

# Peer feedback to enhance learning in problem-based tutorial groups

Citation for published version (APA):

Kamp, R. J. A. (2013). *Peer feedback to enhance learning in problem-based tutorial groups*. [Doctoral Thesis, Maastricht University]. Uitgeverij BOXPress. <https://doi.org/10.26481/dis.20131128rk>

## Document status and date:

Published: 01/01/2013

## DOI:

[10.26481/dis.20131128rk](https://doi.org/10.26481/dis.20131128rk)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

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

[Link to publication](#)

## General rights

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

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

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

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

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

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

**Peer Feedback to Enhance  
Learning in Problem-Based  
Tutorial Groups**

The research reported here was carried out at



in the School of Health Professions Education



and in the context of research school ICO (Interuniversity Centre for Educational Sciences)

**ico**

ISBN 978-90-8891-731-8

© Rachele J. A. Kamp, Maastricht 2013

Cover design:

Mariet Wielders || M3W

Printed & Lay Out by:

Proefschriftmaken.nl || Uitgeverij BOXPress

# **Peer Feedback to Enhance Learning in Problem-Based Tutorial Groups**

## **PROEFSCHRIFT**

Ter verkrijging van de graad van doctor,  
aan de Universiteit Maastricht,  
op gezag van de Rector Magnificus, Prof. dr. L.L.G. Soete,  
volgens het besluit van het College van Decanen,  
in het openbaar te verdedigen op donderdag  
28 november 2013 om 12:00 uur

door

Rachelle Johanna Adriana Kamp

Geboren op 23 augustus 1984 te Sittard

## **PROMOTORES**

Prof. dr. D. H. J. M. Dolmans

Prof. dr. H. G. Schmidt (Erasmus Universiteit Rotterdam)

## **COPROMOTOR**

Dr. H. J. M. van Berkel

## **BEOORDELINGSCOMMISSIE**

Prof. dr. A. J. J. A. Scherpbier (voorzitter)

Prof. dr. D. Gijbels (Universiteit Antwerpen)

Dr. S. M. M. Loyens (Erasmus Universiteit Rotterdam)

Prof. dr. J. F. M. Metsemakers

Dr. I. H. A. P. Wolfhagen





# CONTENTS

## CHAPTER 1

General introduction 9

## CHAPTER 2

Can students adequately evaluate the activities of their peers in PBL? 29

## CHAPTER 3

The relationship between students' small group activities, time spent on self-study, and achievement 49

## CHAPTER 4

The effect of midterm peer feedback on student functioning in problem-based tutorials 73

## CHAPTER 5

Midterm peer feedback in problem-based learning groups:  
The effect on individual contributions and achievement 97

## CHAPTER 6

General discussion 125

Appendices 142

Summary 149

Samenvatting (*Dutch summary*) 155

Dankwoord (*Acknowledgements*) 163

Curriculum vitae 171

SHE dissertation series 175

ICO dissertation series 181



# **CHAPTER I**

## General Introduction



This introduction chapter will illustrate the background of this study, explain the relevant concepts, present the problem definition as well as the research questions, and give a brief overview of the outline of the thesis.

## **PROBLEM-BASED LEARNING**

Problem-based learning (PBL) is a type of small group learning applied mostly in health education (Hak & Maguire, 2000). It can be considered a constructivist approach to instruction, because it supports the assumption that knowledge is constructed by learners through interaction with their environment (Hendry, Frommer, & Walker, 1999). It consists of three phases during which students work together in groups of approximately 10 students (Barrows, 1988). During the first phase, which is the problem analysis, students are presented with a problem that describes a patient or a situation that is relevant for the future profession of the students. Next, students collaboratively analyze the problem, identify what they already know about the subject (activate their prior knowledge), discuss hypotheses, and formulate learning issues. The second phase is the self-study phase, during which students search for relevant information from multiple sources and study in order to meet their learning issues. During the third phase, students report back to the group what they have found related to the learning issues and collaboratively synthesize 'new knowledge'. Learning during these three phases can be considered a cumulative process; learning in each phase is a precondition for learning in the next phase (Yew, Chng, & Schmidt, 2011). Since PBL stimulates students to take charge of what and how they learn, it can be seen as a student centered instructional approach that promotes self-directed learning. This whole process is guided and supported by a tutor. The tutor acts as a mentor and metacognitive coach and has a dual task: facilitate learning as well as the group process (Papinczak, 2010). The tutorial group meetings are accompanied by a limited number of lectures, which provide information that complements the self-study sources.

Although the success of PBL is affected by multiple factors, the quality of the group discussion has a strong effect on academic achievement and is, therefore, a crucial element of PBL (Gijsselaers & Schmidt, 1990). A deep discussion, and therefore a well-functioning tutorial group, is essential for the construction of 'new knowledge'

(Nieminen, Sauri, & Lonka, 2006). Unfortunately, unproductive group sessions are a widespread and serious problem (Moust, Van Berkel, & Schmidt, 2005). With superficial, haphazard discussions and some students not pulling their weight, unproductive group work is detrimental to effective group learning.

This thesis will focus on the quality of students' individual contributions (which will also be referred to as activities) to the tutorial group discussion, the relationship with student achievement, and the role peer feedback can play in enhancing the quality of these contributions. Student achievement will be defined as the performance of an individual student on a test that addresses the subject matter that was discussed during the tutorial group meetings. Peer feedback is feedback generated through peer evaluation; the process whereby students rate the work of their peers (Falchikov & Goldfinch, 2000).

### **Group processes in PBL**

Within PBL research a considerable number of studies have focused on processes (i.e., process variables) that take place within the tutorial group, which are conducive to learning in PBL.

First of all, several studies have emphasized the importance of cognitive processes (Slavin, Hurley, & Chamberlain, 2003; Van den Bossche, Gijsselaers, Segers, & Kirschner, 2006), which are processes that are thought to enhance the collaborative construction of 'new' and shared knowledge and lead to conceptual change. They occur when group members interact with each other by communicating their knowledge, answering each other's questions and identifying and correcting misconceptions within the group. By engaging in these activities, students elaborate, i.e., they restructure their knowledge by connecting their prior knowledge with new ideas (Van Blankenstein, Dolmans, Van der Vleuten, & Schmidt, 2009; Wittrock, 1992). Two other essential and mutually intertwined constructive processes are co-construction, which is the process of developing and adjusting shared mental models, and constructive conflict, which is the process of identifying and discussing contradicting opinions (Van den Bossche et al., 2006). The occurrence of these processes has been confirmed by several studies (e.g., De Grave, Boshuizen, & Schmidt, 1996; Visschers-Pleijers, Dolmans, De Leng, Wolfhagen, and Van der Vleuten, 2006; Yew and Schmidt, 2009). Van den Bossche et al. (2006) even found a causal

relationship between the occurrence of these processes and the development of shared cognition within the group, and subsequently group effectiveness. Although this study was performed in the context of team-based learning instead of PBL, these findings are relevant for PBL tutorial groups because both instructional approaches can be considered examples of small group learning.

Second of all, research has indicated that tutorial productivity is not only affected by cognitive processes but also by social processes (Dolmans, Wolphagen, & Van der Vleuten, 1998; Slavin et al., 2003). Social processes can be seen as processes that affect the cohesiveness between group members. Social cohesion exists when group members are committed to one another and want the group to succeed. This results in a sense of 'team spirit', i.e., students are willing to help group members and encourage each other to learn (Dolmans et al., 1998). Group members that sponge on or withdraw from the rest of the group can have a negative effect on the cohesion within the group (Van der Linden, Erkens, Schmidt, & Renshaw, 2000). In addition, Van den Bossche et al. (2006) proved that the extent, to which the group members are committed to one another in order to reach the shared goals, had a positive effect on the occurrence of the behaviors that enhance the construction of 'new' and shared knowledge.

Third, motivational processes form an explanation for the success of PBL tutorial groups. According to this perspective, the effectiveness of small group learning lies in the motivating nature of the specific setting of learning within a group (Slavin et al., 2003). Discussing a given problem raises students' curiosity and their intrinsic motivation for the subject matter, which results in active participation within the tutorial group. Hidi and Renninger (2006) refer to this as situational interest, which can be defined as a type of motivation that is elicited by a specific (learning) environment and focuses one's attention on the topic at hand. Within PBL tutorial groups, situational interest was found to be related to academic achievement (Rotgans & Schmidt, 2011).

Although the occurrence of these cognitive, social, and motivational processes and their relationship with group effectiveness have been demonstrated in various studies, research on the relationship between students' *individual contributions* to these processes and student achievement is scarce. Therefore, a closer look needs to be taken at the relationships between a student's individual activities within the tutorial group and his or her achievement.

## Evaluating individual contributions

From the above it can be concluded that, in order for PBL tutorial groups to be successful, students need to contribute to the cognitive processes, to the cohesion within the group, and need to be motivated. From now on this will be referred to as cognitive, social, and motivational contributions to the educational process.<sup>1</sup> Individual contributions within PBL tutorial can be evaluated in three ways: through students' self-evaluations, through the tutor/faculty, or through the peers (Eva, 2001).

When used in PBL tutorial group, all three methods have their pros and cons, but peer evaluation has several important advantages over the other two methods (Eva, 2001). First of all, peer evaluations are highly valid. From research it appears that students in PBL tutorial groups are capable of assessing the performance of their peers (Eva, 2001). This is in line with findings of a study performed by Sluijsmans, Moerkerke, Van Merriënboer, and Dochy (2001). They investigated the reliability of peer evaluations within a PBL setting and found moderately positive results. Papinczak, Young, Groves and Haynes (2007) as well as Sullivan, Hitchcock, and Dunnington (1999) demonstrated a strong concordance between peer and tutor/faculty evaluations. In addition, peer evaluations appear to have a stronger predictive validity in comparison with self-evaluations with regard to academic achievement (Schwartz, Donnelly, Sloan, & Young, 1994). Another advantage of peer evaluation is that multiple evaluations can be combined into one judgment (Eva, 2001). This is contrary to tutor evaluations. Furthermore, peer evaluations possess an additional advantage over tutor/faculty evaluations; they stimulate students to reflect on and evaluate their own performance, which enhances self-directed learning (Sluijsmans, Dochy, & Moerkerke, 1998). When applied in PBL tutorial groups, students indicate that peer evaluations positively affect their learning and enhance their sense of responsibility for group members (Papinczak, Young, & Groves, 2007).

Three types of peer evaluation can be distinguished: peer ranking (where students rank their group peers from worst to best on certain criteria), peer nomination (where a group member is nominated as 'the best' with regard to a certain characteristic), and peer rating (where students rate each other on a set of performance criteria) (Kane & Lawler, 1978). Peer ratings are the only type of peer evaluation that generates a judgment about the quality of students' individual

---

<sup>1</sup> In the following chapters, cognitive processes will also be referred to as constructive processes and social processes will also be referred to as collaborative processes.

contributions (Kane & Lawler, 1978). Therefore, it can be concluded that the relationship between students' individual contributions and their achievement should be investigated using peer ratings.

Instruments that focus on the quality of group processes in PBL are scarce. The ones that do exist (e.g., Loyens, Rikers, & Schmidt, 2007; Rangachari & Crankshaw, 1992; Van Berkel & Van Til, 1994; Visschers-Pleijers, Dolmans, Wolfhagen, & Van der Vleuten, 2005), 1) either do not focus on the contributions of the *individual* student, which makes it impossible to evaluate the *quality of students' individual contributions* and how this is related to *individual* achievement in PBL groups, which is the goal of this study; 2) are not developed as a *peer* rating scale (but instead use self- or tutor evaluations); 3) do not address all relevant processes (i.e., cognitive, social, and motivational) that need to take place within the tutorial group, but focus primarily on cognitive and partially on social processes; or 4) use another form of peer evaluation such as peer ranking or peer nomination instead of peer rating. Therefore, a new peer rating instrument needs to be developed that addresses cognitive, social, as well as motivational processes. More specifically, with this instrument students should be able to evaluate the quality of the cognitive, social, and motivational contributions of their peers. This leads to the first central research question that will be addressed in this thesis (see Figure 1).

*RQ1: Are students able to evaluate the cognitive, social, and motivational contributions to the tutorial group of their peers in a reliable and valid way and to what extent do these contributions predict student achievement?*

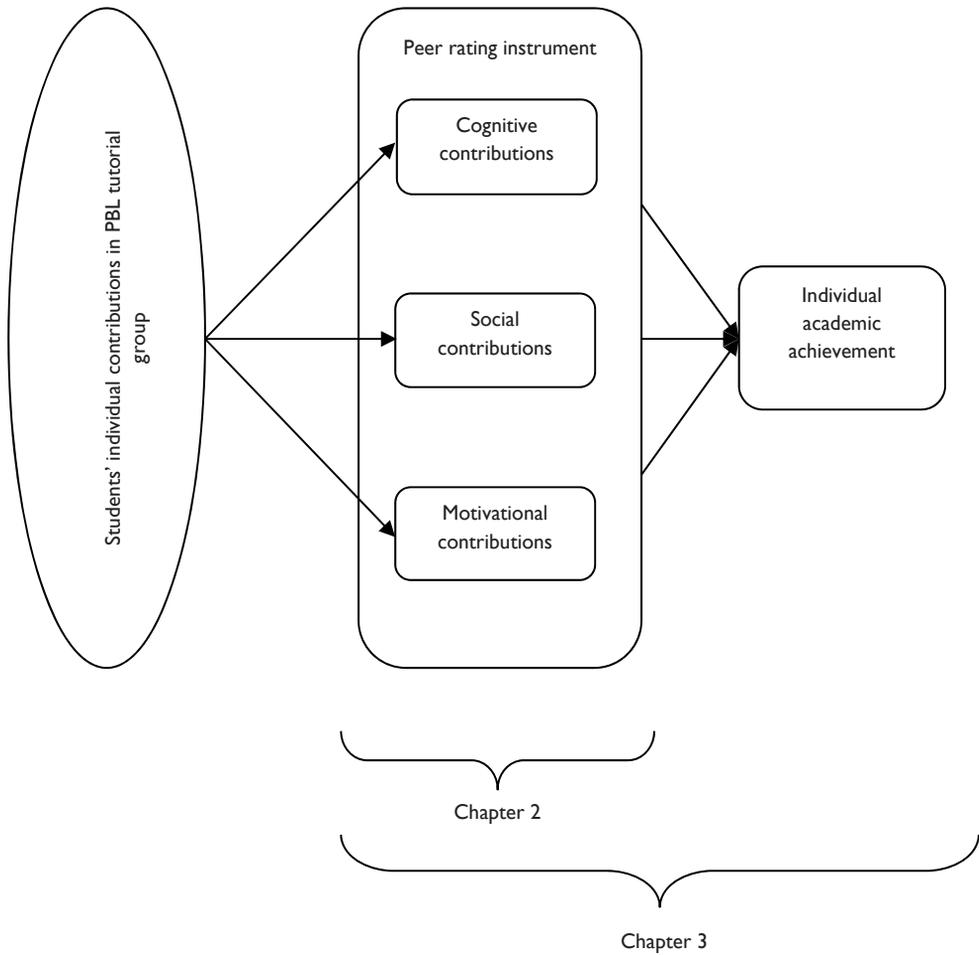


Figure 1. Schematic representation of research question 1.

## FEEDBACK AS AN INTERVENTION

Although it is important for students to actively contribute within the tutorial group, the reality within tutorial groups is often less desirable (Moust et al., 2005). According to Moust et al. “signs of wear” are becoming visible (p. 665). An often seen phenomenon within tutorial groups is that, over time, students tend to spend less time on brainstorming, elaborating, and discussing findings. This leads to a superficial

discussion, which might have detrimental effects on student learning. Not only the cognitive processes are eroded in PBL practice, but the same applies to the social and motivational processes. Decuyper, Dochy, and Van den Bossche (2010) described two high frequency and related problems that occur in small learning groups: free riding and social loafing. Free riding occurs when an individual does not actively participate within the group, but does profit from the achievement of the group. As a consequence of free riding, other students start to withdraw from the group as well. This is detrimental for the cohesion within the group and can cause social loafing (Dolmans et al., 1998). Social loafing implies that students tend to invest less effort in learning when they are learning collaboratively.

The following question, therefore, presents itself: What kind of support is needed to steer the intended group processes that are essential for successful PBL? (Hmelo-Silver, Duncan, & Chinn, 2007). In light of this question peer evaluation seems to be a promising tool. The fact is that one of the major advantages of peer evaluation is that it can be used to generate feedback for individual students (Topping, 2009). Feedback, in general, is thought to promote learning (Hattie & Timperley, 2007). It can make a student aware of the gap between his/her current behavior and the intended or desired behavior (Archer, 2002). Furthermore, feedback affects positive social interdependence by increasing mutual support and collaboration among group members (Hattie & Timperley, 2007). Positive social interdependence implies willingness between group members to help each other because they perceive that they can only achieve their goals if the other members in the group also achieve their goals (Johnson & Johnson, 1995; Johnson, Johnson, & Smith, 2007). Peer feedback is most suitable when it is used in a formative matter (Hattie & Timperley, 2007). This suggests that the peer feedback should serve as input for decisions that the learner needs to make with regard to his or her learning process (Black & Wiliam, 2009).

Providing and receiving peer feedback is, however, by no means a straightforward process and the quality of peer feedback depends on several different factors (Topping, 2010). With regard to these factors two perspectives on the quality of peer feedback can be distinguished. The first perspective focuses on *giving* peer feedback. The central question within this perspective is: Under which conditions can students generate reliable, valid, and constructive judgments about their peers. With regard to this question, Sluijsmans, Brand-Gruwel, and Van Merriënboer (2002), for example, found that it is important to provide students with training on peer

evaluation before they can generate high quality peer feedback. Topping (2009), in addition, emphasizes the importance of clear evaluation criteria and the need to involve students in the clarification of these criteria.

Within this dissertation the focus will, however, be on the second perspective, which focuses on *receiving* peer feedback. Here, the central question is: Under which conditions is the peer feedback received effectively in the sense that it stimulates learning? Within this perspective, it has become evident that it is important to pay attention to the level that the feedback addresses. Hattie and Timperley (2007) discern four levels. First, feedback can address the quality of the task or the product (task level feedback). Second, feedback can focus on the (learning) processes that need to be undertaken by the student in order to produce a high quality product (process level feedback). Third, feedback can address the self-regulation skills needed for the task, (self-regulation level feedback) and fourth, it can address the student himself in an affective and often task unrelated way (self-level feedback). Task level feedback seems to be less effective than feedback that addresses the process and self-regulation level (Hattie & Timperley, 2007). Feedback on students' contributions to the group processes can be mainly classified as process feedback (Kamp, Dolmans, Van Berkel, & Schmidt, 2012).

Experimental research on the effectiveness of peer feedback is, however, scarce (Van Zundert, Sluijsmans, & van Merriënboer, 2010, Strijbos & Sluijsmans, 2010), even more when it concerns peer process feedback in small collaborative learning groups (Geister, Konradt, & Hertel, 2006). The little research that has been performed has largely focused on computer supported collaborative learning (CSCL) environments. Geister et al. (2006), for instance, found that process feedback in virtual teams enhanced motivation, satisfaction, and performance of team members. In another study, which was also performed in a CSCL environment, Phielix, Prins, and Kirschner (2010) provided students with peer feedback (combined with self-evaluation and reflection) on the social and cognitive performance of the group. Especially the social processes of the group seemed to benefit from the peer feedback.

However, in PBL face-to-face groups it remains unclear if, and under which conditions, peer process feedback on the students' contributions affects the quality of these contributions and academic achievement. It has, however, been emphasized that students need to be stimulated to reflect upon the received peer feedback (Prins,

Sluijsmans, & Kirschner, 2006). Reflection can be seen as “a cyclical process which originates from experience and informs future behavior” (Koole et al., 2011, p. 4). During this step it is important to make students aware of the gap between their current and desired behavior and stimulate them to set goals for improvement. Hattie and Timperly (2007) call this feed forward. It is, however, unknown whether PBL tutorial groups can also benefit from peer feedback combined with reflection and goal setting. In addition, the question remains whether it is better to facilitate this reflection and goal setting in PBL tutorial groups collaboratively through a feedback dialogue instead of individually in written form. Even though Phielix et al. (2010) found that individual combined with collaborative reflection had positive effects on the social performance of the group, experimental proof that individual combined with collaborative reflection is more effective than individual reflection alone is still lacking. Furthermore, the study performed by Phielix et al. was executed in a CSCL environment, and not in a face-to-face PBL environment. The second part of the thesis will, therefore, address the facilitation of reflection and goal setting after receiving peer feedback (see Figure 2 for a schematic overview of this research question).

*RQ2: Do reflection and goal setting enhance the effectiveness of a peer feedback intervention, aimed at evaluating students' individual contributions to the tutorial group?*

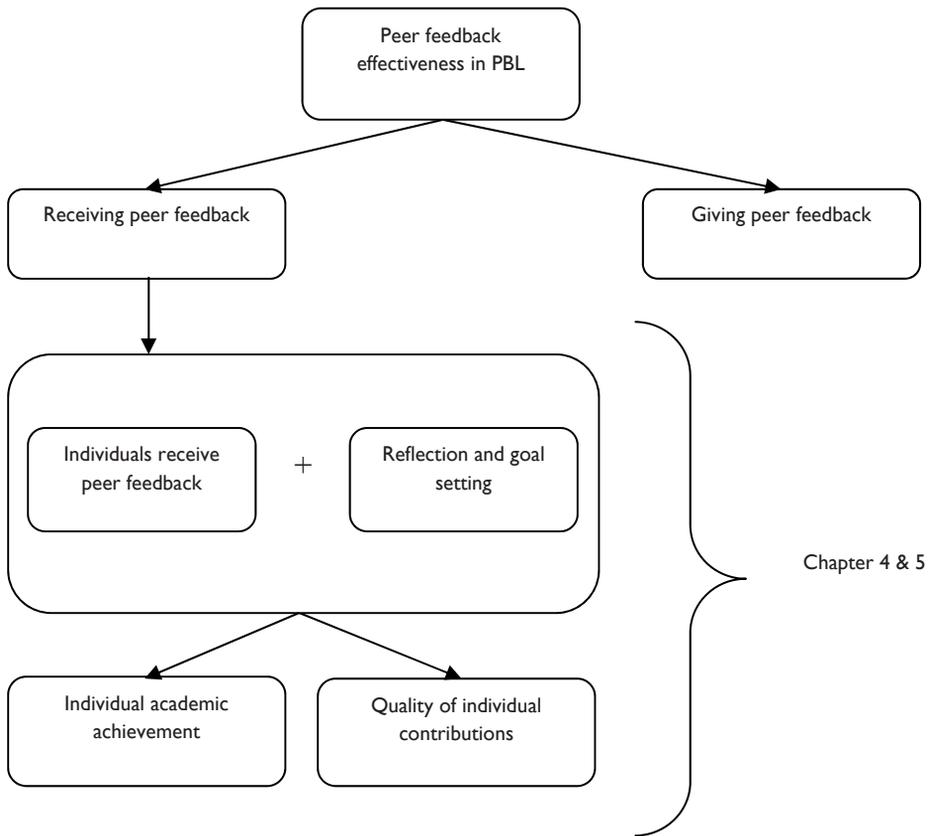


Figure 2. Schematic representation of research question 2.

## OVERVIEW OF THE DISSERTATION

This section will provide a short overview of the remaining chapters of this dissertation. *Chapter 2* describes the development of the Maastricht-Peer Activity Rating Scale (M-PARS). This instrument consists of three subscales: cognitive, social, and motivational contributions to the tutorial group process. In this chapter indications of reliability and validity were investigated. *Chapter 3* focuses on the relationship

between the quality of students' individual contributions, self-study time, and student achievement. The study described in *Chapter 4* investigated to what extent midterm peer feedback improved the quality of individual contributions to the PBL tutorial group, whether this differed for students with different levels of functioning, and how students perceived the effectiveness of peer feedback. This mixed-methods study generated important points of improvement, which were incorporated in the design of the study described in *Chapter 5*. In this chapter the effects of individual versus shared reflection and goal setting on students' individual contributions to the group and their academic achievement were investigated. In addition, the role of prior knowledge on the effectiveness of peer feedback was addressed. Figures 1 and 2 provide a schematic overview of the content of chapters 2 to 5. *Chapter 6* provides an overview of the main conclusions of the four studies. Strengths, limitations, implications, and encountered dilemmas will be discussed. Finally, suggestions for future research will be made.

The chapters in this dissertation are written as individual articles; therefore some repetition of information across chapters may occur.

## REFERENCES

- Archer, J. C. (2010). State of the science in health professional education: Effective feedback. *Medical Education*, 44, 101-108. doi: 10.1111/j.1365-2923.2009.03546.x
- Barrows, H. S. (1988). *The Tutorial Process*. Springfield Illinois, Southern Illinois University School of Medicine.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. doi: 10.1007/s11092-008-9068-5
- Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning: An integrative model for effective team learning in organisations. *Educational Research Review*, 5, 111-133. doi: 10.1016/j.edurev.2010.02.002
- De Grave, W. S., Boshuizen, H. P. A., & Schmidt, H. G. (1996). Problem based learning: cognitive and metacognitive processes during problem analysis. *Instructional Science*, 24(5), 321-341.
- Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (1998). Motivational and cognitive processes influencing tutorial groups. *Academic Medicine*, 73 Suppl 10, S22-S24.
- Eva, K. W. (2001). Assessing tutorial-based assessment. *Advances in Health Sciences Education*, 6, 243-257.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-322. doi: 10.3102/00346543070003287
- Geister, S., Konradt, U., & Hertel, G. (2006). Effects of process feedback on motivation, satisfaction, and performance in virtual teams. *Small Group Research*, 37(5), 459-489. doi: 10.1177/1046496406292337
- Gijsselaers, W. H., & Schmidt, H. G. (1990). Development and evaluation of a causal model of problem-based learning. In Z. M. Nooman, H. G. Schmidt, & E. S. Ezzat (Eds.), *Innovation in medical education: an evaluation of its present status* (pp. 95-113). New York: Springer Publishing.

- Hak, T., & Maguire, P. (2000). Group process: The black box of studies on problem-based learning. *Academic medicine*, *75*(7), 769-772.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. doi: 10.3102/003465430298487
- Hendry, G. D., Frommer, M., & Walker, R. A. (1999). Constructivism and problem-based learning. *Journal of Further and Higher Education*, *23*(3), 359-371. doi: 10.1080/0309877990230306
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111-127. doi: 10.1207/s15326985ep4102\_4
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99-107. doi: 10.1080/00461520701263368
- Johnson, D. W., & Johnson, R. T. (1995). Positive interdependence: Key to effective cooperation. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning*. (pp. 174-199). New York, NY: Cambridge University Press.
- Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, *19*, 15-29. doi: 10.1007/s10648-006-9038-8
- Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2012). The effect of midterm peer feedback on student functioning in problem-based tutorials. *Advances in Health Sciences Education*. Advance online publication. doi:10.1007/s10459-012-9364-1
- Kane, J. S. & Lawler, E. E. (1978). Methods of peer assessment. *Psychological Bulletin*, *85*, 555-586.
- Koole, S., Dornan, T., Aper, L., Scherpbier, A., Valcke, M., Cohen-Schotanus, J., & Derese, A. (2011). Factors confounding the assessment of reflection: A critical review. *BMC Medical Education*, *11*, 104. doi: 10.1186/1472-6920-11-104
- Loyens, S. M. M., Rikers, R. M. J. P., & Schmidt, H. G. (2007). Students' conceptions of distinct constructivist assumptions. *European Journal of Psychology of Education*, *12*, 179-199.

- Moust, J. H. C., Van Berkel, H. J. M., & Schmidt, H. G. (2005). Signs of erosion: Reflections on three decades of problem-based learning at Maastricht University. *Higher Education, 50*, 665-683. doi: 10.1007/s10734-004-6371-z
- Nieminen, J., Sauri, P., & Lonka, K. (2006). On the relationship between group functioning and study success in problem-based learning. *Medical Education, 40*(1), 64-71. doi: 10.1111/j.1365-2929.2005.02344.x
- Papinczak, T. (2010). An exploration of perceptions of tutor evaluation in problem-based learning tutorials. *Medical Education, 44*(9), 892-899. doi: 10.1111/j.1365-2923.2010.03749.x
- Papinczak, T., Young, L., & Groves, M. (2007). Peer assessment in problem-based learning: a qualitative study. *Advances in Health Sciences Education, 12*, 169-186. doi: 10.1007/s10459-005-5046-6
- Papinczak, T., Young, L., Groves, M., & Haynes, M. (2007). An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Medical Teacher, 29*, e122-e132. doi: 10.1080/01421590701294323
- Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior, 26*(2), 151-161. doi: 10.1016/j.chb.2009.10.011
- Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles, and preferences. *Advances in Health Sciences Education, 11*(3), 289-303. doi: 10.1007/s10459-005-3250-z
- Rangachari, P. K., & Crankshaw, D. J. (1992). *Formative evaluation in a tutorial-based course*. Paper presented at the Proceedings of 5th Instructional Show and Tell for Ontario Universities, Guelph, ON.
- Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction, 21*(1), 58-67. doi: 10.1016/j.learninstruc.2009.11.001
- Schwarz, R. W., Donnelly, M. B., Sloan, D. A., & Young, B. (1994). Knowledge gain in a problem-based surgery clerkship. *Academic Medicine, 69*(2), 148-151.
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. (2003). Cooperative learning and achievement: theory and research. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology, Vol. 7*. (pp. 177-198). Hoboken (NJ): John Wiley & Sons Inc.

- Sluijsmans, D. M. A., Brand-Gruwel, S., & Van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, 27(5), 443-454. doi: 10.1080/0260293022000009311
- Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, 1(3), 293-319. doi: 10.1023/a:1009932704458
- Sluijsmans, D. M. A., Moerkerke, G., Van Merriënboer, J. J. G., & Dochy, F. J. R. C. (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation*, 27, 153-173.
- Strijbos, J. -W., & Sluijsmans, D. M. A. (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction*, 20, 265-269. doi: 10.1016/j.learninstruc.2009.08.002
- Sullivan, M. E., Hitchcock, M. A., & Dunnington, G. L. (1999). Peer and self-assessment during problem-based tutorials. *The American Journal of Surgery*, 177, 266-269.
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48, 20-27. doi: 10.1080/00405840802577569
- Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, 20(4), 339-343. doi: 10.1016/j.learninstruc.2009.08.003
- Van Berkel, H. J. M. & Van Til C. T. (1994). *Group performing in a problem-based curriculum: The development and evaluation of an instrument*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Van Blankenstein, F. M., Dolmans, D. H. J. M., Van der Vleuten, C. P. M., & Schmidt, H. G. (2009). Which cognitive processes support learning during small-group discussion? The role of providing explanations and listening to others. *Instructional Science*, 39(2), 189-204. doi: 10.1007/s11251-009-9124-7
- Van den Bossche, P., Gijsselaers, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments. Team learning beliefs & behaviors. *Small Group Research*, 37(5), 490-521. doi: 10.1177/1046496406292938

- Van der Linden, J., Erkens, G., Schmidt, H., & Renshaw, P. (2000). Collaborative learning. In R. J. Simons, J. Van der Linden & T. Duffy (Eds.), *New Learning* (pp. 37-54). Dordrecht, The Netherlands: Kluwer.
- Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction, 20*(4), 270-279. doi: 10.1016/j.learninstruc.2009.08.004
- Visschers-Pleijers, A. J. S. F., Dolmans, D. H. J. M., de Leng, B. A., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2006). Analysis of verbal interactions in tutorial groups: A process study. *Medical Education, 40*(2), 129-137. doi: 10.1111/j.1365-2929.2005.02368.x
- Visschers-Pleijers, A. J. S. F., Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2005). Development and validation of a questionnaire to identify learning-oriented group interactions in PBL. *Medical Teacher, 27*(4), 375-381. doi: 10.1080/01421590500046395
- Wittrock, M. C. (1992). Generative learning processes of the brain. *Educational Psychologist, 27*(4), 531-541.
- Yew, E. H. J., Chng, E., & Schmidt, H. G. (2011). Is learning in problem-based learning cumulative? *Advances in Health Sciences Education, 16*(4), 449-464. doi: 10.1007/s10459-010-9267-y
- Yew, E. H. J., & Schmidt, H. G. (2009). Evidence for constructive, self-regulatory, and collaborative processes in problem-based learning. *Advances in Health Sciences Education, 14*, 251-273. doi: 10.1007/s10459-008-9105-7





# CHAPTER 2

## Can Students Adequately Evaluate the Activities of their Peers in PBL?

Published as:

Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2011). Can students adequately evaluate the activities of their peers in PBL? *Medical Teacher*, 33(2), 145-150



**ABSTRACT**

In problem-based learning (PBL), high-quality discussions are crucial for student learning. The quality of the discussion is affected by the quality of the contributions students make during PBL tutorials. This study investigated whether students are able to evaluate the activities of their peers in PBL groups in a reliable and valid way. For this purpose, the Maastricht-Peer Activity Rating Scale (M-PARS) was developed. The M-PARS was well-founded on the literature about effective tutorial performance. With this scale, students (N=196) were evaluated by their peers on three main aspects: their constructive, collaborative, and motivational contributions to the group. A confirmatory factor analysis demonstrated that the data fitted the three-factor model reasonably well. The generalizability studies demonstrated good internal consistency when students were evaluated by, at least, four of their peers. Furthermore, Hancock's coefficients indicated good construct reliability. The results prove that peers are able to provide reliable and valid information about a student's active participation in the tutorial group, if at least four peer ratings are attainable, out of a group of eight students. In conclusion, the M-PARS is a valid and reliable instrument.

## INTRODUCTION

Problem-based learning (PBL) is a student-centered instructional strategy often used in medical schools all over the world. In PBL, students work together in small groups in which they discuss the subject matter. High-quality discussions are crucial for successful student learning in PBL (Gijsselaers & Schmidt, 1990). The quality of the contributions that students make during tutorial group meetings strongly affects the quality of the discussion and therefore group functioning. The purpose of this study was to test whether students are able to evaluate the tutorial performance of their peers in a reliable and valid way. Therefore, the Maastricht-Peer Activity Rating Scale (M-PARS) was developed, which measures a student's active participation in the tutorial group.

Up till now, several studies which focus on measuring tutorial group effectiveness have been described in the literature. These studies (e.g., De Grave, Boshuizen, & Schmidt, 1996; Das, Mpofu, Dunn, & Lamphear, 1998; Dolmans, De Grave, Wolfhagen, & Van der Vleuten, 1998; Visschers-Pleijers, Dolmans, Wolfhagen, & Van der Vleuten, 2005) are often based on students' self-evaluations. However, research has shown that reliability and validity of self-evaluations are not always optimal (Das et al., 1998; Eva, 2001; Rezler, 1989; Tousignant & DesMarchais, 2002). Peer ratings, on the other hand, are often associated with good reliability and validity (Eva, 2001). Therefore, peer evaluations could serve as a more suitable and effective alternative for the evaluation of tutorial group performance.

Research has, for instance, proved that the congruence between self-evaluations and tutor/faculty evaluations is not sufficient (Papinczak, Young, Groves, & Haynes, 2007b; Rezler, 1989; Sullivan, Hitchcock, & Dunnington, 1999). Papinczak et al. (2007b), Rezler (1989), and Sullivan et al. (1999) compared tutor/faculty evaluations with self-evaluation in a PBL setting on, among others, the following dimensions: reasoning, knowledge, information processing, responsibility and respect, critical analysis, communication skills, and interpersonal skill. In general, all of these studies found only small-to-moderate correlations between self- and tutor/faculty evaluation. Several studies also concluded that self-evaluations have a poor predictive validity, in the sense that they correlate poorly with achievement. Rezler (1989), for example, studied the correlations between a student's self-evaluation and his/her performance on an examination that measured knowledge of the basic sciences. The results demonstrated only a small, nonsignificant correlation between the two. This implies that students who evaluate their own tutorial

performance as 'good' are not always the ones who perform well on an examination. Tousignant and DesMarchais (2002) also compared self-evaluations with examination performance. Here, students were presented with a medical problem that they had to solve verbally. Here too, only a small correlation was found between self-evaluations and performance. The poor predictive validity of self-evaluations could possibly be explained by the fact that students have a tendency to either underestimate or overestimate their own performance (Boud & Falchikov, 1989).

Peer ratings, on the other hand, exhibit a higher standard of predictive validity. Schwartz, Donnelly, Sloan, and Young (1994) found a high correlation between peer evaluations of tutorial performance and learning achievement, suggesting that students who received a good evaluation from their peers also gained more knowledge than the students who received a poor evaluation. Papinczak et al. (2007b) and Sullivan et al. (1999) also compared peer with tutor ratings of tutorial performance and found moderate-to-large correlations. This indicates that peer evaluations bear a stronger resemblance to tutor/faculty evaluations than self-evaluations. Not only has the validity of peer ratings been proved, but Sluijsmans, Moerkerke, Van Merriënboer, and Dochy (2001) also found (moderately) positive results concerning the reliability of peer ratings in PBL tutorial groups.

Furthermore, Papinczak, Young, and Groves (2007a) studied the attitudes and perceptions of students with regard to peer ratings during PBL and found that students experienced an increased sense of responsibility and felt their learning performance improved when they were asked to evaluate their peers. A recent study by Schönrock-Adema, Heijne-Penninga, Van Duijn, Geertsma, and Cohen-Schotanus (2007) confirms these perceptions. They found that assessment of tutorial behavior improves the tutorial behavior of the assessing student. The findings of the previously mentioned studies seem to favor peer evaluations over self-evaluations, in the sense that peers have a more accurate, reliable, and valid picture of a student's tutorial performances than the students themselves.

It is therefore striking that instruments applied to measure a student's activity in a tutorial group based on peer evaluations are scarce. The ones that do exist appear to be either ill-founded on the literature about effective student behavior in tutorial groups (Van Berkel & Van Til, 1994) or not adequately validated as a peer evaluation instrument (Papinczak et al., 2007b). For that reason, a new rating scale was developed for this study,

which was strongly based on what is known from the literature about effective student behavior in a tutorial group.

Three key issues or prerequisites were derived from the literature. The first prerequisite for effective learning in tutorial groups is that students actively construct and reconstruct their knowledge in the group by means of summarizing, asking critical questions, and correcting misconceptions (Dolmans, De Grave, Wolfhagen, & Van der Vleuten, 2005). Visschers-Pleijers, Dolmans, De Leng, Wolfhagen, and Van der Vleuten (2006) and Yew and Schmidt (2009) proved that these processes do occur during tutorials in forms of posing questions, elaborating, and reasoning. In addition, Dolmans et al. (1998) found that these cognitive processes have a positive influence on the group's productivity. The second important prerequisite for effective student learning in tutorial groups is that students collaborate strongly with each other (Slavin, Hurley, & Chamberlain, 2003). Collaborative learning takes place when group members share a common goal, share responsibilities, are mutually dependent and need to reach agreement through interaction (Van der Linden, Erkens, Schmidt, & Renshaw, 2000). Yew and Schmidt (2009) demonstrated that these collaborative processes, such as sharing information, take place in different phases of the PBL process. Effective collaboration is also important because it can lead to increased cohesion between group members and to a feeling of 'team spirit' (Slavin et al., 2003). Research has also demonstrated that students who withdraw from the group seem to have a negative influence on group cohesion and therefore the group's productivity (Dolmans et al., 1998). A third important condition for effective tutorial groups is that students be highly motivated and actively participate in the group discussions. Hidi and Renninger (2006) describe a type of motivation characterized by a willingness to engage in a particular content or situation. They refer to this as situational interest, which originates from an interaction between a person and a specific content or event (Krapp, 2000). More specifically, learning environments such as tutorial groups can facilitate the development of situational interest. According to Hidi and Renninger (2006) situational interest has a positive influence on academic learning.

Based on these three main conditions for effective PBL tutorial group sessions, the M-PARS was developed for this study. Its purpose was to measure the extent in which an individual student's activities, as rated by his/her peers, promotes constructive and collaborative learning, and strengthens motivation for group work. This instrument was then applied to answer the main research question: are students able to evaluate the tutorial group activities of their peers in a reliable and valid way?

## METHOD

### Participants

Participants included 223 first-year health science (PBL) students at Maastricht University. The students were divided into 26 tutorial groups, each group consisting of approximately 8 students. Within these tutorial groups students discussed a given problem and generated learning issues for further self-study. During the following tutorial, students discussed and synthesized their findings. Students met once a week for a 2-h session over a period of eight weeks.

### Instrument

Student activity was measured with the M-PARS, which consisted of 34 items. Students were asked to evaluate each of their peers individually by indicating how much they agreed with every item on a 5-point Likert scale (1, completely disagree; 2, disagree; 3, neutral; 4, agree; and 5, completely agree). These items were divided into three subscales under the assumption that they were representative of the three key issues: constructive, collaborative, and motivational activities. Students were also asked to score the overall activities of each peer in the tutorial group on a scale from 1 (very poor) to 10 (excellent).

### Procedure

This research was initiated with the development of the M-PARS. For this purpose, relevant items that suited one of the three main conditions (constructive, collaborative, and motivational activities) were collected. These items were selected from several existing self- and tutor rating scales (Das et al., 1998; Visschers-Pleijers et al., 2005; Loyens, Rikers, & Schmidt, 2007). Next, the rating scale was distributed among a small group of students and subject experts. They were asked to name characteristics of 'good' and 'poor' peer students and to assess the face validity of the scale. Based on the comments of the students and the experts, the scale was revised and relevant characteristics that were not already apparent in the scale were then added. At the beginning of the 8-week course, students were encouraged to observe their peers' behavior during the tutorial group sessions. The rating scale was administered to the students during the sixth week of this 8-week course. Students received an envelope with approximately nine rating scales - one rating scale for each

group member. Students were instructed that the data would be analyzed on an anonymous basis and would not be reported to their peers. Students were also instructed to submit the completed rating scales at the conclusion of the course and were given a small gift certificate for their participation.

### Analyses

Since the aim of this study was to investigate whether or not students could make reliable and valid evaluations of their peers, the data were aggregated at the student level by computing mean scores on the items for the different students. Next, construct validity of the rating scale was examined with a confirmatory factor analysis using Amos 7.0 (Arbuckle, 2006). This analysis was carried out to determine whether the data conformed with the three-factor model. Model fit was assumed if the greater part of the following seven conditions were met: (1) the chi-squared divided by the degrees of freedom (CMIN/df) should be lower than 3 with a significant *p*-value (Kline, 1998); (2) the root mean square error approximation (RMSEA) should be less than 0.1 (Browne & Cudeck, 1993); (3) the Tucker–Lewis coefficient (TLI) and; (4) the comparative fit index (CFI) should both be higher than 0.90 (Hu & Bentler, 1999); (5) the root mean square residual (RMR) should be less than 0.07; (6) the goodness-of-fit index (GFI); and (7) the adjusted goodness-of-fit index (AGFI) should both be higher than 0.80. These last three indices were specified by Saris and Stronkhorst (1984). The construct validity was also tested by examining the correlations between the items and the overall score and between the subscale scores and the overall score. In addition, generalizability studies were conducted in order to examine how many different peer ratings per individual student are necessary to ensure a reliable evaluation of one student (Crick & Brennan, 1983). According to Grönlund (1988), the generalizability-coefficient (G-coefficient) should be between 0.60 and 0.80. Therefore, a minimal G-coefficient of 0.70 was applied in this study. The standard error mean (*SEM*) is an estimation of the 95% confidence interval divided by 1.96 (Ferguson, 1981). This means that a 5-point scale with a ½ point confidence interval should have an *SEM* lower than 0.26 and a 10-point scale with a confidence interval of 1 point should have an *SEM* lower than 0.51. Furthermore, the construct reliability of the rating scale was examined by computing Hancock’s coefficients for the three factors (Hancock & Mueller, 2001). The following formula was used for the calculation of Hancock’s coefficient:

$$H = 1 / \left[ 1 + \frac{1}{\frac{\rho_1^2}{1 - \rho_1^2} + \dots + \frac{\rho_p^2}{1 - \rho_p^2}} \right]$$

When assessing the reliability of latent constructs, such as the current three-factor model, Hancock's coefficients are preferred over the more commonly used Cronbach's alpha coefficients. The *H*-coefficient directly assesses the reliability of the construct itself, because it is based on the factor loadings of all indicators belonging to one factor. In other words, the *H*-coefficient represents the percentage of variance that is explained by the construct. This is in contrast with the Cronbach's alpha coefficient because it is based only on the interrelatedness of the indicators. The recommended minimum for Hancock's *H*-coefficient lies between 0.70 and 0.80 (Hancock & Mueller, 2001).

## RESULTS

A total of 1111 completed rating scales were submitted. With the application of these rating scales, 196 different students were evaluated and rated by four to eight of their peers. This was a response rate of 87%.

### Construct validity

Initially, the confirmatory factor analysis of the three-factor model demonstrated a poor fit. Therefore, items were removed based on content overlap between items, and on the fit indices generated by Amos. More parsimonious, alternative models of effective student activity during tutorials were also tested. A three-factor model consisting of 14 items generated the closest fit with the data. The items, means, and standard deviations are depicted in Table 1.

Table I

*Items, Means, and Standard Deviations of the Maastricht Peer Activity Rating Scale (M-PARS)*

<b>Items</b>	<b>Mean (1-5)</b>	<b>Standard deviation</b>
<i>Constructive activities</i>		
1	This student was able to make adequate summaries	3.45 0.49
2	This student was able to make a distinction between the main and lateral issues in the subject matter	3.49 0.42
3	This student asked critical questions	3.24 0.57
4	This student corrected misconceptions about the subject matter	3.28 0.50
5	This student contributed to a better understanding of the subject matter	3.58 0.54
<i>Collaborative activities</i>		
6	This student had a positive influence on the group	3.74 0.46
7	This student felt responsible for the group	3.22 0.49
8	This student promoted collaboration between group members	3.41 0.44
9	This student was willing to share his/her information	4.02 0.40
10	This student was committed to the group	3.65 0.50
<i>Motivational activities</i>		
11	This student demonstrated motivation	3.64 0.59
12	This student participated well	3.75 0.58
13	This student actively participated during the brainstorm sessions	3.67 0.52
14	This student contributed more than other group members	2.92 0.67
15	Score this student's overall activity in the tutorial group sessions on a scale from 1 (poor) tot 10 (excellent).	7.23 0.81

*Note.* 1 = completely disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = completely agree.

This three-factor model produced, among others, the following results:  $\chi^2$  (74,  $N = 196$ ) = 187.28,  $p = 0.000$ , TLI = 0.95, CFI = 0.96 and GFI = 0.88 (Table 2). Seven out of eight earlier specified conditions were met, which indicated an acceptable fit. Because the correlations between the three factors were fairly high ( $r = 0.86, 0.88, 0.88$ ), confirmatory factor analyses were also carried out on the one- and two-factor models. These models were based on theoretical assumptions. The fit indices of the one- and two-factor models indicated a poorer fit compared to the three-factor model (Table 2). After this, the three-factor model was cross-validated by splitting the dataset into two random subsets. Both subsets generated acceptable and similar results (Table 2).

Table 2

*Fit Indices for the One-, Two-, and Three-Factor, and for the Two Random Subsets of the Three-Factor Model*

	<b>N</b>	<b><math>\chi^2</math></b>	<b>df</b>	<b>p</b>	<b>Cmin/ df</b>	<b>RMSEA</b>	<b>TLI</b>	<b>CFI</b>	<b>RMR</b>	<b>GFI</b>	<b>AGFI</b>
<i>Number of factors</i>											
1	196	262.0	77	0.000	3.403	0.111	0.924	0.935	0.009	0.831	0.769
2	196	238.1	76	0.000	3.132	0.105	0.932	0.943	0.009	0.845	0.786
3	196	187.3	74	0.000	2.531	0.089	0.951	0.960	0.008	0.876	0.824
<i>Sub-set 1</i>	98	143.4	74	0.000	1.938	0.098	0.945	0.955	0.011	0.823	0.749
<i>Sub-set 2</i>	98	127.0	74	0.000	1.716	0.086	0.951	0.960	0.008	0.843	0.777

Furthermore, strong positive correlations were found between the overall activity score and the 14 individual items ( $r > 0.53, p < 0.01$ ) as well as between the overall activity score and the three factor scores ( $r = 0.88, r = 0.89, r = 0.94, p < 0.01$ ).

## Generalizability

G-coefficients and SEMs were calculated for the three factors, the total score, and the overall evaluation. The results indicated that four or more ratings gave acceptable G-coefficients and SEMs (Table 3).

Table 3

*Generalizability Coefficients (G-Coefficient) and Standard Errors of Measurement (SEM), Based on Three, Four, Five, and Six Ratings*

	Raters, <i>n</i>	G-Coefficient	SEM
Factor 1: <i>Constructive activities</i>	3	.66	.28
	4	.72	.24
	5	.76	.22
	6	.79	.20
Factor 2: <i>Collaborative activities</i>	3	.64	.27
	4	.70	.24
	5	.75	.21
	6	.78	.19
Factor 3: <i>Motivational activities</i>	3	.71	.31
	4	.76	.27
	5	.80	.24
	6	.83	.22
Total scale	3	.72	.25
	4	.77	.21
	5	.81	.19
	6	.84	.18
Overall score	3	.72	.47
	4	.77	.41
	5	.81	.37
	6	.83	.33

### Construct reliability

Table 4 exhibits the means, standard deviations, range, and Hancock's coefficients for the three subscales, the total scale, and the overall score. The Hancock's coefficients for the three factors: 0.93, 0.93, and 0.96, respectively, indicated very good construct reliability for the three-factor model.

Table 4

*Mean Scores, Standard Deviations and Hancock's H Coefficients for the Three Factors, the Total Scale, and the Overall Grade (1 'Poor' – 10 'Excellent') Based on Four to Eight Ratings per Individual Student*

	Items, n	Students, N	Min.	Max.	Mean	Standard deviation	Hancock's coefficient
Factor 1: Constructive activities	5	196	2.0	4.4	3.4	0.4	0.93
Factor 2: Collaborative activities	5	196	1.9	4.3	3.6	0.4	0.93
Factor 3: Motivational activities	4	196	1.7	4.6	3.5	0.5	0.96
Total Scale	14	196	1.9	4.3	3.5	0.4	-
Overall score	1	196	3.7	9.00	7.2	0.8	-

Note. 1 = completely disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = completely agree.

## DISCUSSION

The purpose of this study was to investigate whether students are able to evaluate the performance of their peers in a problem-based tutorial group setting in a reliable and valid way. For this reason, a theoretically grounded rating scale was developed that evaluated three main aspects of tutorial activity: a student's constructive, collaborative, and motivational contributions to the tutorial group.

Subsequently, this rating scale was administered to first-year health science students who participated in tutorial groups. Based on the results of the confirmatory factor analysis and content overlap, several items were removed, which resulted in a valid and reliable rating scale. In accordance with the hypothesis, this study demonstrated that in the specific setting of a tutorial group, students are indeed able to accurately evaluate the performance of their peers. Furthermore, these peer group ratings proved to have high inter-rater agreement. The generalizability studies indicated that a minimum of only four peers rating a student will generate a reliable evaluation. These findings are in line with previous findings reported by Eva (2001), that generally speaking, peer evaluations are highly reliable and valid.

In addition, this study has resulted in a unique instrument with which activities that are conducive for learning in PBL tutorials can be measured. The M-PARS has many advantages over existing instruments. First of all, it is based on peer group evaluations which are thought to be more valid and reliable than self-evaluations (Eva, 2001; Papinczak et al., 2007b). Furthermore, this scale meets two important conditions that previously developed peer rating scales often lacked. First of all, this scale is adequately validated as a peer evaluation instrument. Second, it is well-founded on the literature about effective student performance in a tutorial group. Third, the advantage of the M-PARS is that it is a very small scale (only 14 items) and it requires only four raters per individual student, to carry out the rating, which makes this rating scale very easy to use. With regard to the practical relevance, this rating scale could serve as a monitoring instrument which a tutor can use to improve tutorial group processes. The M-PARS could also be applied as a formative evaluation tool in order to stimulate students to reflect on their own tutorial behavior. This not only applies to the student being evaluated, but also to the one carrying out the evaluation. For the student being evaluated, the scores on the three subscales of the rating scale provide a clear overview of which activities he/she already displays and which activities call for improvement. For the student carrying out the evaluation, it could be beneficial in the sense that it might enhance their own tutorial performance (Schönrock-Adema et al., 2007).

One of the limitations present in this study was that the correlation between the three subscales of the M-PARS was fairly high. This could be an indication that the three subscales actually only measure one and the same factor. However, when the hypothesis that all 14 items represent only one factor was

tested the majority of the fit indices did not meet the requirements. This indicated a poor model fit and as a result, this hypothesis was rejected. Future research could focus on the influence of tutorial performance on other variables that are thought to play an important role in PBL. Examples of these variables include achievement scores and the amount of time devoted to self-study. It would be interesting to know whether students with a high score on one or more of the M-PARS' subscales also devote more time to self-study and achieve higher grades than students with lower scores. Correlating the three subscales with achievement scores and self-study time could elucidate which tutorial activities are the most essential to achieve successful PBL. In other words, it could shed some light on the 'key activities' required by tutorial groups.

## REFERENCES

- Arbuckle, J. L. (2006). *Amos 7.0 user's guide*. Chicago, IL: SPSS Inc.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education, 18*, 529-549.
- Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Crick, J. E., & Brennan, R. L. (1983). *Manual for Genova: A Generalized Analysis of Variance System*. Iowa: American College Testing Program.
- Das, M., Mpofu, D., Dunn, E., & Lanphear, J. H. (1998). Self and tutor evaluations in problem-based learning tutorials: Is there a relationship? *Medical Education, 32*(4), 411-418.
- De Grave, W. S., Boshuizen, H. P. A., & Schmidt, H. G. (1996). Problem based learning: Cognitive and metacognitive processes during problem analysis. *Instructional Science, 24*(5), 321-341.
- Dolmans, D. H. J. M., De Grave, W., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2005). Problem-based learning: Future challenges for educational practice and research. *Medical Education, 39*, 732-741. doi: 10.1111/j.1365-2929.2005.02205.x
- Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (1998). Motivational and cognitive processes influencing tutorial groups. *Academic Medicine, 73*(Suppl 10), 22-S24.
- Eva, K. W. (2001). Assessing tutorial-based assessment. *Advances in Health Sciences Education, 6*, 243-257.
- Ferguson, G. A. (1981). *Statistical analysis in psychology and education. 5th ed.* Auckland: McGraw-Hill.
- Gijsselaers, W. H., & Schmidt, H. G. (1990). Development and evaluation of a causal model of problem-based learning. In Z. Nooman, H. G. Schmidt, & E. Ezzat (Eds.), *Innovation in medical education: An evaluation of its present status* (pp. 95-113). New York: Springer Publishing.
- Grönlund, N. E. (1988). *How to construct achievement tests. 4th ed.* Englewood Cliffs (NJ): Prentice-Hall Inc.

- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit & D. Sörbom (Eds.), *Structural equation modeling: Present and future - A Festschrift in honor of Karl Jöreskog* (pp. 195-219). Lincolnwood (IL): SSI, Scientific Software International.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127. doi: 10.1207/s15326985ep4102\_4
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York (NY): Guilford Press.
- Krapp, A. (2000). Interest and human development during adolescence: an educational-psychological approach. In J. Heckhausen (Ed.), *Motivational psychology of human development: developing motivation and motivating development*. (pp. 109-129). New York (NY): Elsevier Science.
- Loyens, S. M. M., Rikers, R. M. J. P., & Schmidt, H. G. (2007). The impact of student's conceptions of constructivist assumptions on academic achievement and drop-out. *Studies in Higher Education, 32*(5), 581-602. doi: 10.1080/03075070701573765
- Papinczak, T., Young, L., & Groves, M. (2007a). Peer assessment in problem-based learning: A qualitative study. *Advances in Health Sciences Education, 12*, 169-186. doi: 10.1007/s10459-005-5046-6
- Papinczak, T., Young, L., Groves, M., & Haynes, M. (2007b). An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Medical Teacher, 29*, e122-e132. doi: 10.1080/01421590701294323
- Rezler, A. G. (1989). Self-assessment in problem-based groups. *Medical Teacher, 11*(2), 151-156.
- Saris, W., & Stronkhorst, H. (1984). *Causal modeling in nonexperimental research*. Amsterdam, The Netherlands: Sociometric Research Foundation.
- Schönrock-Adema, J., Heijne-Penninga, M., Van Duijn, M. A. J., Geertsma, J., & Cohen-Schotanus, J. (2007). Assessment of professional behaviour in undergraduate medical education: Peer assessment enhances performance. *Medical Education, 41*(9), 836-842. doi: 10.1111/j.1365-2923.2007.02817.x

- Schwartz, R. W., Donnelly, M. B., Sloan, D. A., & Young, B. (1994). Knowledge gain in a problem-based surgery clerkship. *Academic Medicine*, 9(2), 148-151.
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. (2003). Cooperative learning and achievement: theory and research. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology Educational psychology, Vol. 7.* (pp. 177- 198). Hoboken (NJ): John Wiley & Sons Inc.
- Sluijsmans, D. M. A., Moerkerke, G., Van Merriënboer, J. J. G., & Dochy, F. J. R. C. (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation*, 27, 153-173.
- Sullivan, M. E., Hitchcock, M. A., & Dunnington, G. L. (1999). Peer and self-assessment during problem-based tutorials. *The American Journal of Surgery*, 177, 266-269.
- Tousignant, M., & DesMarchais, J. E. (2002). Accuracy of student self-assessment ability compared to their own performance in a problem-based learning medical program: A correlation study. *Advances in Health Sciences Education*, 7(1), 19-27. doi: 10.1023/A:1014516206120
- Van Berkel, H. J. M., & Van Til, C. T. (1994 Apr 15-17). *Group performing in a problem-based curriculum: the development and evaluation of an instrument.* Paper presented at the Paper presented at the Annual Meeting of the American Educational Research Association, San Diego (CA).
- Van der Linden, J., Erkens, G., Schmidt, H., & Renshaw, P. (2000). Collaborative learning. In R. J. Simons, J. Van der Linden, & T. Duffy (Eds.), *New Learning* (pp. 37-54). Dordrecht, The Netherlands: Kluwer.
- Visschers-Pleijers, A. J. S. F., Dolmans, D. H. J. M., De Leng, B. A., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2006). Analysis of verbal interactions in tutorial groups: A process study. *Medical Education*, 40, 129-137. doi: 10.1111/j.1365- 2929.2005.02368.x
- Visschers-Pleijers, A. J. S. F., Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2005). Student perspectives on learning-oriented interactions in the tutorial group. *Advances in Health Sciences Education*, 10(1), 23–35. doi: 10.1007/s10459-004-9348-x
- Yew, E. H. J., & Schmidt, H. G. (2009). Evidence for constructive, self-regulatory, and collaborative processes in problem-based learning. *Advances in Health Sciences Education*, 14, 251–273. doi: 10.1007/s10459-008-9105-7





# CHAPTER 3

## The Relationship Between Students' Small Group Activities, Time Spent on Self-Study, and Achievement.

Published as:

Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2012). The relationship between students' small group activities, time spent on self-study, and achievement. *Higher Education*, 64(3), 385-397. doi: 10.1007/s10734-011-9500-5



**ABSTRACT**

The purpose of this study was to investigate the relationship between the contributions students make to the problem-based tutorial group process as observed by their peers, self-study time, and achievement. To that end, the Maastricht Peer Activity Rating Scale was administered to students participating in problem-based learning tutorial groups. With this rating scale, students had to rate the constructive, collaborative, and motivational activities of their peers within the tutorial group. In addition, time spent on self-study was measured with a self-estimation method and achievement was measured with a unit test and a group assignment. A causal model of these variables was developed, in which the three types of activities were assumed to affect time spent on self-study, which would in turn affect unit test scores and group assignment scores. A structural equation modeling analysis indicated acceptable model fit. Especially apparent was the evidence for the causal relations between a student's constructive activities and his/her unit test score and between a student's collaborative activities and the group assignment score. On the other hand, time spent on self-study was not affected by the students' contributions, nor did it have an effect on the unit test score. These results suggest that there are indeed causal relations between a student's contributions to the tutorial group process and achievement.

## INTRODUCTION

Problem-based learning (PBL) is an instructional method originating from the active learning perspective, allocating the responsibility of learning on learners. PBL is also a constructivist and collaborative learning environment, meaning that people learn by giving meaning to experiences and interactions with others (Savery & Duffy, 1996). As its name indicates, this instructional method starts with an authentic problem that is relevant for the future practice of the profession. Students meet each other twice a week, during which they work together in small groups of approximately eight to 10 students. During the first meeting, students start with activating their prior knowledge about the given problem. Next, the problem is analyzed, possible explanations are generated and learning issues are formulated. After these first steps, students individually study these learning issues and search for relevant information. During the second meeting students discuss and synthesize their findings, which should lead to a deeper understanding of the problem. In short, PBL consists of three phases: problem analysis in the tutorial group, individual self-study, and (re)structuring of the newly acquired knowledge in collaboration with the other group members (Barrows, 1988; Schmidt, 1993). This whole process is facilitated by a tutor, who monitors and if necessary improves the quality of the discussion in the group. The quality of the discussion within the tutorial group seems to play a crucial role in the success of PBL (Gijsselaers & Schmidt, 1990; Savery & Duffy, 1996). Tutorial groups that are perceived as well functioning by the students within the group, achieve better (Gijsselaers & Schmidt, 1990). At the same time it is still unclear what the exact relationship is between students' contributions to the group discussion and achievement.

### Collaborative learning

PBL can be seen as a type of collaborative learning environment where students work together on a common task (Dolmans, De Grave, Wolphagen, & Van der Vleuten, 2005). According to this learning principle learning increases when students share common goals, are dependent on and accountable to each other (Johnson, Johnson, & Smith, 2007; Van der Linden, Erkens, Schmidt, & Rensaw, 2000). The factors that affect the effectiveness of different types of collaborative learning environments have been studied extensively over the years. In a recent study, Van den Bossche, Gijsselaers, Segers, and Kirschner (2006), developed a model of team

effectiveness. Participants in this study were enrolled in higher business education and worked together on an assignment for a period of seven weeks. According to this model team effectiveness is affected by both cognitive and social factors. The cognitive factors that contribute to the effectiveness are referred to as team learning behavior and include construction, co-construction, and constructive conflict. This implies that within a group different viewpoints are articulated and discussed, and a common understanding is negotiated (Van den Bossche et al., 2006). Teams in which these three types of cognitive behavior take place reported a higher shared cognition (i.e., “shared conception of the problem” (Van den Bossche et al., 2006, p. 492)) within the group. This, in turn, seemed to enhance the perceived team effectiveness. Furthermore, the occurrence of these team learning behaviours depends on different social factors. Students have to be committed towards the task, to experience a sense of interdependence, feel safe to express their ideas and belief that their group can work in an effective way (Van den Bossche et al., 2006). If these conditions are met team learning behaviour will more likely occur. Decuyper, Dochy, and Van den Bossche (2010) developed an even more extensive model of team learning. In this model they also demonstrated that shared cognition, and therefore successful collaborative learning, arises from co-construction and constructive conflicts within the group. But according to this model, as opposed to the model of Van den Bossche et al. (2006), the occurrence of co-construction and constructive conflicts is not only affected by (social) variables at the team level, but also by variables at the individual level. At this individual level they distinguish individual characteristics and individual behavior of team members as important input variables for successful collaborative learning. Examples of such individual variables are: motivation, appreciation of team work, construction, responsibility, and participation.

### **Effectiveness of PBL**

The effect of PBL on learning has been subjected extensively to study (e.g., Albanese & Mitchell, 1993; Colliver, 2000; Dochy, Segers, Van den Bossche, & Gijbels, 2003; Schmidt, Cohen-Schotanus, & Arends, 2009a), of which the majority has focused on medical education. Research on effectiveness of PBL has not so much focused on the influencing factors, but more on the comparison of PBL and conventional curricula. This particular learning environment seems to facilitate the development of a flexible knowledge base within students and the development of problem-solving skills and self-

directed learning skills in medical education (Hmelo-Silver, 2004). Dochy et al. (2003) concluded that students enrolled in medical PBL curricula are better at the application of knowledge. Recently, Schmidt, Van der Molen, Te Winkel, and Wijnen (2009b) also found that medical curricula that promote active learning through small group instruction have higher graduation rates and lower study duration, compared to the more conventional curricula that emphasize direct instruction. Another recent Dutch study of Schmidt et al. (2009a) summarized studies in which a single well-established PBL medical curriculum was compared to several conventional medical curricula. This meta-analysis indicated that students graduating from this single PBL curriculum have better interpersonal and practical medical skills than the students from the more conventional, teacher centered, curricula. Medical knowledge and diagnostic reasoning was also slightly better in the PBL curriculum, but these effects were only small. In addition, the drop-out rate was lower and the graduation time was shorter in this curriculum.

### **Group functioning**

Although PBL seems to have a beneficiary effect on student learning, it has also become evident that a high quality group discussion is conditional to the success of PBL (Cohen, 1994; De Grave, Boshuizen, & Schmidt, 1996; Savery & Duffy, 1996). Only within well-functioning PBL groups, cognitive conflicts (i.e., conflicting ideas within the group) will arise and new knowledge will be constructed and tested through negotiation, leading to true constructivist learning (Savery & Duffy, 1996). The quality of interactions in tutorial groups is perceived by students as an important predictor for the productivity of a tutorial group (Dolmans, Wolfhagen, & Van der Vleuten, 1998). This is in line with studies on collaborative learning environments other than PBL (e.g., Decuyper et al., 2010; Van den Bossche et al., 2006). Merely putting students together in a group with a common task does not automatically lead to successful learning. Within collaborative learning groups such as PBL tutorial groups, different factors can impede successful group functioning. In their model Decuyper et al. (2010) list the most mentioned factors in literature. One of these factors is that group members tend to take less responsibility because they can rely on their group members to do the work. This so-called 'free riding' has also been seen in PBL tutorial groups and students have indicated that they believe this has a negative effect on group functioning (Dolmans et al., 1998). The same is true for overly dominant students, because this

might restrain other group members from giving their opinion. Another problem in tutorial groups is that students sometimes avoid constructive conflicts (Moust, Van Berkel, & Schmidt, 2005). Students tend to merely state the undifferentiated main issues that resulted from their self-study, without explicitly discussing differences in viewpoint between students or sources. This closely resembles one of the impeding factors that Decuyper et al. (2010) called 'group thinking'.

Causal models have demonstrated to be a very suitable tool to further elucidate the relationship between group functioning and achievement. Back in 1990, Gijsselaers and Schmidt developed a causal model of learning in a PBL environment in which group functioning played a crucial role. In their model group functioning was influenced by the degree to which a student's prior knowledge linked up with the subject-matter, the quality of the problems and the tutor performance. Students in high functioning groups spent more time on self-study and consequently achieved better on a test. In addition, well-functioning tutorial groups had a positive influence on students' intrinsic interest in the subject-matter. Van Berkel and Schmidt (2000) added another variable to this model, namely group attendance which represented the commitment or willingness of students to engage in PBL. They found that attendance had a mediating effect between group functioning and achievement. Effective tutorial groups had higher attendances, which consequently led to higher achievements. In addition, the results indicated that the better the attendance is in a tutorial group, the less time for self-study is needed.

The weakness of these studies is that they treat group functioning as a single variable, leaving it unclear which specific student activities and group processes are conditional for effective group functioning and consequently for student achievement. In other words, these causal models do not differentiate, by indicating which student activities in the tutorial group are of central importance. Although the functioning of the tutorial group seems to steer learning, it remains unclear which student activities are crucial for good learning achievements (Dolmans & Schmidt, 2006; Hak & Maguire, 2000). Hak and Maguire (2000, p. 1) refer to it as "the black box of studies on PBL".

### **Student activities**

In order to unravel the relationship between tutorial student activities and achievement, we first have to take a look at what is known about students' tutorial behavior and tutorial group processes. According to Slavin, Hurley, & Chamberlain

(2003), there are three major perspectives on the effectiveness of learning in small groups. The first is the cognitive perspective, which comes down to the notion that students learn through interaction with peers and through elaboration (e.g., by explaining in own words or asking critical questions) on the subject matter. Observational studies have demonstrated that students in tutorial groups display these cognitive activities such as summarizing, asking each other critical questions and correcting misconceptions (Visschers-Pleijers, Dolmans, De Leng, Wolfhagen, & Van der Vleuten, 2006; Yew & Schmidt, 2009). Within a group these activities need to occur in order to develop shared cognition (Van den Bossche et al., 2006). The second perspective is the collaborative perspective, meaning that the degree of cohesiveness within a group is thought to have a positive influence on achievement. Van der Linden et al. (2000) state that, in order for collaborative learning to be successful, students working together within groups should be mutually dependent, should share a common goal and should share responsibilities. Yew and Schmidt (2009) identified these collaborative processes in tutorial groups, for instance in the form of sharing information. The third and last perspective on the effectiveness of learning in small groups is the motivational perspective. According to this perspective motivation to complete the learning task at hand prompts students to actively contribute to the group discussion (Slavin et al., 2003). Situational interest, a somewhat related type of motivation, is mentioned by Hidi and Renninger (2006). They refer to it as the willingness to engage, which springs from an interaction between a person and a specific content or even a specific learning environment. Hidi and Renninger (2006) suggest that academic achievement is positively affected by situational interest of students.

### **Self-study time**

Since PBL intends learning to be self-directed (Dolmans et al., 2005), time spent on self-study also takes up an important position within the PBL process. Based on the model of Gijsselaers and Schmidt (1990), which was discussed earlier, time spent on self-study is thought to play a mediating role between group functioning and student achievement. Van den Hurk, Wolfhagen, Dolmans, & Van der Vleuten (1999) found that students in PBL tutorial groups were of the opinion that more useful learning issues were formulated when they participated in a well-functioning group. The same students also indicated that they used these learning issues as a starting

point for self-study. The relation between self-study time and achievement has been extensively investigated. In their review, Frederick and Wahlberg (1980) found that time spent on learning is positively related to achievement, but they also argue that this relation is influenced by the quality of instruction and student ability. Van den Hurk, Wolfhagen, Dolmans, and Van der Vleuten (1998) also suggest that there is a relation between time spent on self-study and achievement but that it may not be as straightforward as one would expect. They investigated the relation between time spent on self-study and achievement in a medical PBL curriculum. The results indicated that there was no significant relation between self-study time and achievement scores for the first and second year students, but that further research is needed into the influence of different qualitative factors. For instance, there is no insight in the influence of group processes within the tutorial group on self-study time.

### **Peer evaluations**

Research on student performance in PBL tutorials is often based on students' self-evaluation and the validity and reliability of self-evaluation is not optimal (e.g., Eva, 2001). Yew and Schmidt (2009) also argued that data gathered from observations instead of self-evaluation is preferable, because it allows you to be as close to the learning activities as possible. In this study we tried to get an even closer look at the contributions of students, by using peer observations. Peers might have a more accurate picture of how students perform during tutorial meetings. There are two possible explanations for this (Eva, 2001). First, students see their peers intensively (twice a week for period of eight weeks) and second, each student is evaluated by multiple peers. Therefore one judgment about an individual student consists of multiple evaluations.

In a previous study (Kamp, Dolmans, Van Berkel, & Schmidt, 2011) the Maastricht Peer Activity Rating Scale (M-PARS) was developed with which students can evaluate three main aspects of tutorial peer activity: a peer's constructive, collaborative, and motivational contributions to the tutorial group. This study also demonstrated that, students are able to evaluate their peers on these three aspects in a reliable and valid way and that, only four evaluations per student are needed for a reliable judgment.

**The current study**

Although it can be concluded from the previous paragraph that group functioning is affected by cognitive, collaborative, and motivational activities, it is still unknown how these activities relate to student achievement. In this study we will, therefore, focus on specific student behavior that positively influences student achievement. The mediating effect of self-study time on achievement was already mentioned and therefore it will also be taken into account in this study. The central research question that is addressed in this study is: ‘Do students who perform better in tutorial groups (i.e., display more constructive, collaborative, and motivational activities according to their peers) also spend more time on self-study and subsequently achieve better or, in other words, does tutorial performance have any predictive value with regard to self-study time and subsequent achievement?’ Based on the causal model of Gijsselaers and Schmidt (1990) it is hypothesized that the student’s activities have an indirect effect on achievement on the unit test and the group assignment and that this effect is mediated by the time students spend on self-study (Fig. 1).

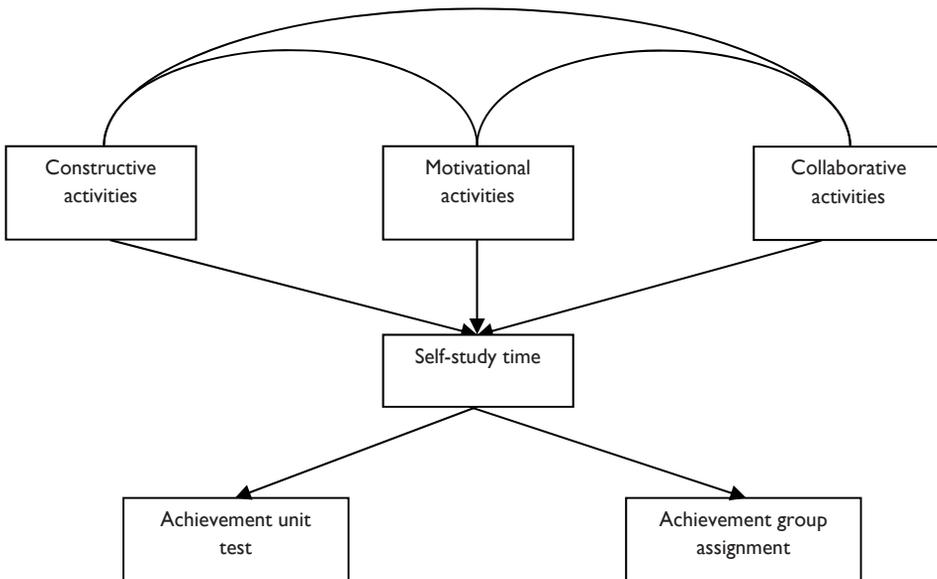


Figure 1. Hypothesized model of the current study.

## METHOD

### Participants

Participants were 650 first- and second-year students attending the medical curriculum (PBL) at Maastricht University during the academic year 2009–2010. These students were divided into 62 tutorial groups, each group consisting of approximately 10 students. Students met each other twice a week for a 2-h session over a period of six (first year students) to 10 (second year students) weeks. During these meetings students analyzed the problem that was presented to them and discussed and restructured the information they gathered during self-study collaboratively.

### Instruments

Students' activities were measured with the Maastricht Peer Activity Rating Scale (M-PARS) (Kamp et al., 2011). This 14-item rating scale consists of three subscales (See Appendix A). The first subscale (5 items), constructive activities, measures activities that promote the development of shared cognition (e.g., This student corrected misconceptions about the subject matter). The second subscale (5 items), collaborative activities, measures activities that contribute to the social aspect of collaboration (e.g., This student promoted collaboration between group members). The last subscale (4 items), motivational activities, measures activities that give the impression the student is motivated for the group work (e.g., This student demonstrated to be motivated). With this scale students can evaluate the performance of their peers in the tutorial group by responding to these items on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Students had to also rate the overall performance of their peers in the tutorial group on a scale from 1 (very poor) to 10 (excellent). Students were asked to fill out the M-PARS during the last week of the unit. Scores on the three types of activities were obtained by averaging all ratings from the same student, resulting in an average score (ranging from 1 to 5) per type of activity.

Student achievement was measured in twofold with a group assignment and a unit test at the end of the unit. For the group assignment students had to collaboratively prepare and give a presentation about a unit content related subject to their peers. The unit test consisted of 75 (for first year students) and 125 (for second year students) true-false questions which corresponded with the content of the unit.

Both the grade of the test as well as the grade of the assignment was transformed to a scale ranging from 1 to 10. With regard to the final unit grade, the grade on the test had a higher weight (80%) than the grade on the group assignment (20%).

Self-study time was measured at the end of the course, before the unit test. Students were asked to estimate the average number of hours per week they spent on self-study during this unit. Moust (1993) already proved the validity of this method. He found a strong correlation between this post hoc self-estimation and a log method, were students had to systematically record the time they had spent on self-study.

### **Procedure**

At the beginning of the course students were informed about the procedure of this study. The M-PARS and the self-study time question were administered to the students during a tutorial meeting in the last week of the course. Students received an envelope with the self-study time question and approximately nine rating scales, one to be completed for each group member. Students were ensured that the data would be processed in a confidential manner and would not be reported back to their peers. The rating scales and the self-study time question were collected a couple days later during the next tutorial meeting. Twenty gift certificates were raffled among the participants who completed all items.

Since the aim of this study was to investigate the causal relations between student activities, self-study time, and achievement, a structural equation modeling (SEM) analysis was undertaken using the AMOS statistical program (Version 17.0). SEM is a widely accepted statistical technique with which a theoretical model can be confirmed. As long as it is used for confirmatory use and not exploratory and as long as the model is based on substantial theory, it can be used to test the hypothesized underlying structural and causal relationships (see Violato & Hecker, 2007). However, as Violata and Hecker also indicate, SEM can only provide evidence and not proof for causality, and it can never rule out any explanatory factors. Model fit will be determined using the following fit indices: Chi-square (CMIN), RMSEA, and NFI. Good fit is indicated by  $CMIN/df < 3$  (CMIN divided by the degrees of freedom) with a significant  $p$ -value,  $RMSEA \leq .05$ , and  $NFI > .95$  (Garson, 2009).

## RESULTS

A total of 375 students completed the rating scales. This is a response rate of 61%. A total of 538 students were evaluated by four or more of their peers in the tutorial group. Inspection of the histograms, the skewness, and the kurtosis of the variables, led to the removal of five outliers. This resulted in a normal distribution for all variables (data may be assumed to be normal if skew and kurtosis is within the range of  $\pm 1.5$  (Schumacker & Lomax, 2004, p. 69)). Multivariate normality was checked by inspecting the Mahalanobis distance values. This led to the removal of six more outliers, resulting in a sample size of 527. Table I shows the descriptives of the variables involved in this study. This table illustrates that the average achievement scores range from 3.8 to 9.5 (scale 1–10), with standard deviations of 0.72 and 0.89. Time spent on self-study varies between one and 36 h, with a standard deviation of 5.96. The mean scores on the three factors of the M-PARS range from 1.6 to 4.9, with standard deviations varying between 0.51 and 0.60. As can be seen in this table, skewness and kurtosis are within the acceptable range.

Since the purpose of this study was to test the relations between the three types of activities, time spent on self-study, and achievement the model depicted in Figure 1 was tested in AMOS, using maximum likelihood as estimation method. The fit indices of the initial model (Fig. 1) indicated a poor fit with the data. The modification indices for the hypothesized relationship between constructive activities and achievement on the unit test, between collaborative activities and achievement on the group assignment, and between achievement on the unit test and achievement on the group assignment were high, indicating that adding these three relationships would significantly improve model fit. Therefore, these three pathways were added. The pathway between self-study time and achievement on the group assignment was dropped because it did not contribute to a better fit of the model. This resulted in the model in Figure 2, which illustrates a significant relation between a student's constructive activities and his/her score on the unit test, between a student's collaborative activities and his/her score on the group assignment, and between a student's score on the group assignment and his/her score on the unit test.

Table I

*Descriptives of the Study Variables*

Variable	N	Mean	Standard deviation	Range	Skew	Kurtosis
Unit test	527	7.10	0.89	3.8-9.2	-0.47	0.41
Group assignment	527	7.23	0.72	5.0-9.5	-0.11	1.00
Time for self-study hours/week	527	13.81	5.96	1.0-36.0	0.95	1.51
Constructive activities (5 items) item mean	527	3.57	0.51	2.0-4.9	-0.49	0.11
Collaborative activities (5 items) item mean	527	3.68	0.52	1.9-4.9	-0.79	0.84
Motivational activities (4 items) item mean	527	3.62	0.60	1.6-4.9	-0.60	0.27
Total score M-PARS (14 items)	527	50.74	7.23	26.7-67.3	-0.66	0.41

### Fit indices

The model depicted in Figure 2 generated the following fit indices in AMOS: Chi-square [ $5df$ ] = 5.167,  $p = 0.396$ , a root mean square error of approximation of 0.008 and a normed fit index of 0.997 (Table 2). These results suggest good model fit. In order to test the stability of the model, the model was cross-validated by splitting the data set into two random sub-sets. The fit indices of both data sub-sets are also shown in Table 2 and are similar to the fit indices of the total data-set.

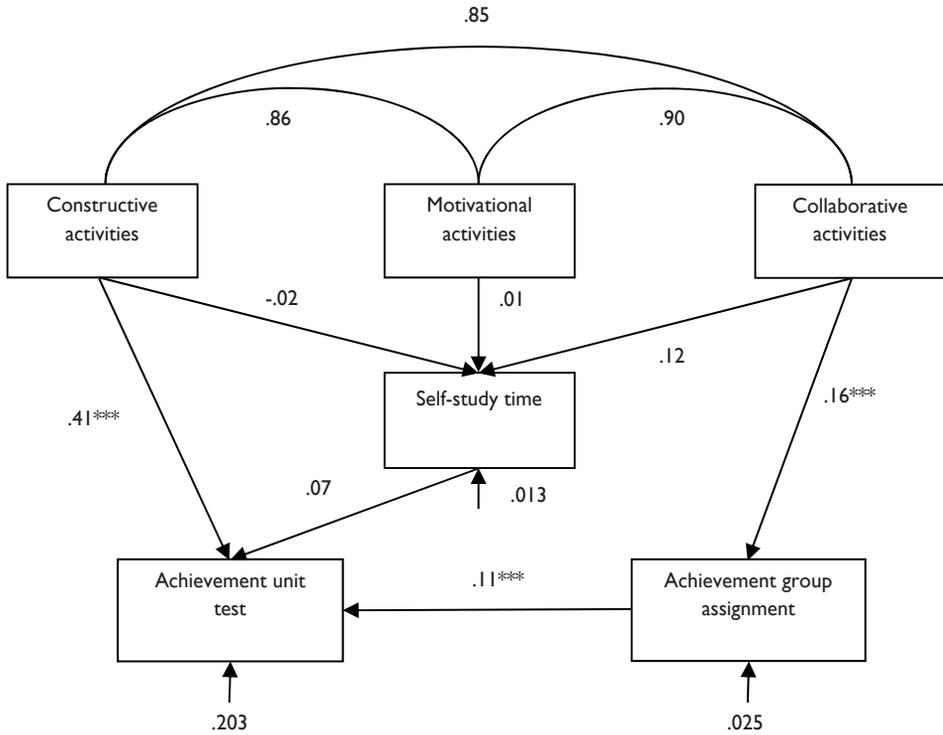


Figure 2. Path diagram with standardized regression weights and percentages of explained variance. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

Table 2

Fit Indices for the Total Data Set and for the Two Random Sub-Sets

Data-set	N	$\chi^2$	df	p	Cmin/df	RMSEA	NFI
Total	527	5.167	5	0.396	1.033	0.008	0.997
Sub-set 1	263	3.609	5	0.699	0.602	0.000	0.997
Sub-set 2	264	7.047	5	0.217	1.409	0.040	0.992

Note. RMSEA = Root mean square error of approximation; NFI = normed fit index. Good fit is indicated by CMIN/DF < 3 (CMIN divided by the degrees of freedom) with a significant p-value, RMSEA  $\leq$  .50, and NFI > .95 (Garson, 2009).

## DISCUSSION

The purpose of this study was to investigate the relationship between the contributions students make to the tutorial group process observed by their peers, self-study time, and achievement. Therefore, the M-PARS was administered to a large group of first- and second-year students attending a bachelor in medicine. With this rating scale students had to evaluate the constructive, collaborative, and motivational contributions of their peers during the tutorial meetings. In a theoretically substantiated model the relation between these three activities, average hours of self-study time per week, and achievement scores were studied. It was hypothesized that the three types of activities would have an indirect effect on the achievement scores of the students. This effect was expected to be mediated by time spent on self-study. Based on the analyses and results there is evidence for the relationships between the three types of activities and self-study time and achievement. This implies there is not only a relationship between social and cognitive factors and successful collaborative learning on the team level (e.g., Van den Bossche et al., 2006), but also on the individual level between individual contributions and individual achievement. This is in line with the model on team learning by Decuyper et al. (2010), which already emphasized the importance of individual behavior in groups in order for collaborative learning to be successful. This study indicated that within groups these inputs at the individual level, may not only promote shared cognition but also affect the individual achievement of a student while learning within a group. More specifically, the results indicated a strong relationship between a student's constructive contributions within the tutorial group and the achievement score on the unit test. In other words, this implies that students who, according to their peers, ask critical questions, are able to correct misconceptions, and can distinguish main from lateral issues, achieve better on the unit test than their peers who make less constructive contributions within the tutorial group. On the group level, Van den Bossche et al. (2006) already demonstrated a similar relationship; teams who display more learning behaviors (e.g., construction) experienced an increased sense of shared cognition and perceived themselves as more effective. In addition, the results indicated an effect of a student's collaborative contributions within the group on their achievement score of the group assignment. Students who are perceived by their peers as more committed to the group and promote collaboration within the group achieve better on the group

assignment than students who are perceived by their peers as less committed to the group. This also confirms previous research which showed that successful group learning is, in addition to constructive learning behavior, affected by social factors within the group (Johnson et al., 2007; Van den Bossche et al., 2006). There also was a small effect of achievement on the group assignment on achievement on the unit test. When one assumes that the achievement score represents the extent to which students have learned about the subject matter, it is plausible that a good score on the group assignment will predict a good score on the following unit test. One could argue that the effect of group assignment achievement on unit test achievement is expected to be larger, but it is also very plausible that both tests address different aspects. The unit test is expected to rather measure factual knowledge and the group assignment more collaborative, social, and presentation skills. What was indeed surprising was that there was no mediating effect of time spent on self-study. Self-study time appeared not to be affected by the contributions students made in the tutorial group, and it showed no effect on the achievement scores of students. A possible explanation for the absence of this effect can be found in an article by Plant, Ericsson, Hill, and Asberg (2005). They argue that learning achievements are not so much influenced by the amount of time a student spends on self-study, but by the manner in which he or she spends this time. Students can spend a lot of time on individual study activities but if they do not use this time effectively it will not represent the extent to which they have learned about the subject matter. Van den Hurk et al. (1998) also suggest that achievement is not so much influenced by the amount of self-study time, but more by the manner in which they spend this time and the individual learning needs of students.

Although the results provide evidence for the causal relationships between the variables, it is impossible to rule out any underlying exploratory factors which have influenced these relationships. One of these underlying factors might be that students who are active in the PBL tutorial also achieve better because they have a more engaged and motivated attitude towards their education and as a consequence invest more into their education.

What would be the implications of these findings for future research in area of PBL? Since a relation was found between a student's contributions in the tutorial group and his or her achievement scores, it is important to monitor the quality of the contributions each individual student makes during tutorial group meetings. To that end, it is worthwhile to investigate whether or not the M-PARS could be used as a

feedback tool for student's tutorial behavior. This tool could then be used by students to provide peer feedback and subsequently improve their tutorial group contributions. As a next step it would be interesting to investigate whether or not this has a positive effect on their achievement. But since we have argued that there might be a discrepancy between what is addressed in the tutorial group and what is measured with the test, it would be a good idea to look for other methods to measure learning outcomes. Another interesting idea for future research is to repeat this study on the level of the whole tutorial group. Since the success of PBL depends on the quality of the collaborative process within the tutorial groups and since the behavior of one group member is heavily influenced by the behavior of other group members, it would be of added value to discover if well contributing groups achieve better than poor contributing groups.

A first limitation of this study is that the relations between the variables in the model are not very strong (with the exception of the one between a student's constructive contributions and the score on the unit test) and, consequently, that the model only explains a small percentage of the variance of the variables. There are two possible extenuating explanations for this result. First of all, the relation between the discussion in the tutorial group and the unit test or group assignment might not be as straightforward as is expected. A test might not always completely agree with what was discussed. In addition, the unit test in these two units measure mainly factual knowledge and might therefore be less suitable for the detection of deep learning which is the intended and expected effect of PBL. In other words, there might be a discrepancy between what is addressed in the tutorial group and what is measured with the test. Second of all, the population of medical students in the Netherlands is highly homogeneous. The greater part of the students have had the same preparatory training in secondary school and have an above average mean grade for their final exams in secondary school. In addition the students are, on the whole, highly motivated. Because of these two characteristics, the variances between our participants might be small. Therefore it can be perceived as a positive result that, despite the small differences within the population, a fitting model was found.

Another limitation is that the score on the group assignment in this study is a group score, meaning that there was no individual grade for each individual group member, but instead that all four group members received the same grade. Students

who contributed poorly to this assignment could, therefore, have received a higher grade than they would have if their individual contributions would have been graded.

A last limitation of this study is the high correlation between the three types of activities measured with the M-PARS. These correlations could indicate that we are dealing with only one type of student tutorial activity. However, a confirmatory factor analysis in a previous study (Kamp et al., 2011) has proved that there are indeed three different factors with regard to the three types of tutorial student activity.

Furthermore, in the model one factor (constructive activities) predicts the score on the multiple choice unit test and the other factor (collaborative activities) predicts the score on the group assignment.

## REFERENCES

- Barrows, H. S. (1988). *The tutorial process*. Springfield, Illinois: Southern Illinois University School of Medicine.
- Cohen, E. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64, 1-35.
- Colliver, J. A. (2000). Effectiveness of problem-based learning curricula: Research and theory. *Academic Medicine*, 75(3), 259-266.
- De Grave, W. S., Boshuizen, H. P. A., & Schmidt, H. G. (1996). Problem based learning: Cognitive and metacognitive processes during problem analysis. *Instructional Science*, 24(5), 321-341.
- Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning: An integrative model for effective team learning in organizations. *Educational Research Review*, 5, 111-133. doi: 10.1016/j.edurev.2010.02.002
- Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, 13, 533-568. doi: 10.1016/S0959-4752(02)00025-7
- Dolmans, D. H. J. M., De Grave, W., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2005). Problem-based learning: Future challenges for educational practice and research. *Medical Education*, 39, 732-741. doi: 10.1111/j.1365-2929.2005.02205.x
- Dolmans, D. H. J. M., & Schmidt, H. G. (2006). What do we know about cognitive and motivational effects of small group tutorials in problem-based learning? *Advances in Health Sciences Education*, 11(4), 321-336. doi: 10.1007/s10459-006-9012-8
- Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. (1998). Motivational and cognitive processes influencing tutorial groups. *Academic Medicine*, 73(10 Suppl), S22-S24.
- Eva, K. W. (2001). Assessing tutorial-based assessment. *Advances in Health Sciences Education*, 6(3), 243-257.
- Frederick, W. C., & Wahlberg, H. J. (1980). Learning as a function of time. *The Journal of Educational Research*, 73(4), 183-194.

- Garson, G. D. (2009). *Structural equation modeling, from statnotes: Topics in multivariate analysis* [online]. Retrieved August 18, 2010 from <http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm>.
- Gijsselaers, W. H., & Schmidt, H. G. (1990). Development and evaluation of a causal model of problem-based learning. In Z. Nooman, H. G. Schmidt, & E. Ezzat (Eds.), *Innovation in medical education: An evaluation of its present status* (pp. 95-113). New York: Springer.
- Hak, T., & Maguire, P. (2000). Group process: The black box of studies on problem-based learning. *Academic Medicine, 75*(7), 769-772.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127. doi: 10.1207/s15326985ep4102\_4
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235-266. doi: 10.1023/B:EDPR.0000034022.16470.f3
- Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review, 19*, 15-29. doi: 10.1007/s10648-006-9038-8
- Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2011). Can students adequately evaluate the activities of their peers in PBL? *Medical Teacher, 33*(2), 145-150. doi: 10.3109/0142159X.2010.509766
- Moust, J. H. C. (1993). *De rol van tutoren in probleemgestuurd onderwijs: contrasten tussen student- en docent-tutoren* [The role of tutors in problem-based learning: Contrasts between undergraduate teaching assistants and faculty]. Unpublished Doctoral dissertation, University of Limburg, Maastricht, The Netherlands.
- Moust, J. H. C., Van Berkel, H. J. M., & Schmidt, H. G. (2005). Signs of erosion: Reflections on three decades of problem-based learning at Maastricht University. *Higher Education, 50*, 665-683. doi: 10.1007/s10734-004-6371-z
- Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30*, 96-116. doi: 10.1016/j.cedpsych.2004.06.001
- Savery, J. R., & Duffy, T. M. (1996). *Problem based learning: An instructional model and its constructivist framework*. In B. G. Wilson (Ed.), *Constructivist learning*

- environments: Case studies in instructional design (pp. 135-148). Englewood Cliffs, NJ: Educational Technology Publications.
- Schmidt, H. G. (1993). Foundations of problem-based learning: Some explanatory notes. *Medical Education*, 27, 422-432.
- Schmidt, H. G., Cohen-Schotanus, J., & Arends, L. R. (2009a). Impact of problem-based, active learning on graduation rates for 10 generations of Dutch medical students. *Medical Education*, 43, 211-218. doi: 10.1111/j.1365-2923.2008.03287.x.
- Schmidt, H. G., Van der Molen, H. T., Te Winkel, W. W. R., & Wijnen, W. H. F. W. (2009b). Constructivist, problem-based learning does work: A meta-analysis of curricular comparisons involving a single medical school. *Educational Psychologist*, 44(4), 227-249. doi: 10.1080/00461520903213592
- Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. M. (2003). *Cooperative learning and achievement: Theory and research*. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology* (Vol. 7, pp. 177-198). Hoboken, NJ: Wiley.
- Van Berkel, H., & Schmidt, H. G. (2000). Motivation to commit oneself as a determinant of achievement in problem-based learning. *Higher Education*, 40, 231-242.
- Van den Bossche, P., Gijselaers, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments: Team learning beliefs and behaviors. *Small Group Learning*, 37(5), 490-521. doi: 10.1177/1046496406292938
- Van den Hurk, M. M., Wolfhagen, I. H. A. P., Dolmans, D. H. J. M., & Van der Vleuten, C. P. M. (1998). The relation between time spent on individual study and academic achievement in a problem-based curriculum. *Advances in Health Sciences Education*, 3, 43-49.
- Van den Hurk, M., Wolfhagen, I. H. A. P., Dolmans, D. H. J. M., & Van der Vleuten, C. P. M. (1999). Student generated learning issues: A guide for individual study? *Education for Health*, 12(2), 213-221.
- Van der Linden, J., Erkens, G., Schmidt, H., & Renshaw, P. (2000). *Collaborative learning*. In P. R. J. Simons, J. L. Van der Linden, & T. M. Duffy (Eds.), *New learning* (pp. 37-55). Dordrecht: Kluwer.

- Violato, C., & Hecker, K. G. (2007). How to use structural equation modeling in medical education research: A brief guide. *Teaching and Learning in Medicine, 19*(4), 362-371. doi: 10.1080/10401330701542685
- Visschers-Pleijers, A. J., Dolmans, D. H., De Leng, B. A., Wolfhagen, I. H., & Van der Vleuten, C. P. (2006). Analysis of verbal interactions in tutorial groups: A process study. *Medical Education, 40*(2), 129-137. doi: 10.1111/j.1365-2929.2005.02368.x
- Yew, E. J. H., & Schmidt, H. G. (2009). Evidence for constructive, self-regulatory, and collaborative processes in problem-based learning. *Advances in Health Sciences Education, 14*(2), 251-273. doi: 10.1007/s10459-008-9105-7



# CHAPTER 4

## The Effect of Midterm Peer Feedback on Student Functioning in Problem-Based Tutorials.

Published as:

Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2012). The effect of midterm peer feedback on student functioning in problem-based tutorials. *Advances in Health Sciences Education*, 18 (2), 199-213. doi:10.1007/s10459-012-9364-1



**ABSTRACT**

Within Problem-Based Learning successful learning depends on the quality of cognitive, social, and motivational contributions students make to the tutorial group. But at the same time, not all students in PBL automatically contribute in a high quality manner, which might impede successful group functioning. This study investigated whether peer process feedback combined with goal setting can be used to improve the quality of students' individual contributions. A mixed-methods explanatory design, in which 74 second-year Health Sciences students participated, combined a pre- and posttest with a focus group. The results indicated that the quality of the contributions only increased for students with a below average score on the pre-test. The qualitative data confirmed that the impact of the feedback could be increased by combining individual reflection by means of goal setting with face-to-face discussion. Another suggestion is to investigate whether midterm peer process feedback is more effective for first year students, because they are still developing their tutorial behavior, as opposed to second-year students.

## INTRODUCTION

Within health sciences Problem-Based Learning (PBL) is a commonly used instructional approach. The success of PBL largely depends on the quality of tutorial group functioning (Savery & Duffy, 1996). Well-functioning PBL tutorial groups are groups in which different viewpoints and ideas are articulated and discussed and new knowledge is constructed collaboratively (Visschers-Pleijers, Dolmans, De Leng, Wolfhagen, & Van der Vleuten, 2006). In addition, the quality of functioning within the group is affected by the social climate within the group (Van den Bossche, Gijssels, Segers, & Kirschner, 2006). But, because group functioning and individual contributions are largely interdependent (Lee & Roth, 2007), well-functioning groups can only exist by the grace of well contributing individuals within the group. In a prior study, a positive relationship was indeed found between the quality of students' individual contributions to the problem-based tutorial group and achievement (Kamp, Dolmans, Van Berkel, & Schmidt, 2012). However, at the same time not all students within a PBL tutorial group automatically contribute in a high quality manner, which might impede successful group functioning. One common problem, for instance, within collaborative groups is that students tend to shirk their responsibility to other group members and rely on them to do the work (Decuyper, Dochy, & Van den Bossche, 2010). This so-called 'free riding' has also been seen in PBL tutorial groups and students have indicated that they believe this has a negative effect on group functioning (Dolmans, Wolfhagen, & Van der Vleuten, 1998). Another common problem is the overly dominant student, because this might restrain other group members from giving their opinion. Furthermore, a problem in tutorial groups is that students sometimes avoid constructive conflicts (Moust, Van Berkel, & Schmidt, 2005). Students tend to merely state the undifferentiated main issues that resulted from their self-study, without explicitly discussing differences in viewpoints between students or sources. This is also called 'group thinking' and is known to have a negative effect on group functioning (Decuyper et al., 2010). In order to overcome these problems, attention should be paid to the quality of individual students' contributions in the tutorial group. The tutor plays an important role in promoting students' learning in PBL (e.g., Dolmans, Wolfhagen, Scherpbier, & Van der Vleuten, 2003; Dolmans & Wolfhagen, 2005). He or she should stimulate students to elaborate (e.g., summarize in own words and search for explanations and contradictions) and to interact with each other. The tutor should

set a positive atmosphere within the group, which will enhance collaborative learning. Dolmans and Wolfhagen (2005) investigated the interaction between tutor performance, group productivity, and learning effectiveness as perceived by students. They found that tutor performance correlated with both group productivity and learning effectiveness. In addition, formative peer feedback might be used to improve the quality of a student's individual contributions (Dominick, Reilly, & McGourty, 1997). Especially because students seem to be better able to evaluate their peers' behaviour as opposed to their own (Eva & Regehr, 2011). Therefore, it would be interesting to investigate whether and how formative peer feedback in a PBL tutorial group setting can improve a student's contributions in the tutorial group.

According to Hattie and Timperley (2007) (peer) feedback is an important tool to promote learning. Its goal is to improve the performance of the receiver by closing the gap between the current performance of a student and a more positive and desired state (Archer, 2010). With regard to its effects on performance, both Hattie and Timperley (2007) and Kluger and DeNisi (1996) conclude that in general feedback seems to have a positive effect on student performance. If students are given the opportunity to revise their performance after receiving peer feedback this seems to positively influence their domain specific performance. Adcroft (2011) points out two possible explanations for the positive effect of feedback on performance. First, feedback demonstrates what is understood by 'good' behavior, and second, feedback can be used to diagnose the gap between a student's current behavior and the desired behavior. According to Hattie and Timperley (2007) feedback can address four different levels: task-, process-, self-regulation-, and self-level. Process level feedback addresses "the main processes needed to understand/perform the tasks" (p. 87) and is thought to enhance deep learning. Feedback on students' contributions to the group processes within PBL tutorial groups can be regarded as process level feedback. Most feedback research has focused on the effect of task feedback and only little attention is paid to the effects of process feedback, especially process peer feedback to individuals within groups (Geister, Konradt, & Hertel, 2006). The scarce research that has been done has demonstrated that peer process feedback has a positive effect on motivation, behavior, communication, and collaboration of group members (Dominick et al., 1997; Geister et al., 2006). It is, however, surprising that research on the effects of peer process feedback is scarce, because it is known from previous studies that students can evaluate their peers in a reliable and valid manner (Eva & Regehr, 2011; Papinzcak,

Young, & Groves, 2007a; Sluijsmans, Moerkerke, Dochy, & Van Merriënboer, 2001). Specifically within PBL tutorial groups, a previous study investigated whether students are able to rate the individual group member's contributions in a reliable and valid way (Kamp, Dolmans, Van Berkel, & Schmidt, 2011). The results indicated that students were able to distinguish and rate three types of student activities within the tutorial group, namely cognitive, collaborative, and motivational activities. Cognitive activities are contributions that lead to the construction of knowledge and constructive conflicts (e.g., summarizing and correcting misconceptions). Collaborative activities contribute to the achievement of a shared learning goal (e.g., willingness to share information, being committed to the group). Motivational activities are activities from which a student appears to be motivated for group work (e.g., amount of contributions, participation).

But as promising as peer feedback seems to be, it is also evident that it is not effective by definition (Topping, 1996). Instead it depends on a complex configuration of multiple variables (Topping, 2010). Although solid experimental proof is often lacking (Gielen, Peeters, Dochy, Onghena, & Struyven, 2010; Prins, Sluijsmans, & Kirschner, 2006), several authors have tried to identify characteristics that affect the effectiveness of peer feedback. Gielen et al. (2010), for instance, investigated the effects of five constructiveness characteristics (relevance and specificity of the peer feedback, presence of justification, suggestion for improvement, and clear formulation) and of the accuracy of the negative comments on subsequent performance on a writing task. The results indicated that only the presence of justification (i.e., explanation of judgment) positively affected performance, but only for students with lower pretest scores. The presence of accurate negative comments also had a positive effect on performance, but this was inferior to the effect of justification. Sluijsmans et al. (2001) and Van Zundert, Sluijsmans, and Van Merriënboer (2010) concluded that, even though students in higher education seem to be able to assess their own peers, peer assessment and feedback should be combined with training or that students should at least have experience with peer assessment. In the study by Sluijsmans et al. (2001), students were asked to rate their peers in a PBL context, students reported that they felt a need for a training prior to making judgments, especially if those judgments were negative. In general, students who have received training before they had to assess their peers seemed to be better able to assess their peers and had higher performances on subsequent tests (Sluijsmans, Brand-Gruwel, & Van Merriënboer,

2002). Another important aspect is to explain the purpose and goal of the feedback to the students before feedback is given and received. Hattie and Timperley (2007) call this *feed up* and argue that it is conditional to effective *feedback*. Archer (2010) agrees with this and states that, in general, effectiveness is increased when students recognize the relevance of the feedback. He also argues that it is important that the recipients agree with the received feedback. However, after the feedback is received, students need help to improve their future performance, which is called *feed forward* (Hattie & Timperley, 2007). In order to improve their future performance students should be stimulated to reflect and act upon the received feedback (Gibbs & Simpson, 2004). Prins et al. (2006) indicated that just providing feedback is not enough, but it should be accompanied by reflection on the feedback. Therefore the feedback should provide clues for behavioral change and focus on future functioning and goal setting to enhance the impact of feedback (Archer, 2010; Hattie & Timperley, 2007; Prins et al., 2006). This also implies that feedback should be facilitative instead of directive and should stimulate students' self-monitoring during the execution of a task. A last important point is that effectiveness of peer feedback might also depend on the competence level of the receiver. In the aforementioned study by Gielen et al. (2010) for instance, the effect of certain characteristics of feedback seemed to be smaller for students with a higher than average pretest score. Therefore, it might be that peer feedback is especially effective with a low to average competence level.

One of the scarce studies on the effectiveness of peer process feedback is the study by Phielix, Prins, and Kirschner (2010), who investigated the effects of peer process feedback in a computer simulated collaborative learning (CSCL) environment. They concluded that a combination of peer process feedback and reflection increases awareness of cognitive and social behavior. This, in turn, can enhance the quality of contributions students make to the social climate within the group (Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011). So, although these studies on the effects of process peer feedback combined with reflection in a CSCL environment showed promising results, studies on the effects of peer process feedback in PBL are limited. A few studies have been done that are described in the next paragraph.

First, Papinczak, Young, Groves, and Haynes (2007b) and Schönrock-Adema, Heijne-Penninga, Van Duijn, Geertsma, and Cohen-Schotanus (2007) investigated the attitudes students have towards peer process feedback. Students indicated that they

gained insight into, and developed an enhanced awareness of the criteria of good tutorial performance after they had provided their peers with feedback. Students who evaluated their group members also indicated that they felt an increased sense of responsibility and that they believed it had improved their learning performance (Schönrock-Adema et al., 2007).

Zumbach, Muehlenbrock, Jansen, Reimann, and Hoppe (2002) performed a study on the effectiveness of feedback on students' individual contributions in computer supported PBL tutorial groups. They provided group members with a quantitative display of their motivation and contributions. The experimental group, compared to the control group who did not receive any feedback, showed an increased amount of contributions. In a later study Zumbach, Hillers, and Reimann (2004) performed a similar study on feedback in distributed online PBL tutorial groups. Here they combined two types of feedback: qualitative feedback on group members' problem-solving strategies, and quantitative feedback on their contributions and their motivation. The first type of feedback appeared to result in better problem solutions, higher grades on knowledge tests, more contributions, and a higher degree of reflection. The second type of feedback was especially advantageous for students' motivation and their attitudes towards the course. Although both of Zumbach's studies did not use peers as a source for their feedback, they support the suggestion that process feedback on students' contributions within PBL tutorial groups has beneficiary effects.

Summarizing, it may be said that although students seem to be favorable towards peer process feedback and that (non peer) feedback on students' contributions also seems to have beneficiary effects on their behavior within computer supported PBL groups, it remains unclear if peer feedback on students' contributions can enhance behavior in a face-to-face tutorial PBL group.

This results in the following research questions:

1. Does midterm process peer feedback improve the quality of individual contributions to the PBL tutorial group?
2. Does this differ for students with different levels of functioning?
3. How do students perceive the effectiveness of process peer feedback and what are points of improvement as perceived by the students?

## METHOD

A mixed-methods explanatory design was used to address the research questions. In an explanatory design the quantitative study is dominant and is followed by the qualitative study. The qualitative study serves as an explanation and refinement of the quantitative results (Creswell & Plano Clark, 2007).

### Participants

Participants were 87 second-year students who attended a PBL bachelor program in Health Sciences at Maastricht University in The Netherlands. In this program students learn about the many factors that either promote or harm health. The intervention took place in the course 'Learning, Cognition, and Personality', which was required for all students. Participation in the feedback intervention was voluntary. During this eight-week course students met each other once a week for a 2-h tutorial group meeting. The students were divided over nine tutorial groups, each consisting of approximately ten students. In these tutorial groups students discuss a given problem and generate learning issues for further self-study. During the next tutorial, students discuss and synthesize their findings. All groups were guided by nine different tutors. During the first three weeks of the first-year of their bachelor program all students received a PBL training 'learning and communicating in tutorial groups'. In this training, students are taught, among other things, how to give feedback to one another within the tutorial group. The next four weeks this remains a point of interest and reflection.

### Instruments

#### *Quality of individual contributions*

The quality of individual contributions was measured with the Maastricht-Peer Activity Rating Scale (M-PARS). This peer rating scale was developed and validated in an earlier study (Kamp et al., 2011) and consists of three subscales, cognitive, collaborative, and motivational activities, which are believed to be conducive to successful learning in PBL. With this 14-item rating scale students can evaluate each of their peers individually by indicating how much they agree with every item on a five-point scale (1 = completely disagree, 3 = neutral, and 5 = completely agree). One student has to be evaluated by at least four peers for one reliable evaluation. See Appendix A for the M-PARS.

*Perceived effectiveness*

Perceived effectiveness of peer feedback and its impeding and promoting factors was evaluated in two different ways. First, students were subjected to a six-item evaluation questionnaire enquiring the instructiveness and usefulness of the received feedback. These items had to be answered on a five-point scale (see Table I for the items). Second, a focus group was conducted with students. The goal of this focus group was to contribute to the interpretation of the quantitative results (Krueger & Casey, 2009). The focus group was audio taped and led by moderator I (HB) who was provided with a topic list (see Appendix B). Students were asked how the feedback had impacted them and several aspects, selected from Archer (2010) and related to the effectiveness of feedback, such as the structure, timing, and feedback procedure were also addressed.

**Procedure**

Before the pretest, all students were asked to sign an informed consent form. The study consisted of a pre- (week 3) and posttest (week 7) of the quality of individual contributions within the tutorial group as perceived by group members. During this pre- and posttest students had to evaluate the contributions of their peers by completing an M-PARS on each group member. The scores were then aggregated per evaluated student. The intervention took place in week 4, when students received the peer feedback on the quality of their contributions in the form of an overview of the aggregated mean scores and standard deviations per item followed by a - (<3.0 = insufficient), a - (3.0–3.5 = needs improvement) or a neutral (3.5–3.9 = average), or a + ( $\geq 4$  = above average). In order to stimulate reflection on the received peer feedback, students were also provided with a list of improvement tips (Appendix C) that they could use whenever they might score low on particular items, which was the second part of the intervention. Suggestions of appropriate improvement tips per type of activity were listed on the back of the peer feedback form (e.g., search for contradictions in the discussion and express these contradictions). The improvement tips were formulated by the researchers and were derived from the theories that were used to develop the M-PARS (see Kamp et al., 2011). It concerns theories that explain the processes that are conducive to tutorial group effectiveness (e.g., Dolmans et al., 1998; Slavin, Hurley, & Chamberlain, 2003).

Next they were encouraged to formulate improvement goals with regard to the items that needed improvement. Tutors were instructed to stimulate students to reflect on the feedback and to discuss group functioning in general after the feedback was received. But since students were instructed that the data would be analyzed anonymously and would not be reported to their tutors, individual feedback results were not available for the tutors and also not discussed. Because there were no more tutorial group meetings after the data of the posttest were analyzed, students were told that they could receive their feedback on the posttest by e-mail on request. When the course was completed, the focus group was conducted to explore students' perceptions of the effectiveness of the peer feedback. Participants were gathered by means of purposive sampling; one student per tutorial group was invited to participate in the focus group by e-mail.

Table 1

*Mean Scores and Standard Deviations for the Items of the Evaluation Questionnaire Filled in by the Students in Week 7*

<b>Items</b>	<b>Mean</b>	<b>Standard deviation</b>
1. The feedback gave a good view of my own functioning	3.31	.98
2. I found it instructive to receive this feedback	3.13	1.15
3. The tutor stimulated me to reflect on the received feedback	2.84	1.03
4. I formulated useful improvement goals	2.65	1.10
5. The anonymity of the feedback was warranted sufficiently	3.97	.94
6. The tutor evaluated functioning of group half way through the course	3.56	1.18

Note. 1 = completely disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = completely agree.

## Analyses

Students were first divided into three groups based on their initial score on the M-PARS (low score: < 33 %, average score: 33–67 %, and high score: > 67 %). In order to assess the impact of the intervention on the three groups a mixed between-within subjects analysis of variance was conducted with total M-PARS score as dependent variable. For the quantitative data of the evaluation questionnaire means and standard deviations were calculated. The audio taped data from the focus group was summarized by moderator II (RK). This summary was sent to all students who participated in the focus group. They were asked to determine the accuracy and, if necessary, to give corrections.

## RESULTS

### Quality of contributions

Of the 87 students, 74 students received peer feedback (response 85 %). Within one tutorial group only three students signed the informed consent form. This group (nine students) had to be removed from the study because a minimum of four evaluators per student is required for a reliable evaluation. The other four students that dropped out of the study attended too few tutorial meetings, making it impossible for their group members to evaluate them in a reliable and valid way.

A mixed between-within subjects analysis of variance was conducted to evaluate the impact of the feedback intervention on the total score on the M-PARS for the three different groups (low score:  $n = 24$ , moderate score:  $n = 25$  or high score:  $n = 25$  on the M-PARS pretest). There was a significant interaction effect between group and time, *Wilks' Lambda* = 0.88,  $F(2, 71) = 4.70$ ,  $p = 0.012$ , *partial eta squared* = 0.12. The main effect comparing the three groups was also significant,  $F(2, 71) = 60.93$ ,  $p < 0.0005$ , *partial eta squared* = 0.63. There was, however, no main effect for time, *Wilks' Lambda* = 0.98,  $F(1, 71) = 1.66$ ,  $p = 0.20$ , *partial eta squared* = 0.02. These results indicated that there was a difference between the three groups with regard to the effectiveness of the intervention. Therefore, three separate paired samples t-tests (one for each group) were conducted. Only the group with an initial low score on the M-PARS showed a significant, but small increase in their quality of individual contributions ( $M = 3.36$ ,  $SD = 0.15$ ;  $M = 3.52$ ,  $SD = 0.33$ ),  $t(23) = -2.29$ ,  $p = 0.031$ ). The results of

the pre- and posttests of the three groups are displayed in Figure 1a. Figures 1b, c, and d show the results per subscale (constructive, collaborative, and motivational) for each group (low, average, and high score).

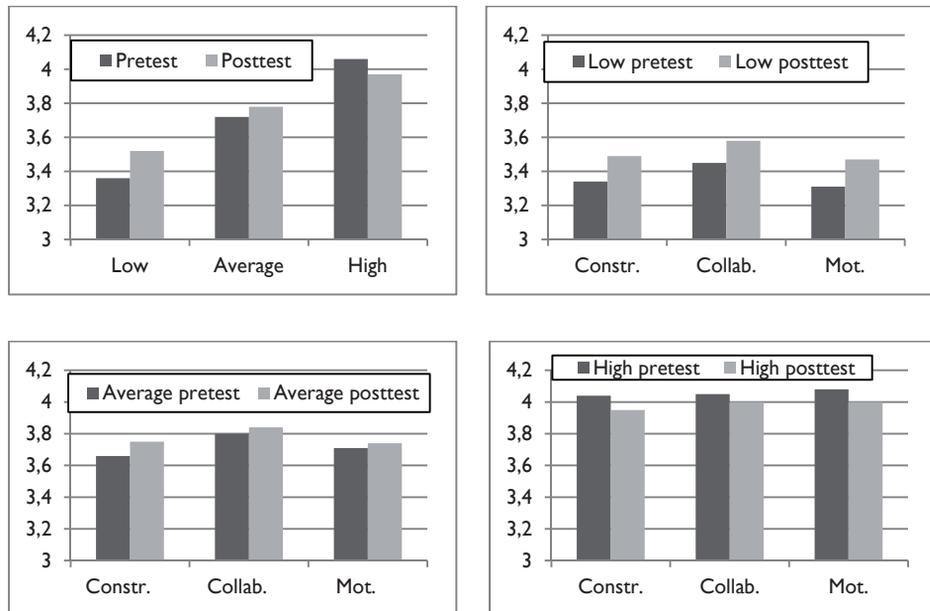


Figure 1. Pre- and posttest mean scores on the three subscales of the M-PARS for a) the three groups, b) the group with a low initial score, c) the group with an average initial score, d) the group with a high initial score.

### Evaluation questionnaire

Of the 74 students who received feedback in this study, 68 students filled in the evaluation questionnaire (response 92 %). Means and SDs were calculated for all items of this questionnaire (see Table 1). The following norm was kept: a mean below 3.0 was considered as insufficient, a mean between 3.0 and 3.5 implies that there is room for improvement with regard to the concerning item, and a mean of 4.0 or higher is considered as good. As can be seen in Table 1, the scores on the evaluation items differed between 2.65 and 3.97. The following 2 items scored below 3. Firstly, students reported that they did not formulate useful improvement goals ( $M = 2.65$ ,  $SD = 1.10$ ). In addition, students perceived that the tutors did not sufficiently stimulate

them to reflect on the received feedback ( $M = 2.84$ ,  $SD = 1.03$ ). Students also indicated that the feedback gave them a reasonably good view of their functioning, but this could still be improved ( $M = 3.31$ ,  $SD = 0.98$ ). Some students also felt that the instructiveness of the feedback could be improved ( $M = 3.13$ ,  $SD = 1.15$ ). These results, therefore, suggest that there is room for improvement.

### **Focus group**

One student per tutorial group was invited to participate in the focus group and seven students agreed to participate. All seven students signed an informed consent form. A short summary of the results of the focus group can be found below.

Students indicated that giving feedback made them more aware of what was considered ‘appropriate tutorial behavior’ and how individual contributions influence the functioning of the group. The distinction between the three types of tutorial contributions (cognitive, collaborative, and motivational) contributed to this awareness. It also made them more critical to each other.

*“I think that, by looking at the quality (of contributions), it made everybody more aware of the quality level that is asked from students, especially because of the different aspects (of good tutorial behavior) that were identified.”*

*“I felt it had a positive effect on the sense of belonging to the group.”*

Students indicated that there was a strong relation between the functioning of the group in general and their willingness to improve their own performance in the tutorial group. Students in ‘good functioning’ groups felt that receiving the feedback had had a positive influence on their own performance as well as the functioning of the group as a whole. One student even said that it had improved her sense of belonging to the group. Students within ‘poorly functioning’ groups were less motivated to improve.

*“I think that it (quality of the group) also influences your own performance in the tutorial group..., because if nobody says anything and you have already read out everything for the last three tutorial meetings, by the fourth time you will think: Figure it out for yourself.”*

Students felt that the time between receiving the feedback and the end of the course was a bit too short to improve their contributions to the tutorial group. They felt that the effectiveness of the peer feedback would have improved when the course would have been longer.

In general, students felt that the numerical design of the feedback (item mean followed by a -, - or +) was very convenient, and they found it interesting to learn from the written feedback and improvement tips how they could improve whenever they received a below average score on an item. Therefore students felt that the improvement suggestions were useful. With regard to the formulation of improvement goals, students reported that this was not always taken very seriously, especially since the formulated improvement goals did not have to be handed in. Not all students recognized the use of formulating these goals. Other students reported that they had kept their goals in mind and were curious whether they had improved on the posttest.

*“I would have been very interested in the second feedback score, to see whether or not I improved.”*

The participating students were asked if they had any points for improvement. Four suggestions were proposed by the students.

First, students felt that the impact of the feedback could have been greater if the advantage of receiving this feedback would have been more emphasized beforehand. They indicated that they would be more motivated if the relation between the quality of students' contributions and achievement had been more stressed.

*“The ‘what’s in it for me’ was a bit missing. It would have been better if it was emphasized that this feedback focuses on your functioning or your points of improvement.”*

Second, the participants thought the peer feedback would be more effective earlier on in the curriculum. According to them, first-year students lack a clear picture of their own individual contributions within the tutorial group and peer feedback could contribute to the illustration of their own functioning. According to several students this could stimulate students to participate more seriously in PBL tutorial groups. Second-year students are already more or less set in their ways.

*“Now (in the second year) you have already found your ways in the tutorial group. First-year students still have to develop their ways and if you then already hear that you do not contribute enough, and that this and this is good, but this and this can improve.....I also think that first-year students take the feedback much more serious, because they still apply the seven-jump.”*

Third, as was already stated, students indicated that they sometimes wondered why they were rated below the mean on a certain item and they expressed the desire to receive some verbal clarification from their peers on their feedback. They believed that the feedback would have had more impact with this verbal clarification.

*“If you personalize it (the feedback) you could explain and nuance it. If you receive it only on paper you tend to shove it aside, but if someone tells you, you want to do something about it.”*

Some students, however, appreciated the anonymity of the peer feedback. They felt that publicly discussing the feedback would have a counterproductive effect because students would be giving more socially acceptable ratings. These students indicated that they had rated their peers in a critical, but honest way because they knew they would remain anonymous.

*“I think it’s counterproductive (to openly discuss the feedback). If you fill in the feedback rating scale anonymously people dare to be more honest.”*

*“I have filled it in very critically, but if it hadn’t been anonymous I might not have filled it in at all.”*

Fourth, even though students felt the improvement suggestions were useful, they thought that the improvement tips could have been more attuned to their personal score. Now everybody received the same tips and had to pick out the ones that were applicable to them.

## DISCUSSION

The purpose of this study was to investigate whether or not peer process feedback combined with goal setting could improve the quality of students' individual contributions within the tutorial group. In order to answer these questions a mixed-methods explanatory design was used, in which the design of the quantitative study was dominant and followed by the qualitative study. The qualitative study served as an explanation and refinement of the quantitative results (Creswell & Plano Clark, 2007). Based on the quantitative data it can be concluded that, with the exception of the students with a low score on the pretest, this peer feedback intervention does not seem to improve the quality of students' individual contributions in the tutorial group. So it seems that only the students who contribute in a poor way seem to benefit from this peer process feedback. The fact that only low performing students seem to improve their individual contributions is in line with the findings of Gielen et al. (2010). They also found that peer feedback is more effective for students with a low to average competence level as opposed to students with a high competence level. The data from the evaluation questionnaire confirmed that the instructiveness of the feedback could be improved. The focus group clarified why the feedback was not effective for the majority of the students, even though students felt that the peer feedback had made them more aware of what good tutorial behavior consists of. This resulted in a number of suggestions for improving the effectiveness of the intervention. First, there was a need for a face-to-face clarification and discussion of the received feedback. In the focus group students indicated that by discussing the feedback as a group they would be more motivated to process the feedback. Archer (2010) agrees that feedback is best discussed face-to-face, but for privacy reasons this was not done in this study. Second, more emphasis should be put on the stimulation of goal setting and reflection. In the evaluation questionnaire students already indicated that they did not formulate useful improvement goals. The reason for this was that students felt a need for more personalized improvement tips instead of the general tips and that the formulated goals were not monitored in the tutorial group. Thus, students should be stimulated more to set goals for, reflect on and monitor their contributions throughout the tutorial group meetings, especially since goal setting and reflection are important conditions for (peer) feedback to be effective (Archer, 2010; Gielen et al., 2010; Prins et al., 2006). Third, more attention should have been paid to the feed up

(explain purpose and goal of the feedback), which is also conditional to effective feedback (Hattie & Timperley, 2007). Students indicated that, if the purpose and goal had been more emphasized, they believed the feedback would have been more effective. Fourth and last, peer feedback was thought to be more effective for first-year students, since they still have to develop their tutorial behavior. Second-year students believed they were already more or less set in their own ways, which made them less susceptible to the feedback.

Based on these findings, one could conclude that the treatment in this intervention needs to be strengthened in order to produce a behavioral change within students. Besides the aspects that were identified by the students themselves, another explanation might be that students simply did not have enough tutorials left after receiving the feedback in order to work on the improvement of their contributions. A last explanation could be that students tend to only focus on the items they received a below average score on and focused less on the aspects that scores relatively high.

Although this study provided rich insights in why and when peer process feedback might be effective in terms of improving students' individual contributions in tutorial groups, this study also has some important limitations. One of these limitations is that the quantitative results could also be explained by regression towards the mean because both the low and high score group trended towards the mean. The low score group did, however, show a significant and greater gain score than the high score group. In order to entirely refute this explanation and to draw solid conclusions about the effectiveness of this intervention and the effect of the different aspects discussed in the focus group, in future research experimental and control groups should be compared (Van Zundert et al., 2010). In such a study the students in the control group should not receive feedback or feedback in a more stripped form. Nevertheless, the aim of this study was not only to explain if peer process feedback is effective but also to clarify why and when peer process feedback might be effective. Another aspect that is worth investigating in an experimental study is the effect of actively involving the tutor in the feedback process. This is interesting because, as was mentioned in the introduction, there is an interaction between tutor performance, group productivity, and learning effectiveness. Another important limitation is that the conclusions of this study are based on students' perceptions and not on observed behavior. Although peer evaluations are seen as valid and reliable

information (Eva & Regehr, 2011), it would be interesting to explore whether peer feedback also affects achievement scores.

## REFERENCES

- Adcroft, A. (2011). The mythology of feedback. *Higher Education Research & Development, 30*(4), 405-419. doi: 10.1080/07294360.2010.526096
- Archer, J. C. (2010). State of the science in health professional education: Effective feedback. *Medical Education, 44*, 101-108. doi: 10.1111/j.1365-2923.2009.03546.x
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage Publications.
- Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning: An integrative model for effective team learning in organizations. *Educational Research Review, 5*, 111-133. doi: 10.1016/j.edurev.2010.02.002
- Dolmans, D. H. J. M., & Wolfhagen, I. H. A. P. (2005). Complex interactions between tutor performance, tutorial group productivity, and the effectiveness of PBL units as perceived by students. *Advances in Health Sciences Education, 10*, 253-261. doi: 10.1111/j.1365-2929.2005.02205.x
- Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., Scherpbier, A. J. J. A., & Van der Vleuten, C. P. M. (2003). Development of an instrument to evaluate the effectiveness of teachers in guiding small groups. *Higher Education, 46*(4), 431-446. doi: 10.1023/A:1027388806218
- Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (1998). Motivational and cognitive processes influencing tutorial groups. *Academic Medicine, 73*(Suppl 10), S22-S24.
- Dominick, P. G., Reilly, R. R., & McGourty, J. W. (1997). The effects of peer feedback on team member motivation. *Group & Organization Management, 22*(4), 508-520. doi: 10.1177/1059601197224006
- Eva, K. W., & Regehr, G. (2011). Exploring the divergence between self-assessment and self-monitoring. *Advances in Health Sciences Education, 16*, 311-329. doi: 10.1007/s10459-010-9263-2
- Geister, S., Konradt, U., & Hertel, G. (2006). Effects of process feedback on motivation, satisfaction, and performance in virtual teams. *Small Group Research, 37*(5), 459-489. doi: 10.1177/1046496406292337

- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*, 3-31. doi: 10.1007/978-3-8348-9837-1
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). The effectiveness of peer feedback for learning: The effects of constructiveness, accuracy and embedding in the learning environment. *Learning and Instruction, 20*(4), 304-315. doi: 10.1016/j.learninstruc.2009.08.007
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi: 10.3102/003465430298487
- Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2011). Can students adequately evaluate the activities of their peers in PBL? *Medical Teacher, 33*(2), 145-150. doi: 10.3109/0142159X.2010.509766
- Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M. & Schmidt, H. G. (2012). The relationship between students' small group activities, time spent on self-study, and achievement. *Higher Education, 64*(3), 385-397. doi: 10.1007/s10734-011-9500-5
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284. doi: 10.1037/0033-2909.119.2.254
- Krueger, R. A., & Casey, M. A. (2009). *Focus groups: A practical guide for applied research* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Lee, Y. J., & Roth, W.-M. (2007). The individual collective dialectic in the learning organization. *The Learning Organization, 14*, 92-107. doi: 10.1108/09696470710726970
- Moust, J. H. C., Van Berkel, H. J. M., & Schmidt, H. G. (2005). Signs of erosion: Reflections on three decades of problem-based learning at Maastricht University. *Higher Education, 50*, 665-683. doi: 10.1007/s10734-004-6371-z
- Papinczak, T., Young, L., & Groves, M. (2007a). Peer assessment in problem-based learning: A qualitative study. *Advances in Health Sciences Education, 12*, 169-186. doi: 10.1007/s10459-005-5046-6
- Papinczak, T., Young, L., Groves, M., & Haynes, M. (2007b). An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Medical Teacher, 29*, e122-e132. doi: 10.1080/01421590701294323

- Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior*, 26(2), 151-161. doi: 10.1016/j.chb.2009.10.011
- Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior*, 27(3), 1087-1102. doi: 10.1016/j.chb.2010.06.024
- Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles, and preferences. *Advances in Health Sciences Education*, 11(3), 289-303. doi: 10.1007/s10459-005-3250-z
- Savery, J. R., & Duffy, T. M. (1996). *Problem based learning: An instructional model and its constructivist framework*. In B. G. Wilson (Ed.), *Constructivist learning environments: Case studies in instructional design* (pp. 135-148). Englewood Cliffs, NJ: Educational Technology Publications.
- Schönrock-Adema, J., Heijne-Penninga, M., Van Duijn, M. A. J., Geertsma, J., & Cohen-Schotanus, J. (2007). Assessment of professional behavior in undergraduate medical: Peer assessment enhances performance. *Medical Education*, 41, 836-842. doi: 10.1111/j.1365-2923.2007.02817.x
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. M. (2003). *Cooperative learning and achievement: Theory and research*. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology* (Vol. 7, pp. 177-198). Hoboken, NJ: Wiley.
- Sluijsmans, D. M. A., Brand-Gruwel, S., & Van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, 27(5), 443-545. doi: 10.1080/0260293022000009311
- Sluijsmans, D. M. A., Moerkerke, G., Dochy, F., & Van Merriënboer, J. J. G. (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation*, 27(2), 153-173.
- Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education*, 32(3), 321-345.
- Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, 20(4), 339-343. doi: 10.1016/j.learninstruc.2009.08.003

- Van den Bossche, P., Gijsselaers, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments. Team learning beliefs and behaviors. *Small Group Research*, 37(5), 490-521. doi: 10.1177/1046496406292938
- Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20(4), 270-279. doi: 10.1016/j.learninstruc.2009.08.004
- Visschers-Pleijers, A. J. S. F., Dolmans, D. H. J. M., De Leng, B. A., Wolffhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2006). Analysis of verbal interactions in tutorial groups: A process study. *Medical Education*, 40(2), 129-137. doi: 10.1111/j.1365-2929.2005.02368.x
- Zumbach, J., Hillers, A., & Reimann, P. (2004). Supporting distributed problem-based learning: The use of feedback mechanisms in online learning. In T. S. Roberts (Ed.), *Collaborative learning: Theory and practice* (pp. 86-102). Hershey, PA: Information Science Publishing.
- Zumbach, J., Muehlenbrock, M., Jansen, M., Reimann, P., & Hoppe, H.-U. (2002). Multidimensional Tracking in Virtual Learning Teams. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community* (pp. 650-651). Hillsdale, NJ: Erlbaum.



# CHAPTER 5

## Midterm Peer Feedback in Problem-Based Learning Groups: The Effect on Individual Contributions and Achievement

Published as:

Kamp, R. J. A., Van Berkel, H. J. M., Popeijus, H. E., Leppink, J., Schmidt, H. G., & Dolmans, D. H. J. M. (2013). Midterm peer feedback in problem-based learning groups: The effect on individual contributions and achievement. *Advances in Health Sciences Education*. doi: 10.1007/s10459-013-9460-x



**ABSTRACT**

Even though peer process feedback is an often used tool to enhance the effectiveness of collaborative learning environments like PBL, the conditions under which it is best facilitated still need to be investigated. Therefore, this study investigated the effects of individual versus shared reflection and goal setting on students' individual contributions to the group and their academic achievement. In addition, the influence of prior knowledge on the effectiveness of peer feedback was studied. In this pretest – intervention - posttest study 242 first year students were divided into three conditions: condition 1 (individual reflection and goal setting), condition 2 (individual and shared reflection and goal setting), and condition 3 (control group). Results indicated that the quality of individual contributions to the tutorial group did not improve after receiving the peer feedback, nor did it differ between the three conditions. With regard to academic achievement, only males in conditions 1 and 2 showed better academic achievement compared with condition 3. However, there was no difference between both ways of reflection and goal setting with regard to achievement, indicating that both ways are equally effective. Nevertheless, it is still too early to conclude that peer feedback combined with reflection and goal setting is not effective in enhancing individual students' contributions. Students only had a limited number of opportunities to improve their contributions. Therefore, future research should investigate whether an increase in number of tutorial group meetings can enhance the effectiveness of peer feedback. In addition, the effect of quality of reflection and goal setting could be taken into consideration in future research.

## INTRODUCTION

Collaborative learning environments are widely implemented in today's organizational and educational practice. The effectiveness of these collaborative learning groups can, however, be affected by cognitive, social, and motivational factors (Van den Bossche, Gijsselaers, Segers, & Kirschner, 2006), leading to a superficial discussion and a lack of cohesion within the group (Decuyper, Dochy, Van den Bossche, 2010). Therefore, interventions aimed at improving the effectiveness of collaborative learning environments, should be directed at the way group members work together within collaborative groups.

Within educational collaborative learning practice, peer feedback has the potential to be a powerful tool to enhance individual performance (Van Gennip, Segers, & Tillema, 2009; Sluijsmans, Brand-Gruwel, & Van Merriënboer, 2002). It is thought to be effective for two reasons. The first reason is that it can facilitate self-evaluation by clarifying what good performance is (Adcroft, 2011) and identifying the gap between a student's current and desired behavior (Archer, 2010). White and Gruppen (2010) have developed a model that described how peer feedback can facilitate self-regulated learning in medical education. According to this model, learning has four components: planning, learning, assessment, and adjustment. External feedback such as peer feedback (assessment) leads to reflection (adjustment), which leads to goal setting and motivation (planning), which finally influences learning. The second reason is that it increases students' sense of responsibility for the other group members (Papinzcak, Young, & Groves, 2007). Students have reported an increased sense of belonging to the group after receiving peer feedback (Kamp, Dolmans, Van Berkel, & Schmidt, 2013). According to Ryan and Deci (2000), these feelings of relatedness can turn extrinsic motivation into intrinsic motivation for learning.

Peer feedback can, however, address different levels (Hattie & Timperley, 2007). One of these levels is the process level (i.e. "the main processes needed to understand/perform the tasks", Hattie & Timperley, 2007, p.87). Feedback on students' contributions to group processes can be regarded as process feedback. Dominick, Reilly, and McGourty (1997) indicate that process peer feedback can be an appropriate intervention to improve the quality of students' individual contributions within groups. Although research on peer process feedback is scarce, it has demonstrated to have a positive effect on motivation, behavior, communication, and collaboration of group

members (Dominick et al., 1997; Geister, Konradt, & Hertel, 2006). Schönrock-Adema, Heijne-Penninga, Van Duijn, Geertsma, and Cohen-Schotanus, (2007) performed an experimental study on the effectiveness of peer process feedback in problem-based learning (PBL) tutorial groups. They found that students who were asked to evaluate the professional behavior of the peers in the tutorial group obtained a higher score with regard to their own professional behavior.

Although process peer feedback has the potential to be an effective tool to promote individual functioning within groups, it is also clear that the effectiveness of peer feedback in general depends on multiple variables (Eva et al., 2012; Topping, 2010). Sluijsmans, Moerkerke, Van Merriënboer, and Dochy (2001) and Van Zundert, Sluijsmans, and Van Merriënboer (2010), for instance, concluded that students need to be trained how to give feedback to their peers. Students find it especially difficult to make negative judgments without any prior training or experience in peer feedback. Peer feedback training appeared to enhance students' ability to assess their peers (Sluijsmans et al., 2002). In addition, it is important that feedback is delivered in such a way that it stimulates students' self-monitoring of their own functioning (Archer, 2010). This implies that the peer feedback process does not stop after the feedback is received, but that students should be stimulated to reflect and act upon the received feedback (Gibbs & Simpson, 2004; Hattie & Timperley, 2007; Prins, Sluijsmans, & Kirschner, 2006). Therefore, students should be provided with clues for behavioral change and should be stimulated to formulate goals for improvement (Prins et al., 2006). It is also important to stimulate reflection, because the effectiveness of feedback depends on how the feedback is received (Eva et al., 2012; Sargeant, Mann, Sinclair, Van der Vleuten, & Metsemakers, 2008). Feedback acceptance and use is, for instance, influenced by the receiver's confidence and experience with the activity that is being evaluated (Eva et al., 2012). Reflection combined with goal setting can, therefore, be a useful tool to increase feedback acceptance and use (Sargeant et al., 2008). Phielix, Prins, Kirschner, Erkens, and Jaspers (2011) found that when midterm peer process feedback is combined with reflection (within a Computer Supported Collaborative Learning (CSCL) environment) it increases awareness of cognitive and social behavior. This, in turn, resulted in an increased satisfaction with the group process and led to a higher quality of contributions to the social climate within the group. The reflection in the study of Phielix et al. (2011) was not only an individual process, but students were also stimulated to reflect collaboratively on the group performance. Students were not

only asked to reflect on their current progress and on whether or not they agreed with the feedback, but also to formulate actions that needed to be taken in order to close the gap between their current progress and the desired behavior.

Within a PBL environment a similar study on the effectiveness of midterm process peer feedback was performed (Kamp et al., 2013). Here students received peer process feedback on their cognitive, collaborative, and motivational contributions to the group. Next, students were asked to reflect on whether or not the feedback was in line with their expectations, on the improvement goals they were going to formulate for the following weeks, and how they were going to put them into practice. Students with a low initial performance showed an increase in the quality of contributions four weeks after the feedback was received. Qualitative analyses indicated that the intervention had increased their awareness of 'good tutorial behavior' and increased their sense of belonging to the group, but that they felt a need for a collaborative face-to-face discussion of the received feedback. This study suggests that there is another factor that influences the effectiveness of peer feedback and that is the performance or competence level of the student. Because, within collaborative learning groups, the functioning of the group (and therefore the quality of individual contributions) and subsequent achievement is influenced by the amount of prior knowledge a student has (Gijsselaers & Schmidt, 1990), it might also influence the extent to which the peer process feedback is effective. This implies that the initial quality of contributions of a student with low prior knowledge might be poor, and is therefore more receptive to peer feedback. In addition, this student will have more room for personal growth.

A last factor that is conducive to the effectiveness of peer feedback is the presence of justification (explanation of given judgment) (Gielen, Peeters, Dochy, Onghena, & Struyven, 2010). This factor had a positive effect on subsequent performance on a writing task, but again only for students with a lower initial score before the intervention.

Although research on peer process feedback has resulted in interesting and valuable findings, some aspects still need further investigation. In general, research on peer process feedback is often based on students' observations and experimental research is scarce, especially experimental research with a control group (Strijbos & Sluijsmans, 2010, Van Zundert et al., 2010). Therefore, it is still unclear how peer feedback can be facilitated optimally. Research has indicated that reflection and goal

setting are essential for peer feedback to be effective, but experimental research on how these conditions can be put into practice optimally is scarce. Research has, for instance, never demonstrated that collaborative or shared reflection is more effective than individual reflection. In addition, there is little experimental research with a control group that has demonstrated that peer process feedback does not only lead to better contributions but also to better achievement. This is unfortunate because, a recent study has indicated that there is a positive relation between the quality of individual contributions and achievement (Kamp, Dolmans, Van Berkel, & Schmidt, 2012). One can, therefore, expect that when peer feedback enhances students' contributions this in turn will enhance students' achievement. A final gap in experimental peer process feedback research is the influence of prior knowledge. The question remains whether students' prior knowledge influences the extent to which the feedback is effective. Therefore, this experimental study will investigate the effects of individual versus shared reflection and goal setting on students' individual contributions to the group and their academic achievement. Furthermore, the influence of prior knowledge on the effectiveness of peer process feedback will be studied. This results in the following research questions:

- RQ1. Does midterm peer process feedback combined with reflection and goal setting improve *the quality of students' individual contributions* within the tutorial group, and is this effect influenced by how reflection is facilitated (individual vs. individual and shared) and students' prior knowledge?
- RQ2. Does midterm peer process feedback combined with reflection and goal setting improve *students' individual achievement* within the tutorial group, and is this effect influenced by how reflection is facilitated (individual vs. individual and shared) and students' prior knowledge?

## METHOD

### Participants

Participants were 242 first-year students enrolled in a PBL bachelor's program

in Health Sciences at Maastricht University in The Netherlands. During this eight-week course students met each other once a week for a two-hour tutorial group meeting. The students were divided into 23 tutorial groups, each consisting of approximately ten students. In these tutorial groups students discuss a given problem and generate learning issues for further self-study. During the next tutorial, students discuss and synthesize their findings. All groups were guided by a tutor. The topic of this course was human biology and students enter this course with different levels of prior knowledge in biology. In Dutch secondary schools students can choose the subject biology or not, but both types of students are admissible to the bachelor's program in Health Sciences.

During the first three weeks of the first-year of their bachelor program (and prior to this study) all students received a two-hour PBL training 'learning and communicating in tutorial groups'. In this training, students are taught, among other things, how to give feedback to one another within the tutorial group. The study was performed after all participants signed an informed consent form.

### Instruments

The quality of individual contributions was measured with the Maastricht-Peer Activity Rating Scale (M-PARS). This peer rating scale was developed and validated in an earlier study (Kamp, Dolmans, Van Berkel, & Schmidt, 2011) and consists of three subscales: 1. cognitive, 2. collaborative, and 3. motivational activities, which have been demonstrated to be related to successful learning in PBL (Kamp et al., 2012). With this 14-item rating scale students can evaluate each of their peers individually by indicating how much they agree with every item on a five-point scale (1 = *completely disagree*, 3 = *neutral*, and 5 = *completely agree*). See Appendix A for the M-PARS.

Academic achievement was measured with a final MCQ exam addressing the subject-matter of the course. This test consisted of 100 multiple choice questions with three choices per question. For every answer students could score a one (correct) or a zero (incorrect), leading to a possible score between zero and 100%. The test had a reliability of 0.82 (Cronbach's alpha).

Prior knowledge was included as a dichotomous variable. In Dutch secondary education students can choose to follow the subject biology. Students who had followed this subject were said to have prior knowledge with regard to this course and students who had not followed this subject were said to have no prior knowledge.

## Procedure

This study adopted an experimental pretest–intervention–posttest design (see Table 1). Participants were randomly divided into 23 tutorial groups. These tutorial groups were again randomly divided over three conditions: condition 1 (midterm peer feedback combined with individual reflection); condition 2 (midterm peer feedback combined with individual and shared reflection), and condition 3 (control group with no midterm peer feedback).

Table 1

*Research Design*

	<b>Evaluation of midterm student functioning (week 3)</b>	<b>Intervention (week 4)</b>	<b>Evaluation of end term student functioning and achievement (week 7)</b>
<i>Condition 1 (individual reflection)</i>	M-PARS	Peer feedback combined with individual reflection	- M-PARS - Final exam (MCQ)
<i>Condition 2 (individual and shared reflection)</i>	M-PARS	Peer feedback combined with individual and shared reflection	- M-PARS - Final exam (MCQ)
<i>Condition 3 (control group)</i>	-	-	- M-PARS - Final exam (MCQ)

The midterm evaluation of the quality of individual contributions took place during the tutorial group meeting of week 3. Conditions 1 and 2 participated in this measurement by filling out one M-PARS for every group member within their tutorial group. Both of these conditions received the midterm peer process feedback during the tutorial group meeting of week 4. This feedback displayed the aggregated evaluations students gave their peers during the midterm evaluation by means of an overview of the mean scores and standard deviations per item followed by a - - (< 3.0 = insufficient), a - (3.0 - 3.5 = needs improvement), a neutral (3.5 – 3.9 = average), or

a + ( $\geq 4$  = above average). See Appendix D for an example.

In order to stimulate individual reflection and goal setting on the received peer feedback, students had to write down the answers the following three questions:

*Is the feedback in line with your expectations?;*

*If not, what were your expectations?;*

*Which improvement goals are you going to formulate for the following weeks (min. 1, max. 3) and how are you going to put them into practice?*

In order to help students formulate appropriate improvement goals they were provided with a list of improvement tips (Appendix C) that they could use whenever they might score low on particular items. Suggestions for appropriate improvement tips per type of activity were listed on the back of the peer feedback form (e.g., search for contradictions in the discussion and express these contradictions). The improvement tips were formulated by the researchers and were derived from the theories that were used to develop the M-PARS (see Kamp et al., 2011). It concerns theories that explain the processes that are conducive to tutorial group effectiveness (e.g., Dolmans, Wolhagen, & Van der Vleuten, 1998; Slavin, Hurley, & Chamberlain, 2003).

The addition in condition 2 was that students were invited by the tutor to ask for clarifications and explanations with regard to the received peer feedback and to share their reflections and formulated goals with their group members. In both conditions the tutor facilitated the process by means of a protocol that was formulated by the researchers. With this protocol tutors were instructed to hand-out the peer feedback, to ask the students to take a good look at the peer feedback, to ask them to answer the reflection and goal setting questions on paper, and to invite students one-by-one to discuss the received peer feedback, their reflections and goals. Students were also stimulated by the tutor to discuss the peer feedback and ask for clarifications and/or justifications from their peers. Based on these answers students were stimulated to adjust their goals. Students in condition 3 did not receive any midterm peer feedback at all. The end term evaluation of the quality of individual contributions took place in week 7. All three conditions participated in this

measurement. After the course had ended all students took the final exam. Finally, all students received their peer feedback based on the end term evaluation by e-mail.

### **Analyses**

Because students in the current study were nested within tutorial groups, multilevel analyses were performed to answer the research questions: 1) one on the M-PARS end term evaluation scores of the three conditions (RQ1); 2) one on the M-PARS gain score (end term evaluation – midterm evaluation) of conditions 1 and 2 (RQ1); and 3) one on the final exam scores of the three conditions (RQ2). In all three analyses prior knowledge in biology (yes vs. no) was included as a covariate, because it was expected that this prior knowledge would affect both students' contributions as well as their achievement on the final exam. In addition, gender and final exam score of the course that ran prior (prior final exam) were included as covariates. Including these variables, controlled for potential confounding due to an unbalanced design and it increased statistical power for the analyses by reducing the unexplained variance.

## **RESULTS**

The 23 tutorial groups (242 participants) were divided into three conditions: condition 1 ( $n = 83$ , individual reflection and goal setting), condition 2 ( $n = 85$ , individual and shared reflection and goal setting), and condition 3 ( $n = 74$ , control group). During the study, 24 students were absent during three or more tutorials. These students were eliminated from the study because they were not able to fully participate in the tutorial group and could, therefore, not be evaluated by their peers. In addition, two outliers were removed. This resulted in a total of 216 students that were evaluated by their peers (C1,  $n = 77$ ; C2,  $n = 70$ ; C3,  $n = 69$ ) and received feedback (response rate 89%). The vast majority of the students in both experimental conditions formulated their reflections and improvement goals. The results of the frequency and descriptive analyses on the variables gender and prior final exam are displayed in Table 2.

Table 2

*Frequencies of Variables Gender and Prior Final Exam in the Three Conditions*

<b>Condition</b>	<b>Gender (male)</b>	<b>Prior final exam (mean)</b>
Condition 1 (Individual reflection)	36%	77.94
Condition 2 (individual and shared reflection)	34%	77.65
Condition 3 (Control group)	22%	78.45

Potential confounding effects caused by the different proportions of males within the three conditions are controlled for by including 'gender' as a covariate in the analyses. In order to examine whether there were significant initial differences between males and females on the pretest M-PARS and three subscales scores four fixed effects models were computed on these pretest scores. In all models gender, prior knowledge in biology, prior final exam score, and experimental condition were included as covariates. All models indicated that there were no significant interaction or main effects for all covariates, indicating that, prior to the intervention, there were no differences between males and females or students in conditions 1 and 2 on the M-PARS scores. Table 3 displays the estimated marginal M-PARS means for males and females, and for students from conditions 1 and 2, adjusted for potential confounding due to gender (for condition) and prior knowledge.

Table 3

*Estimated Marginal Means on the Pretest M-PARS Score for Males and Females and for Conditions 1 and 2 Adjusted for Potential Confounding due to Gender (for Condition) and Prior Knowledge*

		<b>M</b> <b>(total M-PARS)</b>	<b>M</b> <b>(cognitive)</b>	<b>M</b> <b>(collaborative)</b>	<b>M</b> <b>(motivational)</b>
<i>Gender</i>	Males	3.62	3.55	3.67	3.66
	Females	3.58	3.46	3.67	3.64
<i>Condition</i>	1 (individual reflection)	3.58	3.48	3.66	3.61
	2 (individual and shared reflection)	3.62	3.53	3.67	3.67

*RQ 1.* Does midterm peer process feedback combined with reflection and goal setting improve *the quality of students' individual contributions* within the tutorial group, and is this effect influenced by how reflection is facilitated (individual vs. individual and shared) and students' prior knowledge?

### **M-PARS gain score**

A fixed effects model was computed for the gain score on the M-PARS between the midterm and end term evaluation for condition 1 (individual reflection) and 2 (individual and shared reflection). Gender, prior knowledge in biology, prior final exam score, and experimental condition were included as covariates. A random intercepts model was also performed, but the random intercept variance was not statistically significant (residual variance = .059, variance RI (tutor group) = .008, Wald  $Z = 1.221$ ,  $p = .222$ ), indicating that there were no significant group effects of tutorial group. In the fixed effects model there were no significant interaction effects, indicating that the effectiveness of the intervention was not affected by the other covariates (e.g., gender, biology, and prior final exam). In addition, the intervention itself was not significant ( $t(113) = .81$ ,  $p = .423$ ), indicating that there was no difference between the

pre- end posttest for both experimental groups. A more parsimonious fixed effects model (see Table 4) also indicated that there were no main effects on the M-PARS gain score. Table 5 displays the estimated marginal means on the M-PARS gain score for the three conditions adjusted for potential confounding due to gender and prior knowledge.

In addition, separate fixed effects models were performed for the M-PARS gain scores of the three subscales. Again, these models indicated that there were no significant main effects on subscale gain scores. The only exception was the significant effect of gender on the collaborative gain score, indicating that females, on average, had a .18 lower gain score on the collaborative subscale than males ( $t = -2.72, p = .008$ ).

Table 4

*Main Effects of the Fixed Effect Model on the M-PARS Gain Score of Condition 1 (Individual Reflection) and 2 (Individual and Shared Reflection)*

<b>Variable</b>	<b>B</b>	<b>SE</b>	<b>t</b>	<b>Sig.</b>
Intercept	.058	.29	.20	.846
Prior final exam	-.0005	.004	-.14	.887
Biology	.05	.05	.86	.393
Gender	-.10	.06	-1.80	.074
Condition 1	.04	.08	.76	.450

Table 5

*Estimated Marginal Means on the M-PARS Gain Score for Condition 1 (Individual Reflection) and 2 (Individual and Shared Reflection) Adjusted for Potential Confounding due to Gender and Prior Knowledge*

<b>M-PARS Gain Score</b>	<b>Estimated Marginal Mean</b>	<b>95% Confidence Interval</b>
Condition 1	.007	-.06 - .07
Condition 2	-.03	-.10 - .04

### M-PARS end term score

A fixed effects model was performed for the end term M-PARS score for the three conditions. Gender, prior knowledge in biology, prior final exam score, and experimental conditions were included as covariates. A random intercepts model was also performed, but the random intercept variance was not statistically significant (residual variance = .130, variance RI (tutor group) = .024, Wald  $Z = 1.743$ ,  $p = .081$ ), indicating that there were no significant group effects of tutorial group. The results of the fixed effects model indicated no significant main or interaction effects. A more parsimonious fixed effects model with only main effects indicated a main effect for prior final exam score and biology, but not for condition 1 or 2 (see Table 6). Table 7 displays the estimated marginal means on the end term M-PARS score for the three conditions, adjusted for potential confounding due to gender and prior knowledge. In addition, separate fixed effects models were performed for the end term scores of the three subscales of the M-PARS. With regard to the main effect for prior final exam score, these models generated results similar to the results of the total end term M-PARS score. With regard to biology, however, a significant main effect was only found on the cognitive subscale score ( $t = 2.77$ ,  $p = .006$ ). Again, for all three subscales no significant main effect for conditions 1 and 2 were found.

Table 6

*Main Effects of the Fixed Effects Model on the End Term M-PARS Score of Condition 1 (Individual Reflection), 2 (Individual and Shared Reflection), and 3 (Control Group)*

Variable	B	SE	t	Sig.
Intercept	2.32	.35	6.71	$p < .001$
Prior final exam	.02	.004	3.78	$P < .001$
Biology	.12	.06	2.03	.044
Gender	-.06	.07	-.92	.357
Condition 1	-.09	.07	-1.23	.219
Condition 2	-.001	.07	-.01	.990

Table 7

*Estimated Marginal Means on the End Term M-PARS Score for Condition 1 (Individual Reflection), 2 (Individual and Shared Reflection), and 3 (Control Group) Adjusted for Potential Confounding due to Gender and Prior Knowledge*

<b>M-PARS Score</b>	<b>Estimated Marginal Mean</b>	<b>95% Confidence Interval</b>
Condition 1	3.55	3.46 – 3.65
Condition 2	3.64	3.53 – 3.74
Condition 3	3.64	3.54 – 3.74

Another fixed effects analysis was performed on the end term M-PARS score comparing only conditions 1 (individual reflection) and 2 (individual and shared reflection) (see Table 8). Here again, there was only a main effect for biology and prior final exam score. Fixed effects models were also performed for the separate subscales. These results were in line with the results of the total M-PARS score. For the cognitive subscale significant main effects were found for biology and prior final exam score. For the collaborative and motivational subscales significant main effects were only found for prior final exam score.

Table 8

*Main Effects of the Fixed Effect Model on the End Term M-PARS Score of the Conditions 1 (Individual Reflection) and 2 (Individual and Shared Reflection)*

<b>Variable</b>	<b>B</b>	<b>t</b>	<b>Sig.</b>
Intercept	2.22	36.87	$p < .001$
Prior final exam	.02	3.43	.001
Biology	.15	2.03	.045
Gender	-.14	-1.73	.087
Condition 1	-.07	.98	.330

- RQ2. Does midterm peer process feedback combined with reflection and goal setting improve *students' individual achievement* within the tutorial group, and is this effect influenced by how reflection is facilitated (individual vs. individual and shared) and students' prior knowledge?

### **Achievement**

A fixed effects model was performed on the final MCQ exam score with prior knowledge in biology, experimental condition, prior final exam score, and gender as covariates in order to answer the research question (see Table 9). A random intercepts model was also performed, but the random intercept variance was not statistically significant (residual variance = 71.142, variance RI (tutor group) = 6.153, Wald  $Z = 1.238$ ,  $p = .216$ ), indicating that there were no significant group effects of tutorial group. The fixed effects model resulted in a main effect for biology and prior final exam score. Students with prior knowledge in biology on average had a 4.1% higher score on their final MCQ exam and students with a higher score on the final exams of the previous course had a higher score on the final exam of this course. This analysis also indicated a main effect for conditions 1 and 2. Students in condition 1 (individual reflection) had a 6.9% higher score than the students in condition 3 (control group) and condition 2 (individual and shared reflection) combined. Students in condition 2 (individual and shared reflection), in turn, had a 7.0% higher score than the students in condition 3 (control group) and condition 1 (individual reflection) combined. In addition to these main effects, the analysis also indicated an interaction effect between condition 1 and gender; within condition 1 (individual reflection) females performed 7.4% lower than male students. Another interaction was found between condition 2 and gender, indicating that a similar result is seen for condition 2. In this condition, females performed 7.7% lower than the males.

Table 9

*Main and Interaction Effects of the Fixed Effect Model on the Final Exam (MCQ) Scores (0 to 100%) of Condition 1 (Individual Reflection), 2 (Individual and Shared Reflection), and 3 (Control Group)*

<b>Variable</b>	<b>B</b>	<b>SE</b>	<b>t</b>	<b>Sig.</b>
Intercept	7.76	7.13	1.09	.28
Prior final exam	0.71	.09	8.19	<i>P</i> <.001
Biology	4.06	1.20	3.39	.001
Gender	3.20	2.16	1.48	.141
Condition 1	6.92	2.81	2.46	.015
Condition 2	7.04	2.48	2.84	.005
Gender*Condition 1	-7.38	3.24	-2.28	.024
Gender*Condition 2	-7.74	3.05	-2.54	.012

An additional fixed effects model was performed on the final exam score comparing only conditions 1 (individual reflection) and 2 (individual and shared reflection) (see Table 10). Here similar results were found. Prior final exam score and biology had a significant effect and, in both conditions, males performed better than females. Table 11 displays the estimated marginal means for final exam score for the three conditions, adjusted for potential confounding due to gender and prior knowledge.

Table 10

*Main Effects of the Fixed Effect Model on the Final Exam (MCQ) Scores (0 – 100%) for Conditions 1 (Individual Reflection) and 2 (Individual and Shared Reflection)*

<b>Variable</b>	<b>B</b>	<b>SE</b>	<b>t</b>	<b>Sig.</b>
Intercept	15.20	8.13	1.87	.064
Prior final exam	.71	.10	7.01	$p < .001$
Biology	3.51	1.44	2.43	.017
Gender	-4.43	1.58	-2.80	.006
Condition 2	.11	1.39	.082	.935

Table 11

*Estimated Marginal Means on the Final Exam (MCQ) Scores (0 to 100%) for Condition 1 (Individual Reflection), 2 (Individual and Shared Reflection), and 3 (Control Group) Adjusted for Potential Confounding due to Gender and Prior Knowledge*

<b>Final Exam Score</b>	<b>Estimated Marginal Mean</b>	<b>95% Confidence Interval</b>
Condition 1	70.40	68.49 – 72.30
Condition 2	70.26	68.19 – 72.34
Condition 3	68.78	66.77 – 70.78

## DISCUSSION

This experimental study investigated the effect of midterm peer process feedback on the quality of students' contributions and achievement. In addition, two ways of reflection were compared: individual reflection and shared reflection.

With regard to the first research question, it can be concluded that the peer process feedback did not improve the quality of students' contributions. Students did not display an increase in the quality of their contributions to the group process three weeks after receiving the peer feedback. Specifically on the collaborative subscale, females even showed a small negative gain score, indicating a small deterioration. With regard to the second research question, however, results indicate that the peer process feedback does enhance student achievement. Students within conditions 1 (individual reflection) and 2 (individual and shared reflection) received higher grades on the MCQ final exam than students in the control group. In addition, the final MCQ exam was affected by prior knowledge; students with prior knowledge achieved higher. The interaction effect between gender and condition, however, indicated that only males within both experimental conditions achieved higher. The way the peer feedback was facilitated (individual vs. individual and shared) did not appear to affect achievement, indicating that both ways of reflection were equally effective.

The finding that the M-PARS gain score was not affected by the peer feedback is surprising and partially contradicts the findings of Phielix et al. (2011), who did find that peer feedback combined with reflection and goal setting enhanced students' contributions to the social climate within the group. The findings of the current study could be attributed to the fact that students had only three tutorial meetings left to put this feedback into action. It might be that students need more meetings to change their behavior in the tutorial group. It is therefore surprising that, for males, achievement is enhanced by peer feedback. This could be explained by the self-regulated learning model of White and Gruppen (2010). In light of this model one could hypothesize that, by receiving the peer feedback and reflecting on it, students only adjusted their learning with regard to the subject matter. The remaining time might have been too short for students to actually change their behavior in the tutorial group. A similar explanation originates from the self-determination theory of Ryan and Deci (2000). In the current study it might be that receiving the peer feedback enhanced feelings of relatedness within the group, which in turn affected students' motivation and learning. This hypothesis is substantiated by the findings of a prior study on the effectiveness of peer feedback, in which students indicated that receiving the peer feedback enhanced their feelings of relatedness to the tutorial group (Kamp et al., 2013). Another explanation for the lack of an increase in M-PARS gain score could lie in the fact that students have received training in giving and receiving peer feedback at

the beginning of the bachelor program, which might have caused students to be already aware of their behavior within the tutorial group, making them less appreciative to the peer feedback. In addition, at the time of the study, students had had six months to develop their skills in participating in a PBL tutorial. This might also have caused them to be more or less set in their ways and therefore less susceptible to change.

The fact that this increase in achievement was only seen for males is surprising. There are several hypothetical reasons to explain this difference between males and females. The first explanation could be that there were initial differences between the three conditions or between males and females. However, the four fixed effects models on the pretest indicated that there were no significant differences between males and females and between the three conditions on the M-PARS total and subscale scores. Therefore this first explanation can be rejected. The second explanation could be that the difference is due to the nature of the intervention. The fact is that, for male students, motivation is more externally driven (Vallerand & Bissonette, 1992). Therefore, the competitive element of the numerical peer feedback might have affected the motivation of male students stronger than the motivation of female students. Subsequently, this increased motivation influenced learning (Ryan & Deci, 2000). A somewhat similar explanation could be that females are already quite self-aware of their contributions within the tutorial group and that the peer feedback was less of an eye opener for them than for the males, causing the peer feedback to be relative ineffective for the females. A third possible explanation is that males not so much do well with peer feedback, but that females do worse. Because females are more internally motivated than males (Vallerand & Bissonette, 1992), it could be that females are more intimidated by the public peer feedback, especially if it concerns negative peer feedback, which causes them to perform worse than before the feedback intervention. The fact that females showed a small deterioration in gain score on the collaborative subscale also indicates to this explanation.

The question also remains why there was no difference between the effectiveness of both experimental conditions. A possible explanation for this is that the shared reflection and goal setting intervention was not strong enough and, therefore, did not affect the quality of contributions and achievement. Another explanation is that in the current educational program tutors are already encouraged to regularly discuss student behavior within the tutorial group. Therefore, the shared reflection might not have provided an added value on top of the individual reflection

and goal setting.

Even though this study has provided rich insight in the process of peer feedback, it also has some limitations. The first one is that tutors might have influenced the quality of reflection and goal setting. Although tutors were provided with a protocol describing how to process the feedback, reflection, and goal setting, it cannot be determined with complete certainty that there were no individual differences in how the tutors handled the feedback. Quantitative analyses of the handed-in reflection and goal setting forms, however, indicate that the large majority of the students reflected on the received feedback and formulated improvement goals. A second limitation is that there were a limited number of tutorial meetings between receiving the feedback and the end term M-PARS evaluation. Had students had more tutorial meetings between these two events, they would probably have had more opportunities to improve their contributions to the tutorial group. A third limitation is that the written suggestions for improvement could have also had a significant influence on the outcome. Because both intervention conditions received the same suggestions for improvement this could have caused the fact that no difference was found between both conditions.

This study has also generated suggestions for future research. The first suggestion is that it should focus on providing students with more opportunities to work on the received feedback and improve their contributions. A longer collaboration period might be needed to enhance the effectiveness of peer feedback (Phielix, 2012). Another focus for future research is the quality of reflection and goal setting. Relating the quality of reflection and goal setting to the effectiveness of peer feedback might increase the understanding of the feedback process. Such a study, however, entails some difficulties. That is, reflection is a metacognitive process which can only be evaluated by its derivatives. This makes it very difficult to develop a reliable method to assess quality of reflection (Koole et al., 2011). Finally, future studies could focus on the relationship between giving/receiving peer feedback and relatedness and subsequently intrinsic motivation within the tutorial group. These studies could demonstrate whether these variables play an underlying role in the relationship between peer feedback and individual contributions and achievement in PBL groups. Although the M-PARS was developed as a peer rating scale (Kamp et al., 2011), a last suggestion for future research would be to investigate whether tutors can also use this tool to evaluate the quality of students' contributions within the tutorial group. In

addition, future research could possibly compare the peer evaluations with the tutor evaluations to investigate whether there are interesting differences.

## REFERENCES

- Adcroft, A. (2011). The mythology of feedback. *Higher Education Research & Development*, 30(4), 405-419. doi: 10.1080/07294360.2010.526096
- Archer, J. C. (2010). State of the science in health professional education: Effective feedback. *Medical Education*, 44, 101-108. doi: 10.1111/j.1365-2923.2009.03546.x
- Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning: An integrative model for effective team learning in organisations. *Educational Research Review*, 5, 111-133. doi: 10.1016/j.edurev.2010.02.002
- Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (1998). Motivational and cognitive processes influencing tutorial groups. *Academic Medicine*, 73(Suppl 10), S22-S24.
- Dominick, P. G., Reilly, R. R., & McGourty, J. W. (1997). The effects of peer feedback on team member behavior. *Group & Organization Management*, 22(4), 508-520. doi: 10.1177/1059601197224006
- Eva, K. W., Armson, H., Holmboe, E., Lockyer, J., Loney, E., Mann, K., et al. (2012). Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. *Advances in Health Sciences Education*, 17(1), 15-26. doi: 10.1007/s10459-011-9290-7
- Geister, S., Konradt, U., & Hertel, G. (2006). Effects of process feedback on motivation, satisfaction, and performance in virtual teams. *Small Group Research*, 37(5), 459-489. doi: 10.1177/1046496406292337
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3-31. doi: 10.1007/978-3-8348-9837-1
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, 20(4), 304-315. doi: 10.1016/j.learninstruc.2009.08.007
- Gijsselaers, W. H., & Schmidt, H. G. (1990). Development and evaluation of a causal model of problem-based learning. In Z. Nooman, H. G. Schmidt, & E. Ezzat (Eds.), *Innovation in medical education: an evaluation of its present status* (pp. 95-113). New York: Springer Publishing.

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. doi: 10.3102/003465430298487
- Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2011). Can students adequately evaluate the activities of their peers in PBL? *Medical Teacher*, 33(2), 145-150. doi: 10.3109/0142159X.2010.509766
- Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2012). The relationship between students' small group activities, time spent on self-study, and achievement. *Higher Education*, 64(3), 385-397. doi: 10.1007/s10734-011-9500-5
- Kamp, R. J. A., Dolmans, D. H. J. M., Van Berkel, H. J. M., & Schmidt, H. G. (2013). The Effect of Midterm Peer Feedback on Student Functioning in Problem-Based Tutorials. *Advances in Health Sciences Education*, 18(2), 199-213. doi: 10.1007/s10459-012-9364-1.
- Koole, S., Dornan, T., Aper, L., Scherpbier, A., Valcke, M., Cohen-Schotanus, J., et al. (2011). Factors confounding the assessment of reflection: A critical review. *BMC Medical Education*, 11, 104. doi: 10.1186/1472-6920-11-104
- Papinczak, T., Young, L., & Groves., M. (2007). Peer assessment in problem-based learning: a qualitative study. *Advances in Health Sciences Education*, 12, 169-186. doi: 10.1007/s10459-005-5046-6
- Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior*, 27(3), 1087-1102. doi: 10.1016/j.chb.2010.06.024
- Phielix, C. (2012). Enhancing collaboration through assessment & reflection. Unpublished Doctoral dissertation. University of Utrecht, Utrecht, the Netherlands.
- Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles, and preferences. *Advances in Health Sciences Education*, 11(3), 289-303. doi: 10.1007/s10459-005-3250-z
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and facilitation of intrinsic motivation, social development and well-being. *American Psychologist*, 55, 68-78. doi: 10.1037/0003-066X.55.1.68
- Sargeant, J., Mann, K., Sinclair, D., Van der Vleuten, C., & Metsemakers, J. (2008). Understanding the influence of emotions and reflection upon multi-source

- feedback acceptance and use. *Advances in Health Sciences Education*, 13(3), 275-288. doi: 10.1007/s10459-006-9039-x
- Schönrock-Adema, J., Heijne-Penninga, M., Van Duijn, M. A. J., Geertsma, J., & Cohen-Schotanus, J. (2007). Assessment of professional behaviour in undergraduate medical education: peer assessment enhances performance. *Medical Education*, 41(9), 836-842. doi: 10.1111/j.1365-2923.2007.02817.x
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. M. (2003). Cooperative learning and achievement: Theory and research. (In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology* (Vol. 7, pp. 177-198). Hoboken, NJ: John Wiley & Sons Inc.)
- Sluijsmans, D. M. A., Brand-Gruwel, S., & Van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, 27(5), 443-454. doi: 10.1080/0260293022000009311
- Sluijsmans, D. M. A., Moerkerke, G., Van Merriënboer, J. J. G., & Dochy, F. J. R. C. (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation*, 27, 153-173.
- Strijbos, J.-W., & Sluijsmans, D (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction*, 20, 265-269. doi: 10.1016/j.learninstruc.2009.08.002
- Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, 20(4), 339-343. doi: 10.1016/j.learninstruc.2009.08.003
- Vallerand, R.J., & Bissonnette, R. (1992). Intrinsic, extrinsic, and arnotivational styles as predictors of behavioral: A prospective study. *Journal of Personality*, 60(3), 599-620.
- Van den Bossche, P., Gijselaers, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments. Team learning beliefs & behaviors. *Small Group Research*, 37(5), 490-521. doi: 10.1177/1046496406292938
- Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review*, 4, 41-54. doi: :10.1016/j.edurev.2008.11.002

- Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction, 20*(4), 270-279. doi: 10.1016/j.learninstruc.2009.08.004
- White, C. B., & Gruppen, L. D. (2010). Self-regulated learning in medical education. In: T. Swanwick (Ed.), *Understanding medical education: Evidence, theory and practice* (pp. 271-282). London, UK: Wiley-Blackwell.



# **CHAPTER 6**

## General Discussion



In this chapter the main findings of the studies of this dissertation will be summarized and discussed. Furthermore, theoretical implications, strengths, and limitations will be considered. Finally, suggestions for future research and practical implications will be outlined.

## INTRODUCTION

The success of PBL depends largely on the quality of the group discussion (Nieminen, Sauri, & Lonka, 2006). Then again, the quality of this discussion is affected by cognitive, social, and motivational processes (e.g., Slavin, Hurley, & Chamberlain, 2003). In theory, students within tutorial groups should be involved in a profound discussion through which they collaboratively construct new knowledge. In addition, there needs to be a pleasant social climate within the group and a sense of 'team spirit'. With regard to this tutorial discussion there is, however, a gap between theory and practice. In everyday tutorial groups the discussion on occasion remains superficial (Moust, Van Berkel, & Schmidt, 2005) and students withdraw themselves from the group or sponge on other group members (Decuyper, Dochy, & Van den Bossche, 2010). As a result, essential group processes (i.e., cognitive, social, and motivational processes) become eroded and students' individual achievement might be affected negatively. So far, it is unknown to what extent students' *individual* contributions to these tutorial group processes are related to individual achievement.

Although this relationship deserves to be investigated, the problem is that a suitable tool with which these individual contributions can be measured is missing. Existing questionnaires developed to measure the tutorial group processes are often based on students' self-assessment of their functioning within the tutorial group (e.g., Visschers-Pleijers, Dolmans, Wolfhagen, & Van der Vleuten, 2005), whereas it is known that the validity and reliability of self-evaluations is not optimal (Eva & Regehr, 2005). Peer evaluations are a preferable alternative, since research has indicated that they produce more reliable and valid judgments compared with self-evaluations (Eva, 2001). Unfortunately, such a peer evaluation scale, designed specifically for the PBL context, is still lacking. Existing peer evaluation scales (e.g., Loyens, Rikers, & Schmidt, 2007; Rangachari & Crankshaw, 1992; Van Berkel & Van Til, 1994; Visschers-Pleijers et al., 2005) either do not focus on the contributions of *individual* students, do not

discriminate between different types of contributions (*cognitive, social, and motivational*), or are not developed and/or validated as a peer evaluation scale.

This problem generated the first main research question of this dissertation:

*Are students able to evaluate the cognitive, social, and motivational contributions to the tutorial group of their peers in a reliable and valid way and, to what extent do these contributions predict student achievement?*

In order to enhance the success of PBL and, therefore, the quality of the discussion, students need to contribute actively to the group processes. Feedback generated from peer evaluations is a promising tool to enhance students' individual contributions to the group processes in PBL (Dominick, Reilly, & McGourty, 1997) and subsequent achievement (Kluger & DeNisi, 1996). However, the conditions under which peer feedback enhances individual contributions and student achievement need to be further explored through experimental research (Gielen, Peeters, Dochy, Onghena, & Struyven, 2010; Prins, Sluijsmans, & Kirschner, 2006; Van Zundert, Sluijsmans, & Van Meriënboer, 2010). Reflection upon the received feedback and goal setting have been indicated as important conditions for effective peer feedback, but it remains unclear how this is optimally facilitated. The second main research question of this dissertation focused on this problem:

*Do reflection and goal setting enhance the effectiveness of a peer feedback intervention, aimed at evaluating students' individual contributions to the tutorial group?*

These two main research questions were addressed with four separate studies, which were all conducted in health sciences higher education.

## **CONCLUSIONS AND SCIENTIFIC IMPLICATIONS**

The first research question was addressed in Chapters 2 and 3. Chapter 2 focused on the first part of this research question; Are students able to evaluate the contributions of their peers within the tutorial group in a reliable and valid way? In order to answer this question the Maastricht-Peer Activity Rating Scale (M-PARS) was

developed. The contents of this instrument were based on literature about effective tutorial group processes and the items represented three subscales (cognitive, social, and motivational contributions). To answer the first research question a confirmatory factor analysis was performed and generalizability studies were conducted. The confirmatory factor analysis demonstrated a reasonably good fit between the data and the three-factor model. Generalizability studies demonstrated that, at least, four peer evaluations per student in the tutorial group were needed in order to produce a judgment with good internal consistency. Furthermore, good construct reliability was demonstrated with Hancock's coefficients. Based on these analyses it was concluded that students are indeed able to produce reliable and valid evaluations about the quality of students' cognitive, social, and motivational contributions to the tutorial group. This conclusion is in line with earlier conclusions that students are able to evaluate the performance of their peers in a reliable and valid manner (Eva & Regehr, 2011). The added value of this study is that it has generated a suitable peer evaluation instrument; a 14-item peer rating scale called the M-PARS that measures students' individual cognitive, social, and motivational contributions to the tutorial group discussion.

The study described in Chapter 3 attempted to clarify the unknown relationship between the quality of students' individual contributions (cognitive, social, and motivational), time spent on self-study, and student achievement. A structural equation modeling analysis was conducted and the results of this study demonstrated a relationship between the quality of students' cognitive, social, and motivational contributions and achievement. More specifically, a causal relationship was found between students' cognitive activities and his/her unit test score and between students' collaborative activities and the group assignment score. Time spent on self-study was hypothesized to function as a mediator between the quality of students' contributions and achievement. However, the results indicated that time spent on self-study was not affected by the quality of the contributions, nor did it affect student achievement. To conclude, the results of the study discussed in Chapter 3 suggest that there are indeed causal relations between a student's contributions to the group processes in the tutorial group and achievement. The findings of this study have, to a certain extent, elucidated the relationship between students' individual contributions and individual achievement in PBL.

The causal nature of the relationship between students' contributions and student achievement reported in Chapter 3 is in line with earlier studies in which these relationships have been investigated. Gijsselaers and Schmidt (1990) and Nieminen, Sauri, and Lonka (2006) indicated that the overall quality of group functioning has a direct effect on student achievement. Assuming that the quality of the students' individual contributions to the tutorial discussion affects the quality of the overall group functioning, the findings of Gijsselaers and Schmidt (1990) and Nieminen et al. (2006) strongly resemble the findings of the study addressed in Chapter 3. Chapter 3 has also indicated that individual student achievement is not only affected by processes at the level of the overall tutorial group, but also by the contributions to the tutorial discussion at the level of the individual student. Similar conclusions were drawn by De Koning, Loyens, Rikers, Smeets, and Van der Molen (2012), who found that the occurrence of students' observed learning activities (i.e. active participation and asking questions) was an important predictor for achievement in a PBL bachelor program. It is also in line with the findings of Loyens et al. (2007), who demonstrated a relationship between active participation in a PBL course and student achievement. Finally, Yew and Schmidt (2012) also found a relationship between the number of relevant contributions during the PBL cycle and achievement. The study addressed in Chapter 3 added to these earlier insights that evaluations of peers about the students' individual contributions to the group are related to individual achievement. The specific relations that were found in this study are, however, small, but this can be explained by the fact that earlier studies have indicated that achievement in PBL, and in higher education in general, is affected by many different factors at once (De Koning et al., 2012; Schmidt & Moust, 2000). Specifically within PBL, achievement is also known to be affected by students' amount of prior knowledge, quality of problems discussed, and the performance of the tutor (Schmidt & Moust, 2000). Therefore, the fact that these relations are small is not surprising or unexpected.

The second main research question was addressed in Chapters 4 and 5. Chapter 4 focused on the conditions under which the M-PARS could be used to enhance the quality of students' contributions to the tutorial group. Because earlier studies emphasized the importance of reflection and goal setting in order for peer feedback to have an effect (e.g., Archer, 2010; Gibbs & Simpson, 2004; Phielix, Prins, & Kirschner, 2010; Prins et al., 2006), the study discussed in Chapter 4 especially investigated the effectiveness of reflection and goal setting after receiving peer

feedback. In order to investigate the effects of peer feedback combined with reflection and goal setting, a mixed-methods explanatory study was conducted in which a pretest-intervention-posttest was combined with a focus group. The intervention entailed that students received peer feedback based on the M-PARS and were stimulated to individually reflect on the received peer feedback and formulate improvement goals. Based on their pretest score on the M-PARS, students were classified as initially low, moderate, or high contributing students. Gain scores (posttest – pretest M-PARS score) of the three types of students were analyzed. In addition, a focus group was conducted with students in order to investigate the perceived effectiveness of the peer feedback combined with reflection and goal setting.

It was concluded that students who contributed poorly at the pretest showed a small, but significant increase in the quality of their contributions after receiving the peer feedback. For students who contributed on an average or high level at the pretest no significant differences were found between the pre- and posttest. During the focus group, students indicated that the intervention had had a positive effect on their sense of belonging to the group and had made them more aware of what is considered effective tutorial behavior, but that the reflection and goal setting could be further emphasized with a face-to-face dialogue after receiving the peer feedback. The results and conclusions discussed in Chapter 4 are in line with earlier studies. Similar to the significant gain scores of the initially low contributing students in study 3, Gielen et al. (2010), for instance, found that students with an average to low proficiency level benefitted more from certain cognitive characteristics of peer feedback on a writing task than high performing students. The results of the focus group are in concurrence with findings of Papinczak, Young, and Groves (2007) and Schönrock-Adema, Heijne-Penninga, Van Duijn, Geertsma, and Cohen-Schotanus (2007). In both of these studies students indicated that peer feedback increased their awareness of criteria of appropriate tutorial behavior.

Chapter 5 attempted to further clarify the conditions under which peer feedback is effective. More specifically, this study focused on the effectiveness of individual versus shared reflection about the received peer feedback, because students indicated in Chapter 4 that reflection on the peer feedback needed more emphasis by means of a face-to-face dialogue in order to be effective. Therefore, two experimental conditions and one control condition were compared in a pretest-posttest design. In the first experimental condition students were stimulated to reflect on the received

peer feedback and formulate goals for improvement individually; the individual reflection experimental condition. In the second experimental condition students were also stimulated to reflect and formulate goals for improvement individually, and in addition, they were stimulated by the tutor to ask their peers for clarifications and explanations with regard to the received peer feedback. Next, students were asked to share their reflections and goals with their peers and, if necessary, adjust them. This second experimental condition was called the shared reflection condition. Students within the control group participated only in the posttest. Students within the second shared reflection condition were expected to benefit more from the peer feedback than students within the first individual reflection condition and within the control group. In addition to the quality of students' individual contributions, the effect of peer feedback on student achievement was also investigated.

The results of the study addressed in Chapter 5 demonstrated that student achievement increased significantly for male students in both experimental conditions as compared to the control condition, but there was no difference between both experimental groups in achievement. The quality of students' individual contributions was not affected by the peer feedback intervention in general, regardless of how reflection and goal setting was facilitated (i.e., either individually or shared). The absence of positive effects of individual and shared reflection on students' contributions is surprising since it is contradictory to earlier studies, performed in Computer Supported Collaborative learning (CSCL) environments, that reported positive effects of a similar feedback intervention on students' individual contributions (Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011; Zumbach, Muehlenbrock, Jansen, Reimann, & Hoppe, 2002; Zumbach, Hillers, & Reimann, 2004). The fact that the second experimental condition (shared reflection) did not have a superior effect over the first experimental condition (individual reflection) was unexpected, especially because students indicated in Chapter 4 that they felt the peer feedback intervention would be more effective if it were combined with a face-to-face dialogue after receiving the peer feedback. In addition, Gielen et al. (2010) found that peer feedback combined with justification (i.e. explanation of judgment or rating) had a positive effect on performance in a writing task. The lack of improvement in the quality of students' individual contributions to the tutorial discussion could be explained by the short time span of the intervention and the fact that students only completed one rating during

this period. Phielix (2012) recommends multiple ratings over a longer period of time in order for peer ratings to be effective.

Another interesting finding in this study was that there was a difference between males and females with regard to the effect the peer feedback had on their achievement. For males, the peer feedback appeared to have an effect on their achievement, but for females this effect was absent. A possible explanation for this gender difference could be that women are already more self-aware of their behavior in the tutorial group and are, therefore, less affected by the peer feedback, whereas the male students became more aware of the gap between their actual and desired behavior in the group due to their reflections on the peer feedback.

## **STRENGTHS AND LIMITATIONS**

The studies described in this dissertation hold several strengths as well as limitations. Both will be discussed in this paragraph, beginning with the strengths.

### **Strengths**

First of all, this dissertation has used peer ratings instead of self-ratings. In previous studies the problem was that the evaluation of the quality of the interactions within the tutorial group was often based on students' own input. Therefore, it concerned a self-evaluation, and these self-evaluations suffer from validity and reliability problems (Eva & Regehr, 2005). As was explained earlier in this dissertation, peer evaluations have demonstrated greater reliability and validity, and therefore can generate a more reliable evaluation of the quality of interaction in PBL tutorial groups (Eva, 2001; Papinczak, Young, Groves, & Haynes, 2007). In addition, a peer evaluation concerns a judgement that is based on multiple peer observations (Eva, 2001) and it stimulates students' self-directed learning (Sluijsmans, Dochy, & Moerkerke, 1998). Therefore, the fact that peer evaluations were used to measure the quality of students' contributions can be considered a strength of this dissertation.

The second strength of this dissertation is that it has added experimental research to the field of peer feedback, and especially with regard to the conditions under which peer feedback is effective. According to Van Zundert et al. (2010), there is a need for experimental research on the conditions under which peer feedback is

effective. This dissertation has contributed to this appeal, and has indicated that especially the face-to-face discussion, after the feedback is received, is important and does enhance awareness of desired behaviours in the group and the sense of belonging to the group.

### **Limitations**

In addition to these strengths, this dissertation also has several limitations. The three main limitations will be addressed.

The first drawback concerns the time span of the peer feedback intervention which was rather limited in the studies described in Chapters 4 and 5; i.e. only one midterm peer feedback session within a period of seven weeks and seven tutorial group sessions. Phielix (2012) suggests that, in order to realize an enhancement of the cognitive processes in (computer supported) collaborative learning, peer feedback needs to be administered during a long time span and this feedback needs to be given at multiple time points. This offers students the chance to get to know each other better, observe more of each other's behavior, and it provides them with more opportunities to use the peer feedback to enhance their behavior. Due to the structure of the curriculum it was, however, not possible to administer the peer feedback intervention during a unit with a longer time span. Due to the invasive and time consuming nature of the intervention it was, also, not possible to administer the peer feedback at multiple time points.

The second limitation is related to the quantitative nature of this dissertation. With the exception of a part of Chapter 4, all studies can be classified as quantitative research. Although it is considered a strength that this dissertation has taken an experimental approach to the conditions under which effectiveness peer feedback is effective, it can also be seen as a limitation. Qualitative research could have provided a more profound and explanatory insight into the working of peer feedback and how it can enhance learning. More specifically, the quality of reflection through dialogue after receiving peer feedback could be further investigated and possibly enhanced.

The third limitation is the fact that the M-PARS was used as an intervention and as a dependent variable at the same time. One could argue that solely filling in the M-PARS might already increase students' awareness of appropriate behavior in the tutorial group and possibly the quality of their contributions. However, in everyday educational practice, peer feedback cannot originate out of nothingness; in order to

provide peer feedback students first have to rate each. Therefore, in this dissertation the intervention was seen as an integer whole.

## **FUTURE RESEARCH**

The research described in this dissertation has generated several suggestions for future research.

The first suggestion results from the first main limitations of this research: the short time span of the peer feedback intervention. In a next study the effect of receiving peer feedback on the quality of students' contributions and their achievement should be investigated during a unit with a longer time span than the ones in the current studies. In addition it is preferred to administer and address the received peer feedback at multiple times. This will provide students with more opportunities to enhance and monitor their contributions within the tutorial group.

The second suggestion for future research originates from the second limitation. Future research needs to focus on the factors that affect the receptivity of peer feedback. Peer feedback will only be effective if it is received well; meaning that students are willing to process the feedback and to use it to improve their contributions. Although it is difficult to measure quality of reflection (Koole et al., 2011), future research should investigate how the quality of reflection is related to the receptivity of the peer feedback. Furthermore, future research should explore how a face-to-face dialogue can enhance the quality of reflection and the receptivity of peer feedback.

## **PRACTICAL IMPLICATIONS**

This research has demonstrated that, in PBL tutorial groups, students are able to evaluate the quality of the contributions of their peers. In addition, it is important to address the quality of these contributions through peer feedback because it is related to student achievement. The studies addressed in Chapters 4 and 5 have generated several practical points of interest that need to be considered when one wants to implement peer feedback in PBL. These implications can be completed with

implications produced by earlier research (i.e., Hattie & Timperley, 2007; Prins et al., 2006; Sluijsmans, Brand-Gruwel, & Van Merriënboer, 2002). This results in six points of interest.

First, students need to be trained before they are able to provide their peers with reliable and valid feedback. During this training criteria on which the students need to be evaluated need to be explained and illustrated with examples. Second, students should be stimulated to reflect upon the received peer feedback and set goals for improvement. Third, based on the qualitative results discussed in Chapter 4, peer feedback should be combined with a face-to-face discussion to address the received feedback. During this dialogue students can ask for clarification with regard to their feedback. The fourth point of interest is the time span and the frequency of the peer feedback intervention. When administering peer feedback, students should work collaboratively for a time span that provides students with ample opportunities to observe their peers and enhance and practice their contributions to the discussion within the tutorial group. The fifth point of interest is the anonymity of the peer feedback. In this dissertation the peer feedback was given anonymously, because it was expected that this would generate less socially desired and therefore more honest judgments. However, giving peer feedback with disclosed identity might have simplified the face-to-face discussion after receiving the peer feedback. When peer feedback is given publicly the giver is better able to justify the feedback and to provide clear examples. The receiver, on the other hand, is better able to ask for clarification and examples of the evaluated behavior. However, when peer feedback is given anonymously students are less inhibited to give honest ratings, especially since Sluijsmans et al. (2002) reported that students found it difficult to make negative judgments about their peers' performance in PBL tutorial groups. The preceding implies that both options should be considered when implementing peer feedback, but that public peer feedback might be favorable to anonymous peer feedback, provided that there is a safe atmosphere within the tutorial group. The sixth and last point of interest is that the peer feedback should be used in a formative matter and given midterm. Students should have the opportunity to process their feedback and improve their behavior.

## REFERENCES

- Archer, J. C. (2010). State of the science in health professional education: Effective feedback. *Medical Education*, 44, 101-108. doi: 10.1111/j.1365-2923.2009.03546.x
- De Koning, B. B., Loyens, S. M. M., Rikers, R. M. J. P., Smeets, G., & van der Molen, H. T. (2012). Generation Psy: Student characteristics and academic achievement in a three-year problem-based learning bachelor program. *Learning and Individual Differences*, 22(3), 313-323. doi: 10.1016/j.lindif.2012.01.003
- Eva, K. W. (2001). Assessing tutorial-based assessment. *Advances in Health Sciences Education*, 6, 243-257.
- Eva, K. W., & Regehr, G. (2005). Self-assessment in the health professions: A reformulation and research agenda. *Academic Medicine*, 80(10), S46-S54. doi: 10.1097/00001888-200510001-00015
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3-31. doi: 10.1007/978-3-8348-9837-1
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). The effectiveness of peer feedback for learning: The effects of constructiveness, accuracy and embedding in the learning environment. *Learning and Instruction*, 20(4), 304-315. doi: 10.1016/j.learninstruc.2009.08.007
- Gijsselaers, W. H., & Schmidt, H. G. (1990). Development and evaluation of a causal model of problem-based learning. In Z. Nooman, H. G. Schmidt, & E. Ezzat (Eds.), *Innovation in medical education: an evaluation of its present status* (pp. 95-113). New York: Springer Publishing.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. doi: 10.3102/003465430298487
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Koole, S., Dornan, T., Aper, L., Scherpbier, A., Valcke, M., Cohen-Schotanus, J. et al. (2011). Factors confounding the assessment of reflection: A critical review. *BMC Medical Education*, 11, 104. doi:10.1186/1472-6920-11-104

- Loyens, S. M. M., Rikers, R. M. J. P., & Schmidt, H. G. (2007). Students' conceptions of distinct constructivist assumptions. *European Journal of Psychology of Education, 12*, 179–199.
- Moust, J. H. C., Van Berkel, H. J. M., & Schmidt, H. G. (2005). Signs of erosion: Reflections on three decades of problem-based learning at Maastricht University. *Higher Education, 50*, 665–683. doi: 10.1007/s10734-004-6371-z
- Nieminen, J., Sauri, P., & Lonka, K. (2006). On the relationship between group functioning and study success in problem-based learning. *Medical Education, 40*(1), 64–71. doi: 10.1111/j.1365-2929.2005.02344.x
- Papinczak, T., Young, L., & Groves, M. (2007). Peer assessment in problem-based learning: a qualitative study. *Advances in Health Sciences Education, 12*, 169–186. doi: 10.1007/s10459-005-5046-6
- Papinczak, T., Young, L., Groves, M., & Haynes, M. (2007). An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Medical Teacher, 29*, e122–e132. doi: 10.1080/01421590701294323
- Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior, 26*(2), 151–161. doi: 10.1016/j.chb.2009.10.011
- Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior, 27*(3), 1087–1102. doi: 10.1016/j.chb.2010.06.024
- Phielix, C. (2012). *Enhancing collaboration through assessment & reflection*. Unpublished Doctoral dissertation. University of Utrecht, Utrecht, the Netherlands.
- Prins, F. J., Sluijsmans, D. M. A., & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles, and preferences. *Advances in Health Sciences Education, 11*(3), 289–303. doi: 10.1007/s10459-005-3250-z
- Rangachari, P. K., & Crankshaw, D. J. (1992). Formative evaluation in a tutorial-based course. Paper presented at the Proceedings of 5th Instructional Show and Tell for Ontario Universities, Guelph, ON.
- Schmidt, H. G., & Moust, J. H. C. (2000). *Factors affecting small-group tutorial learning: A review of research*. In D. H. Evensen & C. E. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions*. (pp. 19–52). Mahwah, NJ: Lawrence Erlbaum.

- Schönrock-Adema, J., Heijne-Penninga, M., Van Duijn, M. A. J., Geertsma, J., & Cohen-Schotanus, J. (2007). Assessment of professional behaviour in undergraduate medical education: peer assessment enhances performance. *Medical Education*, 41(9), 836-842. doi: 10.1111/j.1365-2923.2007.02817.x
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. (2003). *Cooperative learning and achievement: Theory and research*. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology*, Vol. 7. (pp. 177-198). Hoboken (NJ): John Wiley & Sons Inc.
- Sluijsmans, D. M. A., Brand-Gruwel, S., & Van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, 27(5), 443-454. doi: 10.1080/0260293022000009311
- Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, 1(3), 293-319. doi: 10.1023/a:1009932704458
- Van Berkel, H. J. M. & Van Til C. T. (1994). *Group performing in a problem-based curriculum: The development and evaluation of an instrument*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20(4), 270-279. doi: 10.1016/j.learninstruc.2009.08.004
- Visschers-Pleijers, A. J. S. F., Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van der Vleuten, C. P. M. (2005). Development and validation of a questionnaire to identify learning-oriented group interactions in PBL. *Medical Teacher*, 27(4), 375-381. doi: 10.1080/01421590500046395
- Yew, E. H. J. & Schmidt, H. G. (2012). What students learn in problem-based learning: A process analysis. *Instructional Science*, 40, 371-395. doi: 10.1007/s11251-011-9181-6
- Zumbach, J., Hillers, A., & Reimann, P. (2004). Supporting distributed problem-based learning: The use of feedback mechanisms in online learning. In T. S. Roberts (Ed.), *Collaborative learning: Theory and practice* (pp. 86-102). Hershey, PA: Information Science Publishing.

Zumbach, J., Muehlenbrock, M., Jansen, M., Reimann, P., & Hoppe, H.-U. (2002).  
Multidimensional Tracking in Virtual Learning Teams. In G. Stahl (Ed.),  
Computer support for collaborative learning: Foundations for a CSCL  
community (pp. 650-651). Hillsdale, NJ: Erlbaum.



**APPENDIX A**  
**Maastricht-Peer Activity Rating Scale (M-PARS)**

Group:

Evaluator

Name:

Student IDnumber:

Student to be evaluated

Name:

Student IDnumber:

Explanation:

With this rating scale the performance of your fellow student in the tutorial group can be evaluated. For each fellow student a separate rating scale has to be filled in. The name of the person you have to evaluate is written above. With regard to the results, no feedback will be given to the student you are going to evaluate. Instead they will be analyzed anonymously. When an item does not apply you can skip it. Thank you beforehand for your cooperation.

1	2	3	4	5
Completely disagree	Disagree	neutral	Agree	Completely agree

		1	2	3	4	5
<i>Cognitive activities</i>						
1	This student was able to make adequate summaries	<input type="radio"/>				
2	This student was able to make a distinction between the main and lateral issues in the subject matter	<input type="radio"/>				
3	This student asked critical questions	<input type="radio"/>				
4	This student corrected misconceptions about the subject matter	<input type="radio"/>				
5	This student contributed to a better understanding of the subject matter	<input type="radio"/>				

		1	2	3	4	5
<i>Social activities</i>						
6	This student had a positive influence on the group	<input type="radio"/>				
7	This student felt responsible for the group	<input type="radio"/>				
8	This student promoted collaboration between group members	<input type="radio"/>				
9	This student was willing to share his/her information	<input type="radio"/>				
10	This student was committed to the group	<input type="radio"/>				
<i>Motivational activities</i>						
11	This student demonstrated motivation	<input type="radio"/>				
12	This student participated well	<input type="radio"/>				
13	This student actively participated during the brainstorm sessions	<input type="radio"/>				
14	This student contributed more than other group members	<input type="radio"/>				

**APPENDIX B**  
**Topic list used for the focus group in chapter 4**

The central question presented to the students was:

*How do you perceive the effectiveness of the peer feedback and what are the impeding and promoting factors?*

This question was divided into four sub questions:

1. What do you think of the feedback? (timing, structure, anonymity, improvement points)
2. How did the feedback influence your individual and group functioning?
3. What were promoting and impeding factors for the effectiveness of the feedback?
4. Do you have suggestions for improvement?

## APPENDIX C

### Tips for improvement

	Definition	Improvement tips	Goals for personal improvement ( <i>min. 1, max. 3</i> )
Cognitive contributions	Contributions that contribute to the Construction of new knowledge and the recognition of misconceptions in one's own knowledge	<ol style="list-style-type: none"> <li>1. Summarize the answer to a learning goal in your own words during the discussion.</li> <li>2. Search for contradictions within the discussion and express these.</li> <li>3. Explain the subject matter with an example. Use an example from daily life and not from the literature.</li> <li>4. Identify differences and similarities between different concepts.</li> <li>5. Report your findings during the discussion without checking your notes.</li> <li>6. Indicate what is unclear to you or what you are in doubt of.</li> </ol>	
Collaborative contributions	Contributions that contribute to a good social climate within the group	<ol style="list-style-type: none"> <li>7. Make sure your answers or information is in keeping with the previous comments or question.</li> <li>8. Look someone in the eye when you are talking to them.</li> <li>9. Repeat long answers of group members shortly in your own words.</li> <li>10. Make sure you are well prepared when come to the tutorial group</li> </ol>	
Motivational contributions	Contributions that contribute to a good social climate within the group	<ol style="list-style-type: none"> <li>11. Be the first one to start the discussion by reporting your findings.</li> <li>12. Be the first one to start the brainstorm by telling what you already know.</li> <li>13. Adopt an active attitude during the tutorial Group discussion (sit up straight, hands on the table, open posture).</li> <li>14. Talk with a clear voice, watch your intonation.</li> </ol>	

**APPENDIX D****Example of a personal peer feedback form**

Dear student,

Below you will find an overview of the mean scores you received from your peers per item.

1 – 5 Likert scale	Grade (1-10)
- < 3.0: unsatisfactory (- -)	- < 6.0: unsatisfactory (- -)
- 3.0 – 3.5: open to improvement (-)	- 6.0 – 7.0: open to improvement (-)
- 3.5 – 4.0: average	- 7.0 – 8.0: average
- ≥ 4.0: good (+)	- ≥ 8.0: good (+)

The following standards were used: M: average judgment received from peers, SD: standard deviation

		<b>M(1-5)</b>	<b>SD</b>
<b>Cognitive activities</b>			
<b>1</b>	This student was able to make adequate summaries	3,4 (-)	0,5
<b>2</b>	This student was able to make a distinction between the main and lateral issues in the subject matter	3,8	0,5
<b>3</b>	This student asked critical questions	3,4 (-)	1,0
<b>4</b>	This student corrected misconceptions about the subject matter	3,1 (-)	0,6
<b>5</b>	This student contributed to a better understanding of the subject matter	3,4 (-)	0,5
<b>Social activities</b>			
<b>6</b>	This student had a positive influence on the group	3,5	0,5
<b>7</b>	This student felt responsible for the group	3,6	0,6
<b>8</b>	This student promoted collaboration between group members	3,6	0,6
<b>9</b>	This student was willing to share his/her information	4,0 (+)	0,8
<b>10</b>	This student was committed to the group	3,8	0,4
<b>Motivational activities</b>			
<b>11</b>	This student demonstrated motivation	3,5	0,7
<b>12</b>	This student participated well	3,6	0,5
<b>13</b>	This student actively participated during the brainstorm sessions	3,3(-)	0,7
<b>14</b>	This student contributed more than other group members	3,0 (-)	0,7
<b>Grade</b>		<b>M(1-10)</b>	<b>SD</b>
<b>15</b>	Score this student's overall activity in the tutorial group sessions on a scale from 1 (poor) tot 10 (excellent).	7,1	0,5





# SUMMARY

## SUMMARY

Chapter 1 addresses the aim and general research questions of this dissertation. The success of Problem-Based Learning (PBL) depends largely on the quality of the group discussion. The quality of this discussion is affected by cognitive, social, and motivational processes. The reality in PBL groups is that these processes are often eroded, which might also have a negative effect on student achievement. Peer feedback is a promising tool to enhance students' individual contributions to group processes and subsequent achievement. However, peer rating scales for PBL are scarce and do not measure contributions to the cognitive, social, and motivational processes. Lastly, little is known about the conditions under which peer feedback can enhance individual contributions and student achievement, especially if and how reflection and goal setting after receiving peer feedback can enhance its effectiveness. This results in the following two research questions: 'Are students able to evaluate the cognitive, social, and motivational contributions to the tutorial group of their peers and do these contributions predict student achievement?', and 'Do reflection and goal setting enhance the effectiveness of a peer feedback intervention, aimed at evaluating students' individual contributions to the tutorial group?'

In Chapter 2 a 14-item peer rating scale (M-PARS) is developed in order to investigate whether students are able to evaluate the cognitive, social, and motivational contributions of their peers. The items represented three subscales: cognitive, social, and motivational contributions to the tutorial group. One-hundred-and-ninety-six students were evaluated by their peers on these items. Validity and reliability analyses indicated that students are indeed able to evaluate the contributions of their peers in a valid and reliable manner. Generalizability studies indicated that one student needs to be rated by, at least, four peers in order to produce a consistent evaluation.

In Chapter 3 the relationship between the quality of students' contributions to the tutorial group, time spent on self-study, and subsequent academic achievement is investigated. Therefore, 538 students were evaluated with the M-PARS by their peers and the individual peer ratings were related to time spent on self-study and student achievement using structural equation modeling. Achievement was measured with a unit test on the subject matter that was addressed in the tutorial group meetings and with a group assignment. Results indicated an acceptable model fit. The score on the unit test was especially affected by the quality of a student's constructive contributions to the tutorial discussion. The score on the group assignment, on the other hand, was especially affected by the quality of a student's social contributions.

Time spent on self-study was not affected by the quality of students' contributions, nor did it affect student achievement. These results suggest that there are indeed causal relations between the quality of students' contributions to the tutorial group process and achievement.

Chapter 4 investigates whether reflection and goal setting enhance the effectiveness of a peer feedback intervention in PBL tutorial groups. In this study qualitative and quantitative methods were combined in a mixed-methods explanatory design. Seventy four students participated in a pretest (M-PARS), intervention (receiving peer feedback), and posttest (M-PARS and student achievement). Students were stimulated to reflect and formulate goals for improvement after receiving peer feedback. Based on their pretest score on the M-PARS participants were classified as poorly, average, or highly contributing students. Students' attitudes towards the effectiveness of peer feedback generated by the M-PARS were investigated with a focus group. The results indicated that the peer feedback intervention only increased the quality of contributions for students who initially contributed poorly. The qualitative data indicated that students feel that the peer feedback increased their sense of belonging to the group and their awareness of appropriate tutorial behavior. These results, however, also demonstrated that reflection and goal setting could be further stimulated by adding a face-to-face discussion which enables students to ask for clarifications and explanations regarding the received peer feedback.

Chapter 5 addresses a study in which the effects of *individual* versus *shared* reflection and goal setting on students' individual contributions to the group and their academic achievement were investigated. In this study, with a pretest – intervention – posttest design, 242 first year students were divided into three conditions: condition 1 (*individual reflection and goal setting*), condition 2 (*shared reflection and goal setting*), and condition 3 (control group). The results revealed that students in conditions 1 and 2 scored higher on the achievement test compared with students in condition 3, but this applied only to male students. However, both ways of reflection (*individual* versus *shared*) appeared to be equally effective, since there was no difference between conditions 1 and 2. Furthermore, the quality of individual contributions to the tutorial group did not improve after receiving the peer feedback, nor did it differ between the three conditions.

Finally, Chapter 6 provides a synopsis of the results and general conclusions of the studies addressed in chapters 2 to 5. Here, it is concluded that students' are able

to evaluate their peers' cognitive, social, and motivational contributions to the group, which are related to student achievement. Peer feedback on these contributions combined with reflection and goal setting can be used to enhance students' awareness of how they should contribute to the tutorial discussion and to increase their sense of belonging to the tutorial group. Such a peer feedback intervention can also enhance student achievement. Furthermore, Chapter 6 addresses the strengths and limitations of the research described in this dissertation. A strength of this dissertation is that it has contributed to the clarification of the conditions under which peer feedback is effective, by means of experimental research. An important limitation is the limited number of opportunities students had to improve their contributions. This might have affected the effectiveness of the peer feedback intervention. Therefore, future research should investigate whether a prolonged time-span of tutorial group meetings during which peer feedback is given several times can enhance the effectiveness of peer feedback. In addition, the effect of the quality of reflection and goal setting could be taken into consideration in future research. Finally, Chapter 6 presents the practical implications that need to be considered when implementing peer feedback in PBL tutorial groups. Because the quality of students' individual contributions is related to their achievement in PBL, peer feedback should be used in PBL practice to enhance learning. An important practical implication is that students should be stimulated to reflect and set goals for improvement during a face-to-face discussion after they have received the peer feedback.



# **SAMENVATTING**

Summary in Dutch



In Hoofdstuk 1 wordt het doel en de algemene onderzoeksvraag van dit proefschrift beschreven. Het succes van Probleem Gestuurd Leren (PGO) hangt voor een groot gedeelte af van de kwaliteit van de discussie in de onderwijsgroep. De kwaliteit van deze discussie wordt beïnvloed door cognitieve, sociale en motivationele processen. Echter, in de praktijk verlopen deze processen in de onderwijsgroep vaak niet optimaal, wat weer negatieve gevolgen kan hebben voor de leerprestaties van studenten. Peer feedback (i.e. het proces waarbij studenten feedback geven op het werk van medestudenten) is een veelbelovend instrument voor het verbeteren van individuele bijdragen van studenten tijdens de onderwijsgroep en de daarop volgende leerprestaties. Desondanks bestaan er slechts een klein aantal instrumenten waarmee studenten de kwaliteit van elkaars bijdragen in de onderwijsgroep kunnen evalueren. Het kleine aantal instrumenten waarmee dit wel kan maakt geen onderscheid tussen cognitieve, sociale en motivationele bijdragen van studenten. Ten slotte, is er weinig bekend over de condities waaronder peer feedback de individuele bijdragen en leerprestaties van studenten kan verbeteren, en in het bijzonder of en hoe reflectie en het stellen van doelen ter verbetering na het ontvangen van de peer feedback de effectiviteit hiervan kan versterken. Dit resulteert in de volgende twee onderzoeksvragen: ‘Zijn studenten in staat elkaars cognitieve, sociale en motivationele bijdragen tijdens de onderwijsgroep te evalueren?’ en ‘Versterken reflectie en het stellen van doelen de effectiviteit van een peer feedback interventie, gericht op het evalueren van de individuele bijdragen van studenten in de onderwijsgroep?’.

In hoofdstuk 2 is een peer rating scale (genaamd M – PARS), een instrument waarmee studenten elkaar kunnen beoordelen, ontwikkeld om te onderzoeken of studenten in staat zijn om de cognitieve, sociale en motivationele bijdragen van hun onderwijsgroepsgenoten te evalueren. De M-PARS bestaat uit 14 items die drie subschalen vertegenwoordigen; cognitieve, sociale en motivationele bijdragen aan de onderwijsgroep. Zeshonderd negentig studenten werden beoordeeld door hun collega's op deze items. Validiteits- en betrouwbaarheidsanalyses gaven aan dat studenten inderdaad in staat zijn om de bijdragen van hun groepsgenoten te evalueren op een valide en betrouwbare wijze. Generaliseerbaarheidsstudies gaven aan dat één student door minimaal vier groepsgenoten geëvalueerd moet worden om een consistente evaluatie te produceren.

In hoofdstuk 3 is de relatie tussen de kwaliteit van de bijdragen van de studenten tijdens de onderwijsgroep, tijd besteed aan zelfstudie en leerprestaties

onderzocht. Daarvoor werden 538 studenten door hun groepsgenoten geëvalueerd met behulp van M-PARS. De relaties tussen de beoordelingen die de studenten ontvingen, de tijd die ze besteed hebben aan zelfstudie en hun behaalde leerprestaties zijn onderzocht met behulp van structural equation modeling. De leerprestatie van studenten werd gemeten met een eindtoets over de leerstof die in de onderwijsgroepsbijeenkomsten was behandeld en met een groepsopdracht. Dit resulteerde in een passend model, waarin de score op de eindtoets vooral beïnvloed werd door de kwaliteit van constructieve bijdragen van een student aan de discussie. De score op de groepsopdracht, daarentegen, werd vooral beïnvloed door de kwaliteit van de sociale bijdragen van een student tijdens de onderwijsgroep.

De tijd die studenten besteden aan zelfstudie werd niet beïnvloed door de kwaliteit van de bijdragen van de studenten, noch had dit invloed op de leerprestaties van studenten. Deze resultaten suggereren dat er inderdaad causale relaties bestaan tussen de kwaliteit van de bijdragen van studenten aan de discussie in de onderwijsgroep en hun leerprestaties.

In Hoofdstuk 4 is onderzocht of reflectie op de ontvangen feedback en het stellen van doelen ter verbetering de effectiviteit van een peer feedback interventie in de onderwijsgroep vergroot. In deze studie werden kwalitatieve en kwantitatieve data gecombineerd in een 'mixed methods' verklarend ontwerp. Vierenzeventig studenten namen deel aan een voortest (evaluatie met behulp van de M-PARS), interventie (ontvangen van peer feedback) en nameting (evaluatie met behulp van de M-PARS en leerprestaties van studenten). Na het ontvangen van de peer feedback werden studenten gestimuleerd om te reflecteren op de ontvangen feedback en om doelen voor verbetering te formuleren. Op basis van hun M-PARS score op de voortest werden studenten geclassificeerd als onvoldoende, gemiddeld of sterk bijdragende student. De attitudes van studenten, met betrekking tot de effectiviteit van de peer feedback interventie, werden onderzocht met een focusgroep. De resultaten gaven aan dat de kwaliteit van de bijdragen van studenten ten gevolge van de interventie alleen maar toenam bij studenten die aanvankelijk onvoldoende bijdroegen. De kwalitatieve resultaten toonde aan dat de studenten het gevoel hadden dat de peer feedback het gevoel dat ze onderdeel van de (onderwijs)groep waren vergroten. Het maakten ze ook meer bewust van wat als passend en geschikt gedrag in de onderwijsgroep wordt gezien. De resultaten van de focusgroep gaven echter ook aan dat reflectie en het stellen van doelen ter verbetering verder gestimuleerd konden worden door

studenten in staat te stellen om, tijdens een gezamenlijke discussie, toelichtende en verduidelijkende vragen aan hun groepsgenoten te stellen ten aanzien van de ontvangen peer feedback.

In Hoofdstuk 5 wordt een studie beschreven waarin de effecten van individuele versus gezamenlijke reflectie en stellen van doelen op de kwaliteit van de individuele bijdrage van de studenten tijdens de onderwijsgroep en hun leerprestaties werden onderzocht. In deze studie, met een voortest - interventie - nameting ontwerp, werden 242 eerstejaars studenten verdeeld in drie groepen: groep 1 (individuele reflectie en stellen van doelen), groep 2 (gezamenlijk reflectie en stellen van doelen) en groep 3 (controlegroep). Uit de resultaten bleek dat studenten uit groep 1 en 2 hoger scoorden op de eindtoets in vergelijking met studenten uit groep 3, maar dit gold alleen voor mannelijke studenten. Beide manieren van reflectie en het stellen van doelen (individueel versus gezamenlijk) bleken echter even effectief te zijn, want er was geen verschil tussen groep 1 en 2. Bovendien was de kwaliteit van de individuele bijdragen tijdens de onderwijsgroep niet verbeterd na het ontvangen van de peer feedback, noch waren er verschillen tussen de drie groepen met betrekking tot deze kwaliteit.

Tot slot geeft Hoofdstuk 6 een overzicht van de resultaten en de algemene conclusies van de studies die aan bod zijn geweest in de hoofdstukken 2 tot en met 5. In dit hoofdstuk wordt geconcludeerd dat studenten in staat zijn om de cognitieve, sociale en motivationele bijdragen van hun groepsgenoten tijdens de onderwijsgroepsdiscussie te evalueren. Daarnaast kan geconcludeerd worden dat de kwaliteit van deze bijdragen gerelateerd zijn aan de leerprestaties van leerlingen. Peer feedback op deze bijdragen gecombineerd met reflectie en het stellen van doelen kan worden gebruikt om het bewustzijn van gewenst en optimaal gedrag tijdens de onderwijsgroep van studenten te verbeteren en om hun gevoel van het behoren bij de groep te verhogen. Een dergelijke peer feedback interventie kan ook tot verbetering van de leerprestaties van leerlingen leiden. Daarnaast gaat Hoofdstuk 6 in op de sterke en zwakke punten van de in dit proefschrift beschreven onderzoeken. Een sterk punt van dit proefschrift is dat het heeft bijgedragen, door middel van experimenteel onderzoek, aan de verduidelijking van de voorwaarden waaronder peer feedback effectief is. Een belangrijke beperking van het onderzoek beschreven in dit proefschrift is het kleine aantal onderwijsgroepsbijeenkomsten waarin studenten de kans hadden om de kwaliteit van hun bijdragen te verbeteren. Dit kan eventueel van invloed zijn

geweest op de effectiviteit van de peer feedback interventie. Daarom zou toekomstig onderzoek moet bekijken of het vergroten van het aantal onderwijsgroepsbijeenkomsten na het ontvangen van de peerfeedback leidt tot een verbetering van de effectiviteit van peer feedback. Bovendien komt het effect van de kwaliteit van reflectie en het stellen van doelen in aanmerking voor toekomstig onderzoek. Tot slot worden in Hoofdstuk 6 de praktische implicaties voor de uitvoering van peer feedback in de onderwijsgroep besproken. Omdat de kwaliteit van de individuele bijdragen van studenten tijdens de onderwijsgroep gerelateerd kan worden aan de leerprestaties van studenten, is het aan te raden om peer feedback te gebruiken bij de uitvoering van PGO om zo het leren te verbeteren. Een belangrijke praktische implicatie is dat studenten moeten worden gestimuleerd om te reflecteren en doelen voor verbetering te stellen tijdens een gezamenlijke bespreking van de ontvangen peer feedback.





# **DANKWOORD**

## Acknowledgements

DANKWOORD

En dan, na 5 jaar is mijn proefschrift af. Een last valt van mijn schouders, maar dan moet je nog een dankwoord schrijven en dat is toch nog lastiger dan ik van tevoren had gedacht. Want hoe schrijf je een origineel dankwoord zonder al die clichés? Het wordt waarschijnlijk namelijk door meer mensen gelezen dan de studies zelf. Maar ja, clichés zijn clichés omdat ze waar zijn. Daarom zullen er hieronder ongetwijfeld een aantal voorkomen, dat zij dan zo.

Als eerste wil ik mijn promotie begeleiders danken. Diana en Henk, mijn vader en moeder in wetenschapsland. Al tijdens onze eerste ontmoeting, het sollicitatiegesprek, dacht ik “dit voelt goed”. En dat gevoel bleek helemaal te kloppen. Bovendien vullen jullie elkaar als begeleiders goed aan, waardoor jullie voor mij de ideale begeleiding vormden. Diana, jij was altijd heel snel met goede feedback, had oog voor alle details en bewaakte de voortgang van het project. Henk, jij was meer van de grote lijn, kon de conclusies altijd helder verwoorden en knopen doorhakken. Diana, ik ben blij dat ik met je blijf samenwerken, met name ook binnen de taakgroep PE. Henk, op het moment van mijn verdediging ben je al met pensioen. Jammer dat je de vakgroep gaat verlaten, maar hoop dat je gaat genieten van alle dingen die je je hebt voorgenomen te gaan doen. Ik hoop ook dat ik je (bij O&O) nog af en toe zie om te horen hoe je nieuwe studie, je nieuwe vakantiehuis, etc. je bevallen.

Henk Schmidt, mijn tweede promotor, naast de dagelijkse begeleiding door Diana en Henk keek je op de achtergrond mee. Ik waardeer het zeer dat je, ondanks je drukke bestaan, af en toe tijd voor me vrijmaakte. Al zagen we elkaar niet zo vaak, je advies op hoofdlijnen was zeer waardevol. En als Diana, Henk en ik er met z'n drieën niet uitkwamen stond je klaar met een antwoord of advies.

Naast mijn promotie team zijn er nog andere mensen die me geholpen hebben bij mijn onderzoek die ik hierbij graag persoonlijk wil bedanken. Ron dank voor de analyses die je voor mijn eerste studie hebt uitgevoerd. Toine Smulders, dank voor je hulp bij mijn derde studie. Ik heb veel geleerd van het begeleiden van je thesis als onderdeel van het Honours programma. Herman, dank voor je medewerking bij de vierde studie. En niet alleen dat, door jou heb ik ook nog wat opgestoken van humane biologie. Jimmie, dank voor je hulp bij de vierde studie, maar ook voor alle statistiek vragen waarmee ik altijd bij je terecht kon. En tot slot was het leuk om je nog een paar maanden als kamergenootje te hebben gehad.

En dan mijn kamergenootjes. Als ik anderen vertelde dat ik met 5 collega's op een kamer zat vroegen mensen vaak met een moeilijk gezicht "Gaat dat wel goed?". Gelukkig kon ik daar altijd volmondig "Ja!" op antwoorden. We waren een echt team waarin de sfeer goed was. Zowel op werk- als privégebied hebben we veel met elkaar gedeeld. Niet altijd positief, maar dat maakte de periode wel memorabel en de band met elkaar hecht. Met veel plezier denk ik ook terug aan onze congresbezoeken in binnen- en buitenland. Dank dat jullie allemaal mijn kamergenootjes waren. In willekeurige volgorde wil ik hier wat dieper op ingaan.

Kamergenootjes van eerste uur Janneke en Jeantine. Jeantine, je was altijd bereid om me te helpen met allerlei klusjes m.b.t. mijn onderzoek, bijvoorbeeld de vele pakketjes vragenlijsten die samengesteld en uitgedeeld moesten worden. Voor alle medische vragen of wat voor vragen dan ook kon ik bij jou terecht, maar ook voor een gezellig praatje was je altijd beschikbaar. Janneke, jij was de eerste die ik vertelde dat ik zwanger was van Thomas, waar ik tijdens de EARLI in Exeter achter kwam. Een memorabel en speciaal moment. Ook denk ik met veel plezier terug aan de vele leuke momenten die we samen in o.a. Malaga, Exeter, Egmond aan zee, en op onze kamer hebben gehad. Ik ben dan ook blij dat je één van mijn paranimfen bent. Floris, jij was lange tijd de enige man in ons midden. Bedankt dat ik onderdeel mocht zijn van je harem. Ik weet nog dat ik onverwacht ziek werd en dat jij er voor gezorgd hebt dat mijn vragenlijsten gewoon uitgedeeld en ingenomen werden, waarvoor veel dank. Ik heb daarnaast ook nog mooie herinneringen aan de reis naar de AERA in Denver in 2011. Marjo en Greet, halverwege mijn promotietraject kwamen jullie van Heerlen naar Maastricht. Het duurde niet lang voordat het vertrouwd voelde. Ik hoop dat we het volhouden om af en toe voor koffie of lunch af te spreken. Ellen, onze benjamin. We plagen je wel eens omdat je de jongste bent, maar ik bewonder je wetenschappelijke ethos en ambitie. Ik ken weinig mensen die zo gedreven zijn in de wetenschap. Daarnaast kon en kan ik ook altijd met vragen bij je terecht en vaak kan ik hem, door hem simpelweg aan jou te stellen zelf al beantwoorden. Ik vind het dan ook heel fijn dat jij mijn andere paranimf bent. Katerina, ook al was je vaak bij SBE je was toch een gewaardeerde collega. Emmaline, onze samenwerking was iets korter van duur, maar daardoor niet minder leuk. Bedankt voor de leuke gesprekjes.

Jorrick, de tweede man in ons midden. De man van de creatieve input en de antwoorden op computer gerelateerde vragen. Als er ergens een logo of plaatje voor moet worden ontworpen ga je meteen aan de slag. Ik kan je ietwat cynische humor erg waarderen. Je schreef zelf weleens dat mensen een jaar nodig hebben om je te leren kennen. Het kostte inderdaad misschien wat tijd (wel minder dan een jaar), maar het was zeker de moeite waard. Lorette, het meest recente kamergenootje. Het was kort maar wel leuk. Ik hoop dat we onze samenwerking bij Botox B.V. op dezelfde voet kunnen doorzetten.

Renée, Joost, Mariëtte, Jonne, Ingrid, Annemarie, Frank, Dominique en alle andere collega's bij O&O, bedankt voor de vele gezellige lunches, etentjes, borrels en praatjes op de gang of kamer.

En dan, mijn bestaande collega's in een nieuw jasje bij de taakgroep PE. Ineke, dank voor het vertrouwen dat je in me had waardoor ik nu voor de taakgroep PE mag werken. Renée, dank voor het wegwijs maken in de wereld van PE. Jill, mijn nieuwe kamergenootje. Ik kijk erg uit naar onze verdere samenwerking en hoop nog veel van je te leren. Daniëlle, Diana, Paul en Mariska. Jullie zijn de stille krachten op de achtergrond, maar van onschatbare waarde. Ik heb de laatste tijd veel van jullie geleerd.

Mijn dank gaat ook uit naar Nicky en Lilian. Jullie stonden en staan altijd klaar om vragen te beantwoorden en zijn nooit te beroerd om iets te regelen.

Naast collega's waren er buiten het werk om ook een hoop mensen die indirect een bijdrage hebben geleverd. En dan voornamelijk doordat ik samen met hen veel plezier heb beleefd waardoor de werk/privé verdeling goed in balans bleef. Mijn vriendengroep wil ik dan ook in het bijzonder bedanken. Lionne, Sebastiaan, Kelly, Rob, Imke, Sander, Joost, Sanne en Pieter, dank voor alle vriendenweekendjes, feestjes, stapavonden, meidenavonden, kookclubs en andere dingen die we samen hebben gedaan. Ik heb veel met jullie gelachen en waardeer jullie vriendschap bijzonder. En dan mijn familie. Indirect hebben jullie ook allemaal een bijdrage geleverd. Mijn ouders, Mariet en Léon, wil ik bedanken voor alle zaterdagen bij Brasserie Ludiek. Omdat het altijd een gezellige onderbreking van de werkweek was en jullie al mijn verhalen over mijn onderzoek geduldig aanhoorden. Mariet, fijn dat je iedere week en

dag op Thomas past en voor het ontwerpen van de prachtige cover van mijn proefschrift. Léon, ik waardeer het heel erg dat je altijd bereid om te komen klussen waardoor ons een hoop tijd en frustratie bespaard blijft. En dan mij broers Christiaan en Daniël en mijn schoonzusje Inge. Dank voor alle leuke momenten die we gedeeld hebben. Ik denk o.a. met veel plezier terug aan ons tripje naar het Afrikamuseum en laatst nog de vakantie aan de Kronenburger See. Ik ben blij dat ik jullie (schoon)zus ben! En niet te vergeten mijn kleine nichtje Fay.

Ook wil ik mijn schoonouders, Ursula en Frank, bedanken. Ook jullie hebben mijn werk altijd vol interesse gevolgd en jullie passen iedere week allebei een dag op Thomas. Mede dankzij jullie heb ik altijd met een gerust hart naar mijn werk kunnen gaan. Dank ook voor alle leuke etentjes en vakanties die we samen hebben beleefd.

En dan de twee belangrijkste mensen in mij leven. Mijn man Marcel en zoontje Thomas. Marcel, ik heb veel steun aan je gehad tijdens mijn promotie. Jij was degene die mij, op de momenten dat ik twijfelde of het even niet meer zag zitten, er altijd aan herinnerde dat ik het hele project tot een succesvol einde kon brengen. Je was ook nooit te beroerd om een stuk voor me door te lezen of met allerlei administratieve zaken te helpen. Ik ben blij en trots dat ik jouw vrouw ben en jou aan mijn zijde heb. Thomas, jij hebt nu nog geen besef van dit alles maar je hebt mijn leven meer veranderd dan ik ooit had gedacht. Als ik naar jou kijk vergeet ik even alle dagelijkse beslommeringen en geniet ik van je vrolijkheid en ondeugendheid. Ik ben zo blij dat ik jou in mijn leven heb.

Tot slot zijn er vast mensen die ik vergeten ben, waarvoor mijn excuses. Dank aan iedereen waarmee ik de afgelopen vijf jaar heb samengewerkt en die, op welke manier dan ook, een bijdrage heeft geleverd.

Rachelle





# **CURRICULUM VITAE**



**Rachelle J. A. Kamp** was born on August 23<sup>rd</sup> 1984 in Sittard, the Netherlands. She attended Athenaeum at Were' Di comprehensive school in Valkenswaard where she graduated in 2002. From 2002 to 2007 she studied Health Psychology at Tilburg University. During this study she completed her internship at nursing home and rehabilitation centre 'Volckaert' in Dongen. In 2007 she started a minor and master's degree program in Educational Sciences at Utrecht University. She completed her internship at the Student Service Centre at the Eindhoven University of Technology. After she graduated in 2008, she started working as a PhD student at the department of Educational Development and Research at the Faculty of Health, Medicine, and Life Sciences at Maastricht University. In addition to her research activities, she was a tutor of first- and second-year Health Sciences students, a teacher of communication skills, and a teacher in the master's program of Health Professions Education. From 2009 to 2013 she was a member of the educational committee of the Interuniversity Centre of Educational Sciences (ICO). Currently, she works at the same department as an assistant professor, where she is a member of the taskforce Programme Evaluation.



# **SHE DISSERTATION SERIES**



In the SHE Dissertation Series dissertations are published of PhD candidates from the School of Health Professions Education (SHE) who defended their PhD theses at Maastricht University. The most recent ones are listed below. For more information go to: [www.maastrichtuniversity.nl/she](http://www.maastrichtuniversity.nl/she).

1. Ladhani, Z. (05-07-2013). Competency based education and professional competencies: a study of institutional structures, perspectives and practices in Pakistan.
2. Jippes, M. (01-02-2013). Culture matters in medical schools: How values shape a successful curriculum change.
3. Duvivier, R. J. (12-12-2012). Teaching and Learning Clinical Skills. Mastering the Art of Medicine.
4. De Feijter, J.M. (09-11-2012). Learning from error to improve patient safety.
5. Prescott, L. (09-11-2012). Ensuring the Competence of Dental Practitioners through the Development of a Workplace-Based System of Assessment.
6. Cilliers, F.J. (05-09-2012). The Pre-assessment Learning Effects of Consequential Assessment: Modelling how the Examination Game is Played.
7. Spanjers, I. A.E. (05-07-2012). Segmentation of Animations: Explaining the Effects on the Learning Process and Learning Outcomes.
8. Al-Kadri, H.M.F. (28-06-2012). Does Assessment Drive Students' Learning?
9. Leppink, J. (20-06-2012). Propositional manipulation for conceptual understanding of statistics.
10. Van Zundert, M.J. (04-05-2012). Conditions of Peer Assessment for Complex Learning.
11. Claramita, M. (30-03-2012). Doctor-patient communication in a culturally hierarchical context of Southeast Asia: A partnership approach.
12. Kleijnen, J.C.B.M. (21-03-2012). Internal quality management and organizational values in higher education.
13. Persoon, M.C. (19-01-2012). Learning in Urology; The influence of simulators and human factors.
14. Pawlikowska, T.R.B. (21-12-2011). Patient Enablement; A Living Dialogue.
15. Sok Ying Liaw, (14-12-2011). Rescuing a Patient in Deteriorating Situations (RAPIDS): A programmatic approach in developing and evaluating a simulation-based educational program.

16. Singaram, V.S. (7-12-2011). Exploring the Impact of Diversity Factors on Problem-Based Collaborative Learning.
17. Balslev, T. (24-11-2011). Learning to diagnose using patient video cases in paediatrics: Perceptive and cognitive processes.
18. Widyandana, D. (19-10-2011). Integrating Pre-clinical skills training in skills laboratory and primary health care centers to prepare medical students for their clerkships.
19. Durning, S.J. (09-09-2011). Exploring the Influence of Contextual Factors of the Clinical Encounter on Clinical Reasoning Success (Unraveling context specificity).
20. Govaerts, M.J.B. (08-09-2011). Climbing the Pyramid; Towards Understanding Performance Assessment.
21. Stalmeijer, R. E. (07-07-2011). Evaluating Clinical Teaching through Cognitive Apprenticeship.
22. Malling, B.V.G. (01-07-2011). Managing word-based postgraduate medical education in clinical departments.
23. Veldhuijzen, J.W. (17-06-2011). Challenging the patient-centred paradigm: designing feasible guidelines for doctor patient communication.
24. Van Blankenstein, F. (18-05-2011). Elaboration during problem-based, small group discussion: A new approach to study collaborative learning.
25. Van Mook, W. (13-05-2011). Teaching and assessment of professional behavior: Rhetoric and reality.
26. De Leng, B. (8-12-2009). Wired for learning. How computers can support interaction in small group learning in higher education.
27. Maiorova, T. (29-05-2009). The role of gender in medical specialty choice and general practice preferences.
28. Bokken, L. (04-03-2009). Innovative use of simulated patients for educational purposes.
29. Wagenaar, A. (18-09-2008). Learning in internships. What and how students learn from experience.
30. Driessen, E. (25-06-2008). Educating the self-critical doctor. Using portfolio to stimulate and assess medical students' reflection.
31. Derkx, H. (18-06-2008). For your ears only. Quality of telephone triage at out-of-hours centres in the Netherlands.

32. Niessen, Th. (30-11-2007). Emerging epistemologies: making sense of teaching practice.
33. Budé, L. (05-10-2007). On the improvement of students' conceptual understanding in statistics education.
34. Niemantsverdriet, S. (26-07-2007). Learning from international internships: A reconstruction in the medical domain.
35. Marambe, K. (20-06-2007). Patterns of student learning in medical education – A Sri Lankan study in traditional curriculum.
36. Pleijers, A. (19-01-2007). Tutorial group discussion in problem-based learning.
37. Sargeant, J. (21-09-2006). Multi-source feedback for physician learning and change.
38. Dornan, T. (12-06-2006). Experience-based learning.
39. Wass, V. (12-05-2006). The assessment of clinical competence in high stakes examinations.
40. Prince, K. (21-04-2006). Problem-based learning as a preparation for professional practice.



# **ICO DISSERTATION SERIES**



In the ICO Dissertation Series the dissertations of graduate students from faculties and institutes on educational research within the ICO Partner Universities are published: Eindhoven University of Technology, Leiden University, Maastricht University, Open University of the Netherlands, University of Amsterdam, University of Twente, Utrecht University, VU University Amsterdam, and Wageningen University, and formerly University of Groningen (until 2006), Radboud University Nijmegen (until 2004), and Tilburg University (until 2002). The University of Groningen, University of Antwerp, University of Ghent, and the Erasmus University Rotterdam have been 'ICO 'Network partner' in 2010 and 2011. From 2012 onwards, these ICO Network partners are full ICO partners, and from that period their dissertations will be added to this dissertation series.

1. Radstake, H. (14-05-2009). Teaching in diversity: Teachers and pupils about tense situations in ethnically heterogeneous classes. Amsterdam: University of Amsterdam.
2. Du Chatenier, E. (09-09-2009). Open innovation competence: Towards a competence profile for inter-organizational collaboration in innovation teams. Wageningen: Wageningen University.
3. Van Borkulo, S.P. (26-06-2009). The assessment of learning outcomes of computer modelling in secondary science education. Enschede: University of Twente.
4. Handelzalts, A. (17-09-2009). Collaborative curriculum development in teacher design teams. Enschede: University of Twente.
5. Nievelstein, F.E.R.M. (18-09-2009). Learning law: Expertise differences and the effect of instructional support. Heerlen: Open University of the Netherlands.
6. Visser-Wijnveen, G.J. (23-09-2009). The research-teaching nexus in the humanities: Variations among academics. Leiden: Leiden University.
7. Van der Rijst, R.M. (23-09-2009). The research-teaching nexus in the sciences: Scientific research dispositions and teaching practice. Leiden: Leiden University.
8. Mainhard, M.T. (25-09-2009). Time consistency in teacher-class relationships. Utrecht: Utrecht University.
9. Van Ewijk, R. (20-10-2009). Empirical essays on education and health. Amsterdam: University of Amsterdam.

10. Seezink, A. (18-11-2009). Continuing teacher development for competence-based teaching. Tilburg: Tilburg University.
11. Rohaan, E.J. (09-12-2009). Testing teacher knowledge for technology teaching in primary schools. Eindhoven: Eindhoven University of Technology.
12. Kirschner, F.C. (11-12-2009). United brains for complex learning. Heerlen: Open University of the Netherlands.
13. Wetzels, S.A.J. (18-12-2009). Individualized strategies for prior knowledge activation. Heerlen: Open University of the Netherlands.
14. Zitter, I.I. (04-02-2010). Designing for learning: Studying learning environments in higher professional education from a design perspective. Utrecht: Utrecht University.
15. Koopman, M. (11-02-2010). Students' goal orientations, information processing strategies and knowledge development in competence-based pre-vocational secondary education. Eindhoven: Eindhoven University of Technology.
16. Platteel, T. (11-02-2010). Knowledge development of secondary school LI teachers on concept-context rich education in an action-research setting. Leiden: Leiden University.
17. Mittendorff, K. M. (12-03-2010). Career conversations in senior secondary vocational education. Eindhoven: Eindhoven University of Technology.
18. Moolenaar, N.M. (01-06-2010). Ties with potential: Nature, antecedents, and consequences of social networks in school teams. Amsterdam: University of Amsterdam.
19. Duijnhouwer, H. (04-06-2010). Feedback effects on students' writing motivation, process, and performance. Utrecht: Utrecht University.
20. Kessels, C.C. (30-06-2010). The influence of induction programs on beginning teachers' well-being and professional development. Leiden: Leiden University.
21. Endedijk, M.D. (02-07-2010). Student teachers' self-regulated learning. Utrecht: Utrecht University.
22. De Bakker, G.M. (08-09-2010). Allocated online reciprocal peer support as a candidate for decreasing the tutoring load of teachers. Eindhoven: Eindhoven University of Technology.
23. Groenier, M. (10-09-2010). The decisive moment: Making diagnostic decisions and designing treatments. Enschede: University of Twente.

24. Bonestroo, W.J. (24-09-2010). Planning with graphical overview: Effects of support tools on self-regulated learning. Enschede: University of Twente.
25. Vos, M.A.J. (30-09-2010). Interaction between teachers and teaching materials: On the implementation of context-based chemistry education. Eindhoven: Eindhoven University of Technology.
26. Kostons, D.D.N.M. (05-11-2010). On the role of self-assessment and task-selection skills in self-regulated learning. Heerlen: Open University of the Netherlands.
27. Bruin-Muurling, G. (21-12-2010). The development of proficiency in the fraction domain: Affordances and constraints in the curriculum. Eindhoven: Eindhoven University of Technology.
28. Slof, B. (28-01-2011). Representational scripting for carrying out complex learning tasks. Utrecht: Utrecht University.
29. Fastré, G. (11-03-2011). Improving sustainable assessment skills in vocational education. Heerlen: Open University of the Netherlands.
30. Min-Leliveld, M.J. (18-05-2011). Supporting medical teachers' learning: Characteristics of effective instructional development. Leiden: Leiden University.
31. Van Blankenstein, F.M. (18-05-2011). Elaboration during problem-based small group discussion: A new approach to study collaborative learning. Maastricht: Maastricht University.
32. Dobber, M. (21-06-2011). Collaboration in groups during teacher education. Leiden: Leiden University.
33. Jossberger, H. (24-06-2011). Towards self-regulated learning in vocational education: Difficulties and opportunities. Heerlen: Open University of the Netherlands.
34. Schaap, H. (24-06-2011). Students' personal professional theories in vocational education: Developing a knowledge base. Utrecht: Utrecht University.
35. Kolovou, A. (04-07-2011). Mathematical problem solving in primary school. Utrecht: Utrecht University.
36. Beusaert, A.J. (19-10-2011). The use of personal developments plans in the workplace. Effects, purposes and supporting conditions. Maastricht: Maastricht University

37. Favier, T.T. (31-10-2011). Geographic information systems in inquiry-based secondary geography education: Theory and practice. Amsterdam: VU University Amsterdam.
38. Brouwer, P. (15-11-2011). Collaboration in teacher teams. Utrecht: Utrecht University.
39. Molenaar, I. (24-11-2011). It's all about metacognitive activities; Computerized scaffolding of self-regulated learning. Amsterdam: University of Amsterdam.
40. Cornelissen, L.J.F. (29-11-2011). Knowledge processes in school-university research networks. Eindhoven: Eindhoven University of Technology.
41. Elffers, L. (14-12-2011). The transition to post-secondary vocational education: Students' entrance, experiences, and attainment. Amsterdam: University of Amsterdam.
42. Van Stiphout, I.M. (14-12-2011). The development of algebraic proficiency. Eindhoven: Eindhoven University of Technology.
43. Gervedink Nijhuis, C.J. (03-2-2012). Culturally Sensitive Curriculum Development in International Cooperation Enschede: University of Twente.
44. Thoonen, E.E.J. (14-02-2012). Improving Classroom Practices: The impact of Leadership School Organizational Conditions, and Teacher Factors Amsterdam: University of Amsterdam.
45. Truijen, K.J.P (21-03-2012). Teaming Teachers. Exploring factors that influence effective team functioning in a vocational education context Enschede: University of Twente.
46. Maulana, R.M. (26-03-2012). Teacher-student relationships during the first year of secondary education. Exploring of change and link with motivation outcomes in The Netherlands and Indonesia. Groningen: University of Groningen.
47. Lomos, C. (29-03-2012). Professional community and student achievement. Groningen: University of Groningen.
48. Mulder, Y.G. (19-04-2012). Learning science by creating models. Enschede: University of Twente.
49. Van Zundert, M.J. (04-05-2012). Optimising the effectiveness and reliability of reciprocal peer assessment in secondary education Maastricht: Maastricht University.
50. Ketelaar, E. (24-05-2012). Teachers and innovations: on the role of ownership, sense-making, and agency. Eindhoven: Eindhoven University of Technology.

51. Logtenberg, A. (30-5-2012). Questioning the past. Student questioning and historical reasoning Amsterdam: University of Amsterdam.
52. Jacobse, A.E. (11-06-2012). Can we improve children's thinking? Groningen: University of Groningen.
53. Leppink, J. (20-06-2012). Propositional manipulation for conceptual understanding of statistics Maastricht: Maastricht University.
54. Van Andel, J. (22-06-2012). Demand-driven Education. An Educational-sociological Investigation. Amsterdam: VU University Amsterdam.
55. Spanjers, I.A.E. (05-07-2012). Segmentation of Animations: Explaining the Effects on the Learning Process and Learning Outcomes. Maastricht: Maastricht University.
56. Vrijnsen-de Corte, M.C.W. Researching the Teacher-Researcher. Practice-based research in Dutch Professional Development Schools Eindhoven: Eindhoven University of Technology.
57. Van de Pol, J.E. (28-09-2012). Scaffolding in teacher-student interaction. Exploring, measuring promoting and evaluating scaffolding Amsterdam: University of Amsterdam.
58. Phielix, C. (28-09-2012). Enhancing Collaboration through Assessment & Reflection [Samenwerking Verbeteren door middel van Beoordeling en Reflectie]. Utrecht: Utrecht University.
59. Peltenburg, M.C. (24-10-2012). Mathematical potential of special education students Utrecht: Utrecht University.
60. Doppenberg, J.J. (24-10-2012). Collaborative teacher learning: settings, foci and powerful moments Eindhoven: Eindhoven University of Technology.
61. Kenbeek, W.K. (31-10-2012). Back to the drawing board. Creating drawing or text summaries in support of System Dynamics modeling Enschede: University of Twente.
62. De Feijter, J.M. (09-11-2012). Learning from error to improve patient safety Maastricht: Maastricht University.
63. Timmermans, A.C. (27-11-2012). Value added in educational accountability: Possible, fair and useful? Groningen: University of Groningen.
64. Van der Linden, P.W.J. (20-12-2012). A design-based approach to introducing student teachers in conducting and using research. Eindhoven: Eindhoven University of Technology.





