Teaching *for* versus *through* Problem Solving: Impact on Teaching and Learning.

Annelies Raes, Celine Vens, Pieter Vanneste, Fien Depaepe
ITEC, KULeuven, Belgium also at imec Kapeldreef 75, B-3001 Leuven, Belgium
Annelies.raes@kuleuven.be, Celine.vens@kuleuven.be, Pieter.vanneste@kuleuven.be, Fien.depaepe@kuleuven.be

Abstract: This paper presents the results of a quasi-experimental study with third bachelor students in Medicine comparing the teaching *for* problem solving and teaching *through* problem solving condition. This study focuses both on the outcome and process effects. On the cognitive level this includes students' conceptual learning gain. On the affective level this includes students' experienced motivation both in the instruction phase and the problem-solving phase. This study complements previous research by taking a zoomed-in perspective including multimodal data streams (i.e. audiovisual, self-report, logdata and psycho-physiological data) to fully understand the effects on the teaching and learning mechanisms. Although we did not find statistically significant effects regarding the conceptual understanding, the results of the affective outcome and process variables are consistent with the growing body of evidence that generating solutions to novel problems prior to instruction can improve deeper learning and can make teaching more interactive and student-centered.

Theoretical and empirical framework

Problem solving is often suggested as the new way of learning to cognitively activate students and prepare them for the 21st Century skills (Felmer, Pehkonen, Kilpatrick, 2016; Lester & Cai, 2016). During the past 30 years there have been significant advances in our understanding of the cognitive, metacognitive, and affective aspects of problem solving in domains as science and mathematics. It is known that problem solving has many learning advantages, but it is also known that problem solving places higher cognitive demands and is often difficult for novices in a certain field with inconsistent and limited prior knowledge as indicated within the instructionist-constructivist debate (Kirschner, Sweller, & Clark, 2006). To reduce this burden on the cognitive resources and mitigate problems associated with learner disengagement and learner frustration, many teachers still opt for direct instruction or knowledge transfer preventing struggle and saving time, and to make sure the correct information is taught (Sweller & Chandler, 1991). Although there are several reasons to believe in the effectiveness of problem solving based on considerable research on teaching problem solving in classrooms, there remain far more questions than there are consistent answers about this complex form of activity. One of these questions is if problem solving should be taught as a separate topic in the curriculum or if it should be integrated throughout the curriculum? Schroeder and Lester (1989) identified these two opposite types of approaches to problem solving as 1) 'teaching *for* problem solving' (further abbreviated as TforPS) starting with the content and concepts and then moving to solving problems and 2) 'teaching *through* problem solving' (abbreviated as TthroughPS) meaning students learn the content and concepts through real contexts, problems, situations and problems followed by an instruction phase. Although not equivalent, the latter approach can in a way be related to the theoretical framework of Kapur (2008) defining *productive failure* as a teaching and learning approach which engages students in solving problems requiring concepts they have yet to learn, followed by a consolidation and instruction phase on the targeted concepts. This approach is based on the work of Schmidt and Bjork (1992) suggesting that experimental manipulations that may hinder performance in the shorter term can actually be productive for learning in the longer term. They even concluded that maximizing performance in the initial learning may not necessarily be the ones that maximize learning in the longer term. Previous research has found that starting with problem solving prompts the activation of prior knowledge and idea generation from different perspectives (Schwartz, Chase, Oppezzo, & Chin, 2011). These studies moreover indicated that approaches as productive failure are specifically beneficial for retention and transfer (see e.g. Kapur, 2013). In recent work, Kapur (2016) deepens this approach and suggests to distinguish four possibilities for designing learning: i.e. productive or unproductive success, and productive or unproductive failure. Within our conceptualization, we assume that the TforPS has a higher chance to lead to productive success, whereas TthroughPS has a higher chance to lead to productive failure.

This study compared the TforPS and TthroughPS condition in a quasi-experimental study with third bachelor students in medicine. This study complements to previous work as it aims to better understand the effects of the opposite approaches on both outcome and process variables including multiple streams of data. Not only the student's knowledge gain and behavior is being studied, but also their psycho-physiological state is being monitored. During the experiment, wearables attached to the teacher's and some of the students' wrists, measure their skin conductance level, skin conductance phasic responses and skin temperature. In doing so, the relative stress level a student experiences during a certain moment in the experiment can be determined.

Research questions & hypotheses

The first research question focuses on the effect of both approaches on the outcome variables on the cognitive and on the affective level. The outcome variables include students' conceptual knowledge gain and their self-efficacy on the cognitive level and students' experienced mental effort, experienced competence satisfaction, and intrinsic motivation and perceived value of the learning activities (instruction phase and problem-solving phase) on the affective level. Thanks to the wearable data, also the stress level that students and the teacher experience is studied. Based on previous research, the hypothesis of this study is that in the TthroughPS condition, on the one hand, the cognitive demands will be higher both for students and for the teacher, but on the other hand, students' agency and therefore also their interest, during the task and during the following instruction will be higher. Regarding the effect on the learning gain, we did not expect significant differences on the posttest immediately after the intervention, but hypothesize based on previous literature that the benefits of TthroughPS will unfold on the retention test on the longer term (Schwartz et al., 2011). As the effects are highly dependent on the way the intervention is implemented and the way it includes the expected cognitive mechanisms that facilitate learning from productive failure (Loibl, Roll, & Rummel, 2016), evaluating these mechanism was part of our second research question.

The second research question focuses on the effect of both approaches on the process variables and aims to more qualitatively analyze the effects on the learning mechanisms during the problem-solving phase and the teaching mechanisms during the instruction phase. Within the scope of this study and based on the theoretical assumption that TthroughPS is more time consuming and more demanding, we compared the duration and effort during the problem-solving phase in both conditions and hypothesized that starting with the problem solving phase is more engaging, but also more demanding, and consequently more time consuming. Based on previous literature indicating that instruction after problem solving provides more opportunities for interactions between students and teacher, both the amount and duration of the interactions in the instruction phases of both conditions were analyzed and compared.

Methodology: intervention and measurements

This study was conducted during students' regular class periods of the course 'Biostatistics' taught at the University of Leuven, campus Kulak Kortrijk in Belgium. The specific study was scheduled September 29, 2017. During previous academic years, the content about sampling methods and bias within sampling was given during a lecture based on a power point presentation in which the content is taught using a case study of the National Health Study. Within the scope of the current study, this specific lecture has been redesigned and a transfer task has been added to the instruction phase in which the students were asked to set up a survey to evaluate the psychosocial wellbeing of students within their campus. The focus of the task was not to formulate the questions of the survey, but students had to focus on how to select the method of data collection, how to select a good sample, how to deal with randomness, how to deal with non-response and how to solve the problem of over-and under-sampling through weights. This year, this curriculum was given to 48 third year bachelor students in Medicine. For the purpose of the study, as depicted in Figure 1, students were randomly assigned to one of the two conditions: TforPS and TthroughPS.

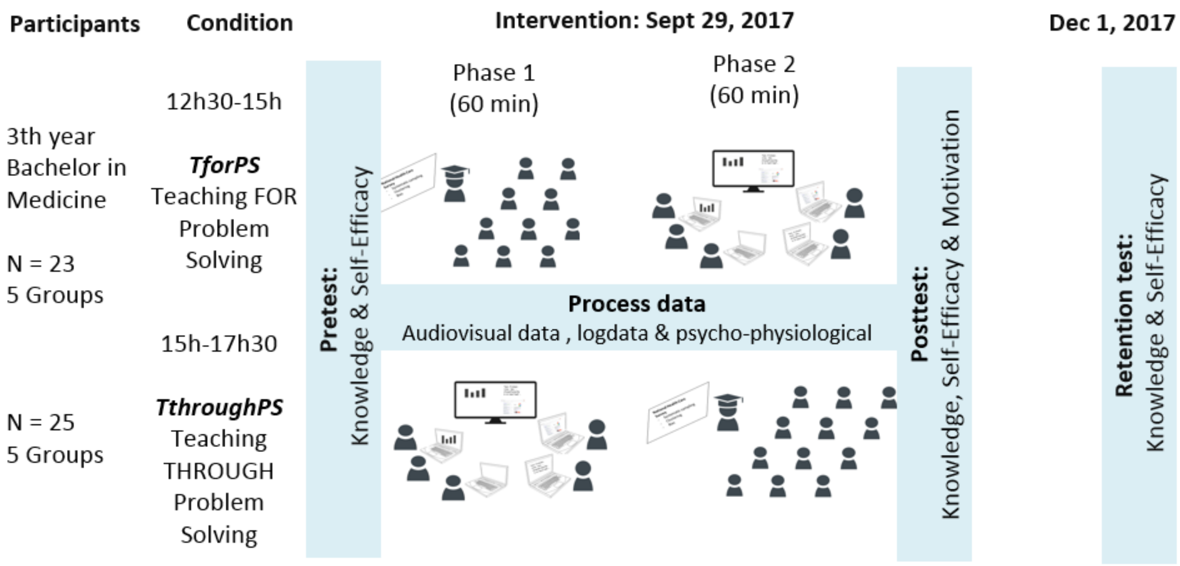

Figure 1. Procedure and design of the quasi-experimental study including both outcome and process variables.

In the TforPS condition, students first got direct instruction in which the instructor modeled and contrasted the different sampling methods with appropriate instructional facilitation and explanation to make them attend to critical features and possible bias within the sampling methods. The direct instruction phase was followed by the collaborative problem-solving phase. In the TthroughPS condition the instruction phase followed the problem-solving phase and aimed to explain the exact same content, but afforded opportunities to compare and contrast relevant student-generated solutions. The learning setting of this study is illustrated in Figure 2.

In this study the main dependent outcome variables were situated on the cognitive and affective level. To assess students' knowledge gain, students administered three equivalent test forms, one as pretest, one as posttest, and one as retention test two months after the intervention. In line with previous research (e.g. Kapur, 2013), the pretest included 4 multiple-choice questions and one open-response question; the post- and retention test included 6 multiple-choice questions and the same open-response question. Based on Bandura's work (1997), students' perceived self-efficacy was also measured by asking students to rate on a scale from 0 – 100 how confident they felt about the correctness of their given answer. To measure students' mental effort and motivation during both the instruction and problem-solving phase the Intrinsic Motivation Inventory (IMI) (Deci & Ryan, 2000) was administered immediately after the given session.

To assess and compare the learning and teaching mechanisms, in both conditions, different data streams have been triangulated. First, audiovisual data have been recorded during the complete intervention from three camera positions to capture student-student and student-teacher interactions for observing and capturing the teacher's and students' actions and reactions throughout the instruction phase. One camera moreover captured the student-student interactions of one group in both conditions. Second, students' reasoning and answers during the problem solving phase were logged on the technology platform WISE on which the problem solving task was provided to students and students' answers were evaluated by means of the knowledge integration rubric (Slotta & Linn, 2009). Third, psycho-physiological data were collected by using individual sensors measuring skin conductance, skin temperature and acceleration.

This overall study design has been evaluated and approved by the Social and Societal Ethics Committee and all students gave their informed consent.

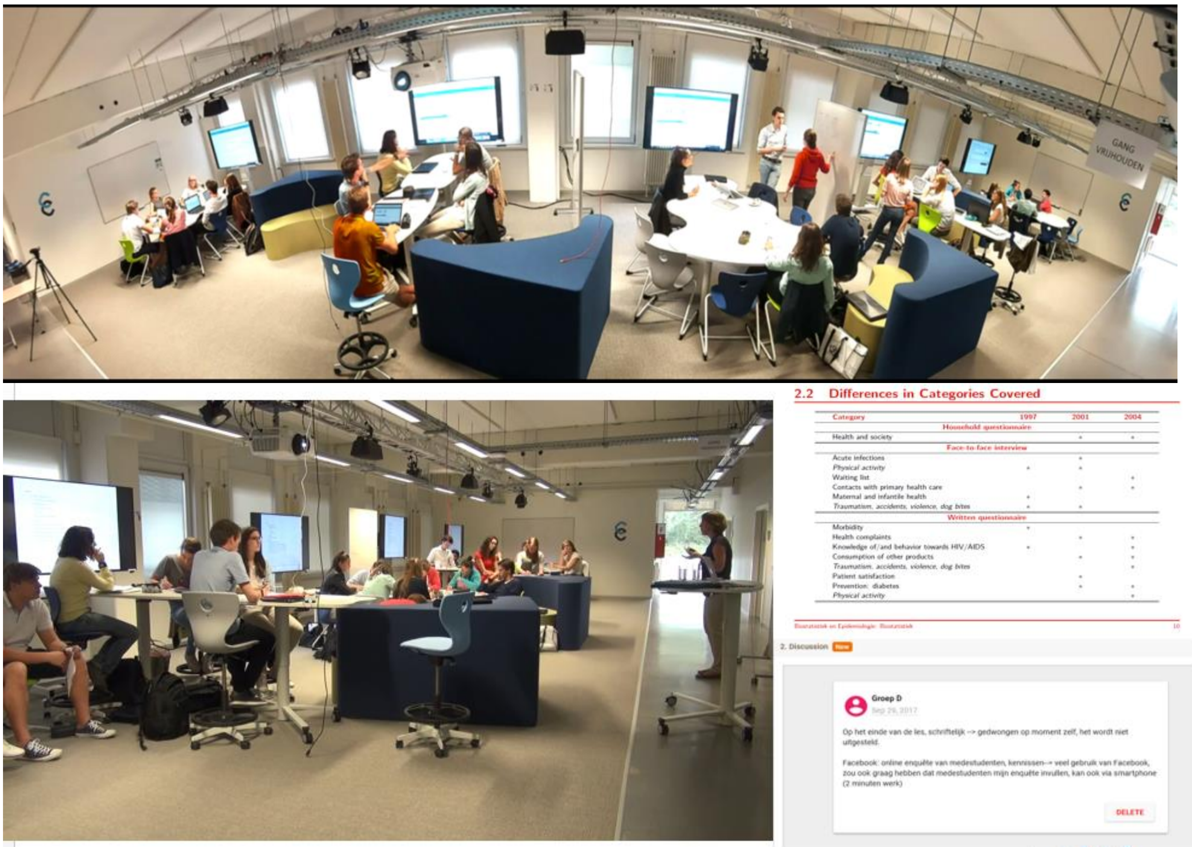

Figure 2: Learning context of the study: upper picture depicts the problem-solving phase, the lower picture displays the instruction phase within the TthroughPS condition in which both slides and students input were shown on the teacher's screen as displayed at the lower right side.

Results

RQ 1. Effects on students' progress in conceptual understanding, self-efficacy and motivation

Based on paired t-tests, an overall increase between pretest and posttest was found with respect to students' conceptual knowledge ($t(46)= 2{,}14$ $p < .001$) and regarding students' self-efficacy ($t(27)= 11{,}56$ $p < .001$). Moreover, one-way analyses of covariance (ANCOVA's) were conducted with posttest scores as dependent variable, condition as independent factor, and pretest scores as covariate to discover whether there are differences between both conditions on the posttest measure regarding conceptual knowledge and self-efficacy, after adjustment for the pretest scores. ANCOVA indicated no significant difference between the conditions as depicted in Figure 3. Moreover, no effect was found on the retention test that was administered two months later.

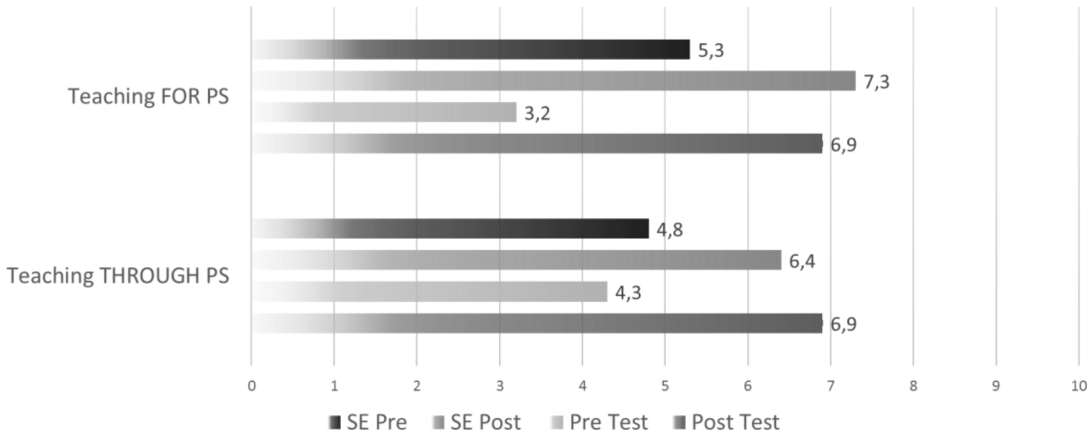

Figure 3: Students' progress in conceptual understanding and self-efficacy

To test students' motivation during both the instruction and problem-solving phase the IMI was administered immediately after the given session. Figure 4 displays the results regarding the different subscales of the IMI, including the experienced enjoyment during respectively the instruction and the problem-solving phase, the perceived competence, the effort they put into the activity, the perceived value or importance of the instruction and problem solving, the perceived autonomy or choice and finally the relatedness students experienced. First, we found significant within-subject differences in motivation between the evaluation of the instruction phase and the evaluation of the problem-solving phase. The mean interest and relatedness for example was significantly higher during the group work (respectively ($t(46)= 0.29$, $p = 0.045$ and $t(46)= 0.55$, $p < 0.001$). Yet, the perceived value and competence was significantly higher regarding the instruction phase. Although these are interesting results, the main focus of this study was the ordering effect the problem-solving task first versus second. Based on between-subject analyses we found two trends. Regarding the instruction phase on the one hand, we found that students' perceived choice or experienced autonomy was higher during the instruction given after the group work. On the other hand, we found a lower perceived competence during the problem solving phase given before the instruction phase. It is important to interpret these results with caution as they are based on a small scale study. Nevertheless, these trends are in line with previous findings in the literature.

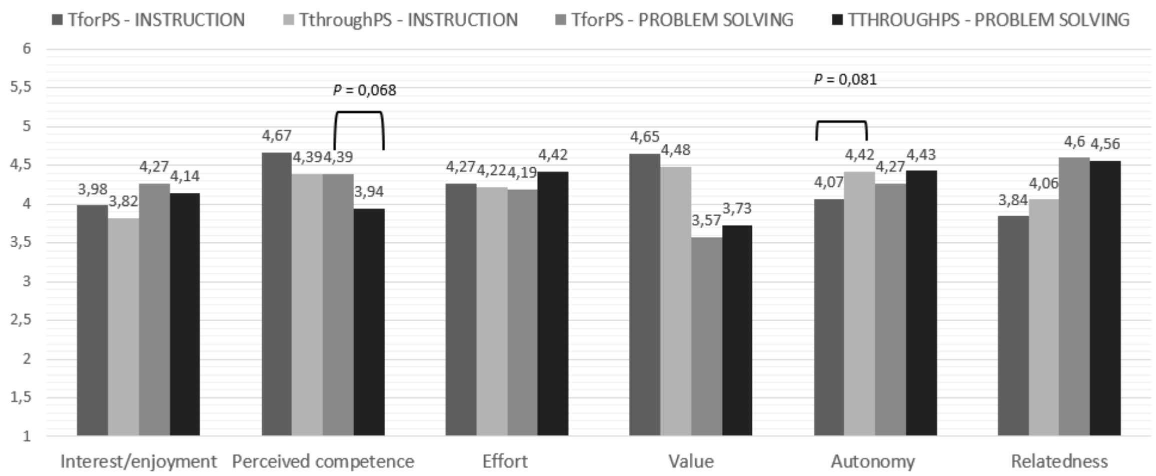

Figure 4: Students' perceived motivation based on the post-experimental intrinsic motivation inventory (IMI)

RQ 2. Effects on the learning and teaching mechanisms

To assess and compare the learning and teaching mechanisms in both conditions, first audiovisual data have been analyzed. Table 1 summarizes the results of the analysis and indicates that in both conditions the time spent to the instruction phase was comparable (i.e. circa 1 hour). However, the time students spent on the problem-solving task differed significantly between both conditions. In the TforPS condition in which the instruction was followed by the group work, one group already finished the task after 25 minutes. The maximum time spent on the group work was 41 minutes. The time spent on the problem-solving task in the TthroughPS condition was higher (mean time spent was 46minutes versus 36 minutes in the TforPS condition) and the variation was lower. The minimum time spent on the task was 44 minutes.

When having a closer look at the quality of answers during the problem-solving task, we found a substantial difference between both conditions. To indicate this difference, Table 1 displays the differences in lengths of the answer on the most challenging subtask of the group work. This subtask asked to postulate their sampling method to select 120 students for the campus survey keeping in mind the prerequisite of selecting the same number of students from all students years and guaranteeing the presence of students from large as well as small faculties in the sample.

Table 1: Comparing the duration and interactions in both conditions

	Teaching FOR PS **Instruction followed by group work: N = 23; 5 groups**	**Teaching THROUGH PS** **Group work followed by instruction. N = 25; 5 groups**
Duration of instruction phase	59 minutes	1h4minutes
Duration of problem solving phase of five groups	Min. time spent:25 minutes Max. time spent:41 minutes Mean time spent:36.6 minutes	Min. time spent: 44 minutes Max. time spent: 47 minutes Mean time spent: 46 minutes
Student-teacher interactions	3 interactions 2 initiated by the teacher 1 initiated by a student asking for a clarification	6 interactions 5 initiates by the teacher, 1 initiated by student asking a question
Student-teacher interaction time	2min10sec	10min20sec
Number of students interacting with teacher	2 out of 23 students	6 out of 25 students
Quality of group work: **indicator length of reasoning regarding task 3.1 'Sampling')**	Min. 9 words Max. 44 words Mean: 24	Min: 43 words Max: 171 words Mean: 78 words

First, it was revealed that the mean length of answers of student groups within the TforPS condition was significantly lower than the lengths of answers of student groups within the TthroughPS condition (*M*=24 versus *M*=78). Within the TforPS condition, although the answers were shorter, the answers were correct in four out of the five groups, but were less reflective. Three groups used in their answer the correct conceptualization of systematic sampling and stratification. It was also interesting to see that three groups made the sampling method visible in the drawing tool in the same way the sampling method was demonstrated by the teacher during the instruction phase. One group described and drew the sampling method in the correct manner, but did not specify the conceptualization of the method. We found that the one group in the TforPS task that only spent 25 minutes on the task, did not provide a correct and satisfying answer and also exhibited off-task behavior (i.e. drawing a funny smiley in the drawing tool which was meant to visualize their sampling method, see Figure 5). Within the other groups within both conditions no off-task behavior was indicated. The answers of the student groups within the TthroughPS condition better described the procedure of the sampling method they proposed, yet only one student group described their method correctly as stratified sampling. No group used the drawing tool to make visible their method, yet, as depicted in Figure 2, two groups had actively used the provided white boards to support their reasoning. One dome camera captured the student-student interactions of one group in both conditions, the results of these analyses will be presented at the conference.

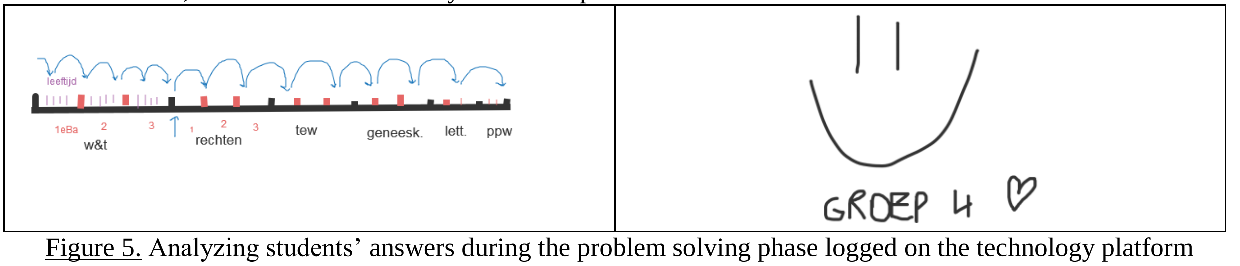

Figure 5. Analyzing students' answers during the problem solving phase logged on the technology platform

Based on previous literature indicating that instruction after problem solving provides more opportunities for interactions between students and teacher, both the number and duration of the interactions in both instruction phases were analyzed and compared. Again, we found remarkable differences between both instruction phases. In the instruction phase within the TforPS only 3 teacher-students interactions were observed which only lasted 2 minutes, two were initiated by a question of the lecturer, one was initiated by a clarifying question of a student. The interactions included 2 different students. During the course, the lecturer asked 4 more questions (e.g. is everything clear?), but students did not respond. This was not the case in the instruction phase in the TthroughPS condition. In this condition, 6 student-teacher interactions were identified and the interaction time lasted 10 minutes. The interactions included 6 different students and the discussions dealt about the different strategies they applied during their group work.

Regarding the psycho-physiological data of the students, no significant differences in skin conductance level between the two conditions on the one hand, neither between the problem-solving and instruction phase on the other hand, were found. It is however still possible that different students experience a certain part of the course differently than others. Therefore, the covariance between the measured stress and the level of stress indicated in the self-report will further be explored. The teacher's psycho-physiological data depicted in Figure 6 shows that the teacher experienced more stress or arousal during instruction phase within the TthroughPS condition. This seems logical as it was the first time the teacher gave her course in this way and the intervention expected her to build upon the student-generate solutions by comparing and contrasting them with the correct solution, which takes the instructor out of the comfort zone. During the problem-solving phase, the teacher did not interfere, which is reflected in the data as well. After the intervention, the teacher remarked that it was not easy to refer to students-answers in the instruction phase and she reflected about the need of a co-teaching context.

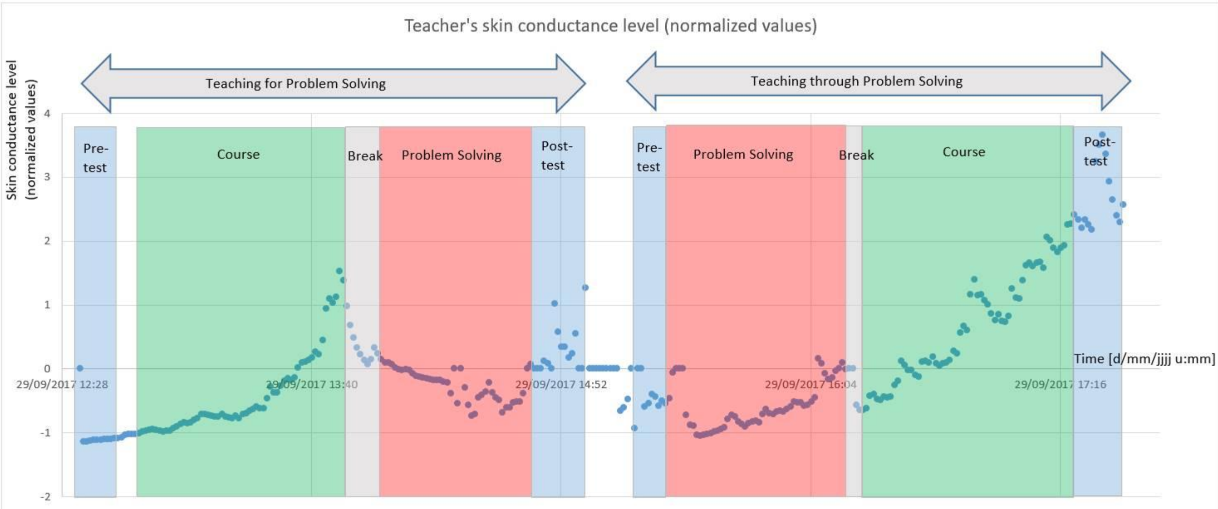

Figure 6: Teacher's skin conductance level during the experiment.

Conclusion and discussion

This study is based on considerable research on teaching problem solving in classrooms and helps to better understand and solve the remaining questions about this complex form of activity. The field stresses that there is much more than the two extremes in teaching and learning, i.e. pure direction instruction and unguided discovery learning, however more research is needed to provide evidence-based design guidelines regarding these problem solving design within the middle position (cfr. productive failure and productive success) to achieve optimal learning for all students.

More specific, this quasi-experimental study focused on the order of the collaborative problem solving task. While the first research question focused on the effect on the outcome variables, the second research questions focused on the process variables including the teaching mechanisms during the instruction phase and the learning mechanisms during the problem-solving phase. Although we did not find significant effects regarding the conceptual understanding both on the short and longer term, the results of the affective outcome variables and the qualitative analysis are consistent with the growing body of evidence that generating solutions to novel concepts and problems prior to instruction can improve learning and teaching compared to students who received instruction first (DeCaro & Rittle-Johnson, 2012; Kapur 2013). The analysis of the quality of answers revealed that starting with problem solving prompts activation of prior knowledge and idea generation from different perspectives.

Although this is a rather small intervention study on a short term, this study complements to the field of the Learning Sciences by bringing together multiple streams of data from both a qualitative and quantitative perspective to fully understand the learning and teaching mechanisms during different approaches towards problem solving and instruction. Taking this zoomed-in perspective gave us the opportunity to capture the interaction processes between the several actors (students, groups and the teacher) and sources (technology, peers and the teacher). The additional analysis of the group work will moreover give us better insight in the collaborative process and the way different students seek for help and/or use the provided support which is offered in the task environment. We only started to use psychophysiological data within this educational context and the results within this study did not results in significant differences in the student data, however, we found a considerable trend in the teacher's data; we will further explore these data as we believe that these data can tell us more detailed information about the affective states of students during teaching and learning in problem-solving contexts. A replication study is currently deployed in secondary education (grade 11 and 12) including the same measurements on a larger scale, including more students and spread over a longer time (4 times 50 minutes instead of 2 times 50 minutes). The first part of the study will be comparable with this design including structured problem solving by means of the web-based inquiry science environment. In the second half of the intervention, however, the problem solving task will be more open and less guided compared to the first one. Consequently this experimental condition will compare problem solving with a higher chance to lead to success followed or advanced by instruction and problem solving with a higher chance to lead to failure followed or advanced by instruction. According to Kapur (2016) such experimental comparison providing a breakdown of results comparing students who achieved problem-solving success with those who did not is lacking in the field. This extended replication study will provide fruitful insight into the complexity of problem solving, which is situated between the two extremes, i.e. pure direction instruction and unguided discovery learning.

References

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control.* New York: W.H. Freeman.

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. Journal of Experimental Child Psychology, 113, 552–568.

Deci, E. L. & Ryan, R. M. (2000). *The Intrinsic Motivation Inventory (IMI).* Retrieved from http://selfdeterminationtheory.org/intrinsic-motivation-inventory/

Felmer, P., Pehkonen, E., & Kilpatrick, K.. (2016). *Posing and Solving Mathematical Problems : Advances and New Perspectives.* Springer International Publishing

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*, 379–424

Kapur, M. (2013). Comparing learning from productive failure and vicarious failure. *The Journal of the Learning Sciences, 23*, 651–677.

Kapur, M. (2016). Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning. *Educational Psychologist, 51*(2), 289–99.

Kirschner, P. A., Sweller, J., & Clark, R.E. (2006). Why Minimal Guidance during Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist, 41*(2), 75–86.

Loibl, K., Roll, I., & Rummel, N. (2016). Towards a Theory of When and How Problem Solving Followed by Instruction Supports Learning. *Educational Psychology Review* 1–23.

Lester, F. K. & Cai, J. (2016). Can Mathematical Problem Solving Be Taught? Preliminary Answers from 30 Years of Research in *Posing and Solving Mathematical Problems*. Cham: Springer International Publishing.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217.

Schroeder and Lester (1989). Developing understanding in Mathematics via problem solving. In P.R. Trafton (Ed.), *New directions for elementary school mathematics* (pp. 31-42). Reston, VA: NCTM.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*, 759–775.

Slotta, J. D. &. Linn, M. C. (2009). *WISE Science, Web-Based Inquiry in the Classroom*. New York: Teachers College Press.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8,* 351–362.