data to develop means of improving communication. By advancing knowledge surrounding team communication behavior, the presented work enables the development of a visual support system for teams working on collaborative problem-solving (CPS).

Scientific Impact

The thesis demonstrates the value of combining multiple literature streams when investigating the relationship between team communication and team performance in the context of CPS. It draws from the fields of social signal processing (SSP), social network analysis (SNA), computer-supported collaborative work (CSCW), and psychology to show how these distinct domains can synergize to produce richer analyses of team behavior.

At its core, this work contributes to the scientific body of literature by demonstrating the value of utilizing SNA metrics for social signal modeling. Furthermore, by integrating insights from psychology on team cognition for social signal modeling, this thesis overcomes the limitations of existing studies, many of which have struggled to identify significant social signals in their analyses of combinations of problem-solving activities. These new insights can be used by researchers and practitioners working in the CSCW domain to develop automated support systems.

Finally, this thesis provides evidence for the effectiveness of analyzing social interactions and providing real-time feedback on team performance. This thesis takes a first step toward expanding the capabilities of present intelligent systems by enabling computers to understand the cognitive stage of a team to present relevant and actionable interventions.

Societal Impact

The findings of this thesis have significant implications for organizations looking to support their teams. Understanding how to effectively promote team communication is crucial for future business agendas, which will see teams meeting more frequently and handling a broader array of meeting formats, from remote to hybrid to face-to-face. The system developed in Chapter 5 represents a first step toward developing a real-time support system for teams. Developing an understanding of the process for implementing this form of feedback is essential for developing effective team interventions and training tools because research on the effectiveness of feedback for skill development shows that real-time feedback is significantly more effective than near–real-time feedback.

Apart from the practical implications that can be derived from the knowledge generated by this thesis, the papers included also show that it is feasible to collect comprehensive data on team interactions without significantly violating team members' privacy. Although we recorded audio of interactions between subjects during the project's data collection phase, none of this thesis' analyses or findings depend on our ability to recognize the spoken words comprising the encounter's semantic content. Therefore, all insights generated are suitable for repetition in privacy-sensitive settings, and the procedures provide examples of avenues for researchers and practitioners to work within the restrictions of data protection and privacy regulations.

I have endeavored to disseminate the presented research results widely to ensure that the findings benefit other researchers, professional practitioners, and the broader community. Notably, because one of this project's goals was to develop connections between research domains of interest, I have tried to establish connections with these various research communities. Accordingly, these research results have been presented at various seminars, both locally at Maastricht University and internationally at research symposiums and conferences, including the 3rd Meeting Symposium, held in Brussels in May 2022, and the International Conference on Computational Social Science, held in the United States in 2020. Finally, the papers included here have been published in both social science and computer science journals.

Summary

Although team meetings are a crucial aspect of organizational life, research indicates that a significant percentage of meetings are unproductive, with managers often citing low engagement and poor communication as reasons for the lack of productivity. This can be detrimental to workers in knowledge-intensive fields, where job performance depends heavily on communication during collaborative problem-solving activities. This thesis investigates the relationship between team communication and team performance in the context of collaborative problem-solving (CPS) by utilizing social network analysis metrics for social signal modeling. Furthermore, by integrating insights into team cognition from psychology, this thesis overcomes the limitations of existing studies, many of which have struggled to identify significant social signals in their analyses of combinations of problem-solving activities.

Thus, adopting an interdisciplinary approach, this research explores the composition of modalities for different CPS tasks (Chapters 2 & 3), investigates the impact of these modalities on performance (Chapter 4), and examines how researchers and designers can benefit from these insights when researching CPS or designing support tools for teams working on CPS (Chapter 5). The outcomes of this thesis have the potential to inform the design of human–computer interaction systems and contribute to knowledge across multiple fields, including social and computer sciences. The interdisciplinary nature of this research should promote the flow of ideas between these fields and provide a holistic understanding of the relationship between team communication and team performance in the CPS context.

Chapter 2 provides evidence that types of problem-solving activities differ in terms of the social signals that can be collected, demanding separate analysis. Together with my co-authors, I demonstrate that

modeling the collected signals reveals significant pattern differences between different problem-solving activities. In doing so, we reveal what we term the sociometric DNA of an interaction. This notion describes how specific interactions have specific properties of social signals that differ distinctly from other interactions.

Chapter 3 builds on these findings. By capturing and analyzing social signals of rotation in team dynamics, my co-authors and I demonstrate how understanding the sociometric DNA of a problem-solving activity can enable context detection. These findings have value for researchers and practitioners working on building systems that provide more relevant recommendations to teams.

Chapter 4 concludes this thesis' first research strand, providing evidence that social signals relate to team performance and that this relationship varies between problem-solving activities. Certain social signals contribute positively to one activity but less positively to others, a finding that is important for building context-aware systems that can improve team performance.

Chapter 5 utilizes the results of the previous studies to test the impact of a custom-built, real-time visual support system on the performance of teams working on a particular problem-solving activity. My co-authors and I provided teams working on divergent-thinking tasks with a visual support system designed to encourage equal participation by visualizing participation contributions. The findings show that the system positively impacts team performance.

Curriculum Vitae

About

Steffi Kohl was born in Düren, Germany, on February 24, 1991. In 2014, she graduated with distinction (honors) from University College Maastricht with a Bachelor of Liberal Arts and Sciences with a focus on social psychology and business studies. From 2015 to 2017, she continued her education at Maastricht University in the School of Business and Economics, graduating with a Master of Science in International Business: Strategy & Innovation as well as a Master of Science in International Business: Strategic Marketing.

Later in 2017, Steffi joined the Marketing and Supply Chain Management Department as a doctoral researcher. Her research on automated support systems for team communication was supervised by Prof. Dr. Jos G.A.M. Lemmink and Dr. M. Graus. She was invited to present her thesis at the European Conferences on Computer-Supported Cooperative Work (18th Doctoral Colloquium) in 2020 and has published her doctoral work in both social science and computer science journals as well as conference proceedings. From 2017 to 2021, Steffi was the Ph.D. representative for the Marketing and Supply Chain Department at the Ph.D. committee, representing her department at the faculty level. From 2019 to 2021, she was part of the Central Ph.D. platform as the Ph.D. representative for the School of Business and Economics, representing the Ph.D. body of her faculty at the university level in matters of education, research, and communication and working with the University Rector regarding school-wide issues.

In 2021, Steffi joined the Human Data Interaction Lab at Zuyd University of Applied Sciences as a post-doctoral researcher. She conducts research on the interpretation and communication of data. She is also involved as a researcher with projects organized through the ELSA Lab for Poverty and Debt, a coalition of researchers from knowledge institutions and public and private organizations that contribute to knowledge about the development and application of reliable, human-centered artificial intelligence.

Publications

2023

Kohl, S., Calero Valdez, A., & Schröder, K. (2023). Using speech contribution visualization to improve team performance of divergent thinking tasks, In *Proceedings of the 15th conference on creativity and cognition*, Virtual Event, USA, Association for Computing Machinery. https://doi.org/10.1145/3591196.3596824

Schroeder, K., Kohl, S., & Sinsel, J. (2023). The key to the future–mapping sustainable energy production, In *Proceedings of the 2023 chi conference on human factors in computing systems, workshop: Physicalization from theory to practice: Exploring physicalization design across domains*

2022

Ajdadilish, B., Kohl, S., & Schröder, K. (2022). Enhancing Evaluation of Room Scale VR Studies to POI Visualizations in Minimaps, In *Eurovis 2022*, The Eurographics Association. https://doi.org/10.2312/evp.20221136

Kohl, S., Graus, M., & Lemmink, J. G. A. M. (2022a). Context is key: Mining social signals for automatic task detection in design thinking meetings, In *Design, user experience, and usability: Ux research, design, and assessment*, Cham, Springer International Publishing. https://doi.org/10.1007/978-3-031-05897-4_2

Kohl, S., Schröder, K., Graus, M., Efendic, E., & Lemmink, J. G. (2022b). Zooming in on the effect of sociometric signals on different stages of the design process, In *Proceedings of the 14th conference on creativity and cognition*, Venice, Italy, Association for Computing Machinery. https://doi.org/10.1145/3527927.3532810

Schröder, K., & Kohl, S. (2022). Vicon - towards understanding visual support systems in collaborative video conferencing, In *International conference on digital human modeling and applications in health, safety, ergonomics and risk management*. Springer. https://doi.org/10.1007/978-3-031-05890-5_22

Schröder, K., Kohl, S., & Ajdadilish, B. (2022a). Netimmerse-evaluating user experience in immersive network exploration, In *International conference on digital human modeling and applications in health, safety, ergonomics and risk management*. Springer. https://doi.org/10.1007/978-3-031-06018-2_27

Schröder, K., Kohl, S., de Jongh, F., Putzu, M., Ziefle, M., & Calero Valdez, A. (2022b). Rethinking pension communication–the role of metaphors in information visualization, In *International conference on digital human modeling and applications in health, safety, ergonomics and risk management*. Springer. https://doi.org/10.1007/978-3-031-06018-2_29

2020

Kohl, S. (2020). Improving collaboration in design thinking teams through automated coaches, In *Proceedings of 18th european conference on computer-supported cooperative work*. European Society for Socially Embedded Technologies (EUSSET). https://doi.org/10.18420/ecscw2020_dc09

Kohl, S., Graus, M., & Lemmink, J. (2020). Deciphering the code: Evidence for a sociometric dna in design thinking meetings, In *Hci international 2020*, Switzerland, Springer, Cham. https://doi.org/10.1007/978-3-030-60700-5_7