

Learning by Preparing to Teach: Fostering Self-Regulatory Processes and Achievement During Complex Mathematics Problem Solving

Krista R. Muis, Cynthia Psaradellis, Marianne Chevrier, Ivana Di Leo, and Susanne P. Lajoie
McGill University

We developed an intervention based on the learning by teaching paradigm to foster self-regulatory processes and better learning outcomes during complex mathematics problem solving in a technology-rich learning environment. Seventy-eight elementary students were randomly assigned to 1 of 2 conditions: learning by preparing to teach, or learning for learning (control condition). Students' conceptualizations (task definitions) of the problem, self-regulatory processes, and mathematics achievement were then compared across the 2 conditions. To measure task definitions of the mathematics problem, students developed concept maps of the problem using a tablet application. To capture self-regulatory processes, students were asked to think out loud as they solved the problem. Results revealed that students in the learning by preparing to teach intervention developed a more detailed and better-organized concept map of the problem compared with students in the control condition. Students in the learning by preparing to teach intervention also engaged in more metacognitive processing strategies and had higher levels of mathematics problem solving achievement compared with students in the control condition. No differences were found, however, in planning and goal setting or in use of cognitive strategies across the 2 conditions. Implications of this research suggest students' initial task definitions may be a key factor in differences found when learning by teaching compared with solely learning for learning.

Keywords: concept maps, self-regulatory processes, mathematics achievement, learning by teaching

When it comes to complex mathematics problem solving, students struggle—especially at the primary grades (Durnin, Perrone, & Mackay, 1997). One reason students struggle with real-world complex problems is that they have difficulty developing appropriate schemas or problem representations that allow them to identify the necessary solution methods to successfully solve the problems (Chi, Feltovich, & Glaser, 1981; Fuchs et al., 2006; Quilici & Mayer, 1996). Another issue that students face is that they lack the self-regulatory skills necessary to navigate complex problems (Zimmerman & Martinez-Pons, 1990). As previous research has shown, key to successful mathematics problem solving is the self-regulation of one's learning (de Corte, Verschaffel, & Op't Eynde, 2000; Muis, 2004; Schoenfeld, 1994; Zimmerman & Labuhn, 2012). According to Schunk and Ertmer (2000), self-regulated learning is defined as self-generated thoughts, feelings, and actions that are oriented toward learning goals. Central to

self-regulated learning is metacognition (Efklides, 2008; Muis, 2007; Winne & Hadwin, 2008). Students who engage in more metacognitive processes typically achieve better learning outcomes (van der Stel & Veenman, 2010; Zimmerman, 2002).

It is unfortunate that younger learners, especially elementary students, are not very good at self-regulating or monitoring their learning (Butler & Winne, 1995; Zimmerman & Martinez-Pons, 1990). As such, it is imperative to develop learning environments that foster a better understanding of a problem's representation and increase key self-regulatory processes to improve learning outcomes. To address this, in collaboration with two teachers¹ who recently implemented a one-to-one tablet program² at their schools, we developed an intervention designed to foster a better problem representation, greater self-regulatory strategy use and better learning outcomes in the context of complex mathematics problem solving. We framed our intervention within the learning by teaching (e.g., Biswas, Jeong, Kinnebrew, Sulcer, & Roscoe, 2010; Biswas, Leelawong, Schwartz, Vye, & The Teachable Agents Group at Vanderbilt, 2005; Palinscar & Brown, 1984; Roscoe, 2014), and self-regulated learning (Greene & Azevedo, 2009; Muis, 2007; Winne & Hadwin, 2008) literatures. Specifi-

This article was published Online First August 24, 2015.

Krista R. Muis, Cynthia Psaradellis, Marianne Chevrier, Ivana Di Leo, Susanne P. Lajoie, Department of Educational and Counselling Psychology, McGill University.

Support for this research was provided by a grant from the Fonds de la Recherche du Québec sur la Société et la Culture and by a grant from the Social Sciences and Humanities Research Council of Canada.

We thank Rhianon Sparkes Szollosy for inviting us to come on this journey to explore the benefits of learning by teaching using technology.

Correspondence concerning this article should be addressed to Krista R. Muis, McGill University, 3700 McTavish Street, Montreal, Quebec, H3A 1Y2. E-mail: krista.muis@mcgill.ca

¹ One teacher previously developed her own approach to learning by teaching by having students create concept maps of their understanding of a content area. After students developed their concept maps, students developed teaching videos to explain their understanding of that content area. We used this approach as the foundation for the current study. The teachers also chose the problem as well as the applications for solving the problem.

² All students are required to purchase their own iPad for school purposes, which they use on a daily basis for all content areas.

cally, the purpose of our research was to explore the effects of learning by preparing to teach compared with the effects of learning for learning (control condition) on students' problem representations, self-regulatory processes and learning outcomes during complex mathematics problem solving. We describe relevant theoretical and empirical work next.

Learning by Teaching

Fiorella and Mayer (2013) define learning by teaching as a learning environment in which a student is given the role of the teacher and is asked to teach academic content to others for instructional purposes. Others may include peers or computer-based agents (Biswas et al., 2010; Roscoe & Chi, 2007). Four lines of research have been conducted within the learning by teaching paradigm, including learning by preparing to teach (Annis, 1983; Bargh & Schul, 1980; Fiorella & Mayer, 2013; Renkl, 1995), learning by (actually) teaching (Annis, 1983; Fiorella & Mayer, 2013, 2014), learning through peer tutoring (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; De Backer, Van Keer, & Valcke, 2012; King, Staffieri, & Adelgais, 1998; Roscoe & Chi, 2007), and teachable agents, such as computer-based agents (Biswas et al., 2005, 2010; Roscoe, Segedy, Sulcer, Jeong, & Biswas, 2013).

We developed an intervention wherein elementary students used tablet applications to solve a complex mathematics problem and created a video wherein they explained how to solve the problem to be used to teach others. As previous research has shown, teaching others can be an effective way to enhance learning (e.g., Biswas et al., 2010; King et al., 1998; Palinscar & Brown, 1984; Peets et al., 2009; Roscoe & Chi, 2007) across a wide range of age groups including college (Annis, 1983), high school (Cloward, 1967; Morgan & Toy, 1970), middle school (Jacobson et al., 2001), and elementary school (Fuchs et al., 1996). That is, through learning by teaching, individuals theoretically learn content more deeply by teaching it to others compared with learning the content just for oneself. However, the reasons for this positive effect on learning remain unclear (Fiorella & Mayer, 2013; Galbraith & Winterbottom, 2011; Peets et al., 2009; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003; Roscoe & Chi, 2007), and some research has shown no positive effects (Renkl, 1995) or negative effects (Ehly, Keith, & Bratton, 1987) on learning. In their seminal article, Bargh and Schul (1980) proposed that the expectation of teaching content to others results in a change in the way individuals study that material compared with normal studying for oneself. They argued that to teach, individuals must develop a good understanding of the domain knowledge and then structure that knowledge in a way that can be presented to others. When learning by preparing to teach, students arguably devote more resources toward selecting the most relevant material and organizing it into meaningful representations (Roscoe & Chi, 2007).

Research on learning by preparing to teach provides support for this hypothesis (Annis, 1983; Benware & Deci, 1984; Biswas, Schwartz, & Bransford, 2001; Fiorella & Mayer, 2013). For example, Fiorella and Mayer (2013) explored the relative effects of learning by preparing to teach and by actually teaching on learning. Students were given the task of studying a lesson on the Doppler effect without the expectation of later teaching the material and then took a comprehension test on the same material (control group). Other students were given the same material but

were told they would actually teach the content by preparing a brief video of the material. Half of these students were given the comprehension test immediately after studying (preparation group), whereas the others prepared a lecture and then were given the comprehension test (teaching group). Results revealed that the preparation and teaching groups significantly outperformed the control group on the comprehension test (effect sizes were $d = .82$ for the difference between the teaching vs. control group, and $d = .59$ for the difference between the preparation vs. control group). However, in a second experiment with the same design, they found that only the teaching group outperformed the control group on a 1-week delayed test ($d = .79$).

The majority of research that has explored the effects of learning by teaching has been drawn primarily from the peer-tutoring literature. From the peer-tutoring literature, Roscoe and Chi (2007) propose that tutors benefit from instructing others because they must be able to explain the content to others. To explain content well to others requires that tutors are able to evaluate their own understanding, gaps in knowledge, or confusions that arise during learning. They must also be able to recover from those confusions, all of which requires substantial self-monitoring and evaluation (King, 1998). Tutors also need to organize the content in well-structured ways to allow them to provide clear explanations. As such, it is likely during the preparing to teach phase that learning gains occur given that tutors must organize the content in ways that allow them to effectively teach it to others. According to Biswas et al. (2010), this initial structuring of knowledge is likely fostered through self-explanations. Individuals are more likely to engage in these self-explanations to ensure they can teach the content to others. As previous research has shown, when individuals engage in more self-explanation processes, this should facilitate learning (e.g., Chi, DeLeeuw, Chiu, & LaVancher, 1994; Matthews & Rittle-Johnson, 2009) via increased metacognitive processes (Kwon & Jonassen, 2011).

For example, De Backer et al. (2012) explored the role of reciprocal peer tutoring to promote university students' metacognitive knowledge and regulation skills. Students enrolled in an instructional science course engaged in reciprocal peer tutoring over eight sessions with actual course material. Using a pretest-posttest design, students self-reported their metacognitive knowledge and regulation skills. Additionally, a think aloud protocol was used to measure students' actual use of metacognitive strategies (orientation, planning, monitoring, and evaluation of learning). Results at posttest revealed no difference in students' self-reported metacognitive knowledge and strategy use, but differences were observed in actual metacognitive strategies used. Specifically, posttest results revealed significant increases in orientation wherein students were more likely to analyze the task ($d = 3.12$), structure task instructions ($d = .75$), and orient themselves to specific content ($d = 1.52$). Differences in metacognitive monitoring were also observed, wherein students focused more on comprehension monitoring ($d = 1.72$) and progress on task ($d = 1.67$), and on understanding ($d = .90$) and elaborating the text ($d = 2.29$). Similarly, students engaged in more evaluation of their learning outcomes at posttest compared with pretest ($d = 2.46$), but no differences were observed in students' planning activities from pretest to posttest.

In another study, King et al. (1998) assigned seventh graders in pairs to one of three peer-tutoring conditions: explanation only,

inquiry plus explanation, or sequenced inquiry plus explanation. Students engaged in both tutor and tutee roles after teacher-led lessons on systems of the human body. Tests were conducted at pretreatment, posttreatment, and 8 weeks following the posttest to assess students' comprehension of factual material as well as their ability to make inferences and integrate material. Students' metacognitive awareness and self-regulation of their use of the tutoring protocol were assessed using self-report scales, which focused on their understanding of the content, implementation of supportive communication techniques, and how well they explained new material. Results revealed that at posttest and 8 weeks following posttest, there were no differences between groups on factual material learned. However, students who engaged in inquiry plus explanation and sequenced inquiry plus explanation performed better on inference and integration tasks compared with students in the explanation only condition (no effect sizes were reported). It is interesting to note that differences between groups were not found for self-reported metacognition, but there was a significant increase in students' metacognitive awareness and self-regulation improved over time.

Despite results from these and other studies, it is still unclear what the underlying mechanisms might entail with regard to why greater learning gains may occur. For example, in King et al.'s (1998) study, there were no differences in self-reported metacognitive strategies despite theoretical assumptions that increases in metacognitive strategies should occur. For De Backer et al.'s (2012) study, the researchers were not able to make substantive causal claims regarding why there were increases in use of metacognitive strategies given the lack of a control group. One explanation, initially proposed by Roscoe and Chi (2007), is that differences in learning processes and outcomes might arise due to knowledge building versus knowledge telling in tutors' behaviors. Knowledge building is defined as a process of metacognitively monitoring one's own knowledge and understanding, integrating new and prior knowledge, and generating new ideas through inference and reasoning. In contrast, knowledge telling is defined as a process of summarizing or restating source materials with little deep reasoning or reflection. As Roscoe (2014) argued, knowledge telling may occur due to tutors' inadequate evaluation of their own understanding of the material. Indeed, Roscoe found that tutors' comprehension monitoring and domain knowledge, as well as tutees' questions, were significant predictors of knowledge building. However, we argue that one central component that has been overlooked in the learning by teaching literature is the initial structuring of the task and content that occurs during the preparing to teach phase (Bargh & Schul, 1980). To elaborate this, we turn to the self-regulated learning literature.

Self-Regulated Learning

Students who self-regulate their learning plan how to approach a learning task, set goals, implement strategies to carry out the task, and evaluate progress and products throughout the learning process. Planning, implementation, and evaluation are key phases in Muis's (2007) theoretical framework. Muis's model was chosen given its focus on metacognition as the hub of self-regulated learning, which is central to mathematics problem solving (Jacobse & Harskamp, 2012; Schoenfeld, 1982). Muis's model was also chosen given its focus on the task definition phase of self-regulated

learning, which is influenced by the instructional context provided for a task. We elaborate this below.

Muis's (2007) Model of Self-Regulated Learning

Based on goal-oriented (Pintrich, 2000) and metacognitively driven (Winne & Hadwin, 2008) models of self-regulated learning, as shown in Figure 1, Muis (2007) proposed a model of self-regulated learning that includes four phases of learning: task definition, planning and goal setting, enactment, and evaluation. According to Muis (2007), in the first phase, task definition, the learner defines the task based on external conditions, such as the instructional context (e.g., instructions provided by teachers to complete the task), task features (e.g., the kind of problem) and internal conditions such as prior knowledge and motivation. It is during this phase of self-regulated learning that learners begin to encode and elaborate the initial givens of a problem and develop a representation of that problem (i.e., develop an understanding of the problem), which are all critical for successful mathematics problem solving (Bédard & Chi, 1992; Chi et al., 1981; Fuchs et al., 2006; Schoenfeld, 1994).

It is important to note that when students are given different instructions to complete a task, their task definitions might differ, which may then lead them to implement different learning strategies to carry out the task (Chevrier, Muis, & Di Leo, 2015; Muis, 2007; Winne & Hadwin, 2008). For example, if students are told they will be given a complex mathematics problem to read first, create a concept map of the problem, solve the problem, and then create a teaching video that explains to others how they solved the problem, their task definitions may differ from students who are told to read the problem, create a concept map, and then just solve the problem. As Bargh and Schul (1980) proposed, students in the teaching condition might define the task as one in which they must develop a good understanding of the mathematics problem and then structure each facet of the problem in a way that can be explained to others. When creating the concept map, students in the teaching condition might include more critical information about the problem, and organize information better in terms of the problem's structure given that the standards these students set for understanding the problem may be higher compared with those set by students in a problem-solving only condition.

These variations in task definitions may then result in differences in the plans and goals that individuals set during the second phase. Planning includes selecting the types of learning strategies to carry out the task and identifying the type of information on which to focus during learning. For example, students in the teaching condition may plan to use more metacognitive strategies to ensure progress on understanding the problem is sufficient and during problem solving to ensure a correct solution to each facet is derived; the specific level of understanding the students set to achieve would be identified as a goal. A goal is modeled as a multifaceted profile of information (Butler & Winne, 1995) and each standard in the profile is used as a basis to compare the products created when carrying out the task.

The third phase begins when a learner implements the learning strategies that were planned to carry out the task. Then, in the last phase, individuals evaluate the successes or failures of each phase or products created for the task, or perceptions about the self or context. Products created during learning are compared with the

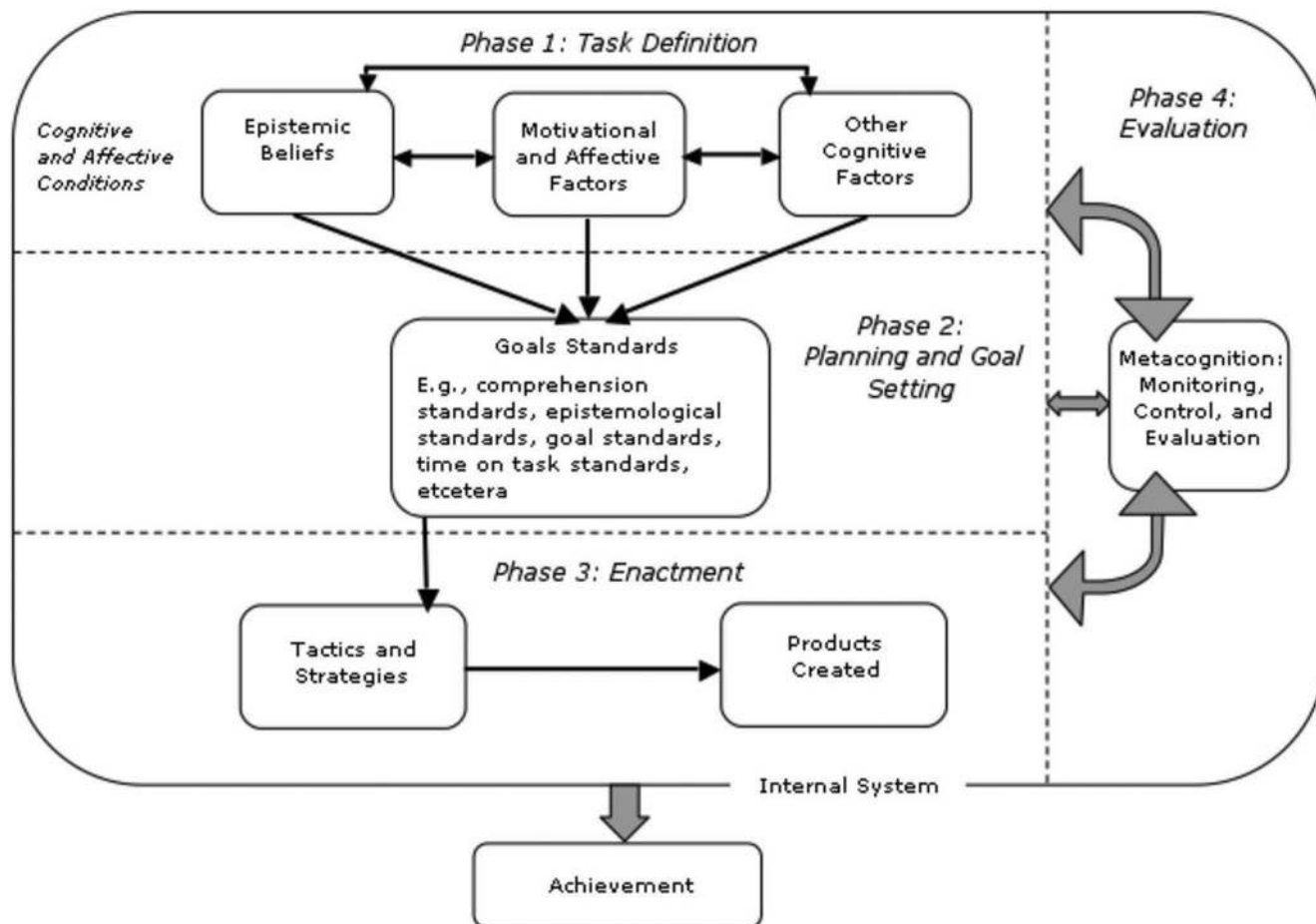


Figure 1. Muis's (2007) model of self-regulated learning. Reprinted from "The Role of Epistemic Beliefs in Self-Regulated Learning, by K. R. Muis, 2007, *Educational Psychologist*, 42, p. 177. Copyright 2007 by Taylor & Francis. Reprinted with permission.

standards set via metacognitive monitoring. If monitoring reveals an inadequacy of one or more of the products created (e.g., an answer is incorrect), a learner may engage in control processes wherein other cognitive strategies are employed to reduce the discrepancy. According to Muis (2007), metacognitive processes can occur during all four phases of self-regulated learning. That is, monitoring, control, and reaction/evaluation can be ongoing throughout the learning process, and goals and plans may also change or be updated as feedback about progress becomes available. Moreover, products created across all four phases can feed into other phases, which reflect the cyclical nature of self-regulated learning in her model.

The Current Study

Taken together, Muis's (2007) model provides a theoretical explanation with regard to why differences in learning outcomes might occur when learning by teaching versus learning for learning. Within the broader learning by teaching paradigm, research has focused primarily on learning outcomes (e.g., Biswas et al., 2010; Palinscar & Brown, 1984; Peets et al., 2009; Roscoe & Chi, 2007). Few studies have explored self-reported or actual cognitive

or metacognitive processes during learning by teaching (De Backer et al., 2012; King et al., 1998; Roscoe, 2014; Roscoe & Chi, 2007). Moreover, to our knowledge, no study has been conducted to assess what learners' initial understanding of the content entails and how they structure that content in preparation for teaching. Research is needed wherein individuals' task definitions and cognitive and metacognitive processes are traced as they occur during problem solving across each phase of self-regulated learning. As such, the purpose of this research was to explore whether learners' task definitions and self-regulatory processes differed when learning by preparing to teach versus learning for learning in the context of complex mathematics problem solving in a technology-rich learning environment. The research was conducted in a classroom context during regular school time with a sample of elementary students from two different schools.

From a self-regulated learning perspective, learners' task definitions should differ when learning solely for oneself versus when given the task to teach others. To capture individuals' task definitions, concept maps can be used to evaluate what information learners think is important and how they structure that information (Pintrich, Marx, & Boyle, 1993), a method that has previously

been used in research on mathematics (Laturno, 1994; Williams, 1998). That is, students' concept maps can be analyzed for quantity of information as well as how that information is organized. To date, research within the learning by teaching framework has not explored this possibility. As such, our first research question was, Do students' task definitions differ when learning by preparing to teach versus learning for learning? Students were asked to create a concept map that included four features: blue for the title, red for the first step to solve the problem, black for important information, and green for calculations needed to carry out the problem. We hypothesized that students in the learning by preparing to teach condition would include more important information about the problem in their concept map and would hierarchically structure the information better (i.e., green calculations subsumed under related important information or vice versa) than students who were asked to just solve the problem.

Differences in task definitions should then theoretically result in differences in planning and goal setting (Muis, 2007; Winne & Hadwin, 2008). That is, if students define the task as one in which they need to develop a good understanding of the problem to be able to explain to others how to solve it, then these individuals may, for example, plan to use more metacognitive strategies to ensure sufficient progress and understanding compared with students who are told to simply solve the problem. Students in a teaching condition may also set more goals, like ensuring their work is done well, compared with students who just solve the problem.

Based on differences in planning and goal setting, differences should also arise during the enactment and evaluation phases wherein various cognitive and metacognitive strategies are employed (Muis, 2007). As such, our second research question was, Are there differences in the frequency of self-regulatory processes, such as planning and goal setting, cognitive processes, and metacognitive processes, when solving a complex mathematics problem when learning by preparing to teach versus when learning for learning? To capture self-regulatory processes of planning and goal setting, cognitive strategies, and metacognitive strategies, a think-aloud protocol was used (Azevedo, 2005; Greene & Azevedo, 2009; Muis, 2008). We hypothesized that students in the learning by preparing to teach condition would engage in more planning and goal setting, and use more cognitive and metacognitive processes during problem solving compared with students in the control condition (Bargh & Schul, 1980; De Backer et al., 2012).

Our final research question was, Does learning by preparing to teach result in higher levels of mathematics problem solving achievement compared with learning for learning? Given that conceptual understanding of the problem and metacognitive processes are central to successful mathematics problem solving (Chi et al., 1981; Fuchs et al., 2006; Schoenfeld, 1994), we hypothesized that students in the learning by preparing to teach condition would have a higher achievement score on the complex mathematics problem compared with students who just solved the problem.

Finally, given that previous research has found gender differences in self-regulatory processes at the elementary-school level (Zimmerman & Martinez-Pons, 1990), gender was included as a variable. We hypothesized that females would engage in more self-regulatory processes compared with males (Zimmerman & Martinez-Pons, 1990) and, as a result, have a higher achievement score on the mathematics problem (Hyde, Fennema, & Lamon, 1990; Voyer & Voyer, 2014). As previous research has found relations between prior knowledge and self-regulated learning (Zimmerman & Martinez-Pons, 1990), prior knowledge in mathematics was included as a covariate for all analyses.

Method

Participants

Eighty-two students were invited to participate from two different schools across four different classrooms, and 78 agreed ($n = 34$ females, 0% attrition rate, and parents provided consent and students provided assent). Students were from the same English public school board in the province of Quebec, Canada. Both schools were bilingual wherein students spent 50% of their time learning in English and the other 50% learning in French. Seventy-six students were first-language English (EFL), and two were first-language French (FFL) but were fully fluent in English. Of the 78 students, 75 were Caucasian, two were Indo Canadian, and one was African Canadian. See Table 1 for a complete summary of the demographic information.

Materials

Demographics. Demographic information was obtained from the parental consent forms, which included students' age (by date

Table 1
Demographics by Frequency or Average

School	Females	Males	EFL	FFL	Age		IEP	LBT	LFL
					<i>M</i>	<i>SD</i>			
School 1	22	18	40	0	11	0.31	5	20	20
School 2	14	24	36	2	11	0.31	4	19	19
Total	36	42	76	2	11	0.31	9	39	39

Note. Individualized education plans (IEP; adapted or modified) are used to help describe and organize the support measures and personalized follow-up that are necessary to help students with special needs progress in their schooling and to foster their success. The IEP is a collaborative process by which an educational plan is created for a student. This plan involves identifying the needs and strengths of the student, creating short- and long-term goals and objectives, and specifying the accommodations or strategies that will be used to achieve those objectives. For the purposes of this study, coding of students' mathematics solution was the same regardless of IEP status. LBT = learning by teaching; LFL = learning for learning; EFL = English first language; FFL = French first language.

of birth), gender, and primary and secondary languages spoken at home.

Prior knowledge. Students' achievement score on a compulsory provincial exam was used to obtain a measure of prior knowledge. The provincial exam was completed 1 week prior to the beginning of the research study (all students in the province must complete this exam, which counts for 30% of their final grade in mathematics). Commencement of the research study was chosen to immediately follow the provincial exam to ensure a valid and standardized assessment of students' prior knowledge. The exam included a series of multiple-choice questions that assessed students' knowledge of the mathematics content covered over the school year. Cronbach's alpha reliability coefficient for the prior knowledge test was .94.

Complex mathematics problem. The situational problem *Start Your Engines* was drawn from the 2009 provincial exam. The objective is to have students develop a coherent solution to a situational problem that meets the following conditions: (a) the procedure required to solve the situational problem is not obvious, because it involves choosing a significant number of previously acquired mathematical concepts and processes and using them in a new way; (b) the situation focuses on obstacles to overcome, which requires various learning strategies; and, (c) the instructions do not suggest a procedure to be followed or the mathematical concepts and processes to be used (Ministère de l'éducation, du loisir et du sport, 2009). In the first phase of problem solving, students read the problem and then developed a concept map of the problem (see next section). In the second phase, students were required to solve the problem, and show all steps and decisions made along the way. For this particular problem, students had to: create a seven-sided polygon for the racetrack design that ranged in length between 4.5 km to 5 km; include at least one acute angle, one obtuse angle, and one angle greater than 180°; create spectator areas with 15 squares per section to seat 120,000 spectators; draw a starting line frieze pattern that was one third white, reflected twice; and calculate the cost of the paint for the starting line given indications of the price per unit.

Mathematics achievement. A rubric (see Appendix A) was developed to score each student's solution to the complex mathematics problem, and total score on each student's solution was used as the measure of mathematics achievement. This particular problem was graded based on the inclusion of all required elements of the task as listed above. Students' calculations for each facet of the problem were also graded. Each element of the problem was given a particular value, and full points were awarded for successfully completing each element. Partial points were given when aspects were missing, and no points were given if an element was completely missing or wrong. For example, if a student created a six-sided polygon, he or she was awarded six points, rather than the full seven points (i.e., one point for each side, but one point was taken off for each additional side over seven). For the perimeter, students were given four points if the track measured between 4.5 and 5.0 km. Two points were awarded if the track was within 0.5 above or below the range. Calculations were given full points if done correctly, and partial points were given if a minor mistake was made (e.g., one number was copied incorrectly from the racetrack to the calculations page). The total number of points to be earned was 50. Krista R. Muis and Cynthia Psaradellis coded 10 students' solutions (randomly selected) to

establish interrater agreement. Interrater agreement was assessed for each facet of the problem (e.g., whether both raters awarded 7 points for the seven-sided polygon, four points for the perimeter). Interrater agreement was 100% ($\kappa = 1.0$). Cynthia Psaradellis then coded all remaining solutions. Coding was blinded to ensure no bias.

Concept map. To assess students' understanding of the problem, that is, their task definitions, students used the tablet application *Popplet* to create a concept map (students had been using this concept mapping tool for various content areas for seven months at one school, and for one month at the other school). Students were provided specific criteria developed by the teachers to create their concept map for the mathematics problem. Criteria included using four different colored borders to represent various facets of the problem. Students were told to use a black border to represent important information, a green border for calculations needed to solve various aspects of the problem, a red border for the first question that needed to be answered (some aspects of the problem needed to be solved first before others, so order was important in some instances), and a blue border for the title of the problem. For important information, students were asked to create one popple (i.e., concept bubble) for each piece of important information, which they included inside the popple. For calculations, students were similarly asked to create a popple for each calculation they needed to carry out, and were asked to insert the description of the calculations but not actually solve the problem within the popple. Moreover, prior to the study, the teachers trained the students to create the popplets using the four colors and to organize the information whereby calculations and associated important information are linked.

Coding the concept map. To score the students' concept maps, a rubric was developed (see Appendix B) that included all aspects of the problem. An expert concept map was also created to allow for ease of comparison (see Figure 2). As shown in Figure 2, there were four main aspects of the problem: racetrack design, spectator area, starting line frieze pattern, and cost of starting line paint. We allowed for flexibility in the hierarchical structure of the concept map wherein students could organize red popples first followed by black or green, or vice versa. Key to coding the organizational structure was the information that was linked as well as each popple's color. For example, there were six possible questions students could solve first (coded in red), and students were given one point for correctly identifying one of the six as the first question. If, however, students indicated two first steps (two red popples), they lost 0.5 marks. For each of the four aspects of the problem, there were a specific number of popples that students could create related to that aspect of the problem. Each of those popples was further color-coded as green or black. Students were given one point for each correctly color-coded popple (i.e., both content and color had to match). If a popple was incorrectly colored (e.g., correct calculation description but the popple was black instead of green), students lost 0.5 of a mark. For organization of the concepts, for each correctly linked popple students were given 0.5 of a mark (marks were not deducted for incorrect links). Total possible score on the concept map was 20. Krista R. Muis and Cynthia Psaradellis then coded 10 concept maps to establish interrater agreement based on scores for each aspect of the problem. Interrater agreement was 100% ($\kappa = 1.0$). Given high inter-

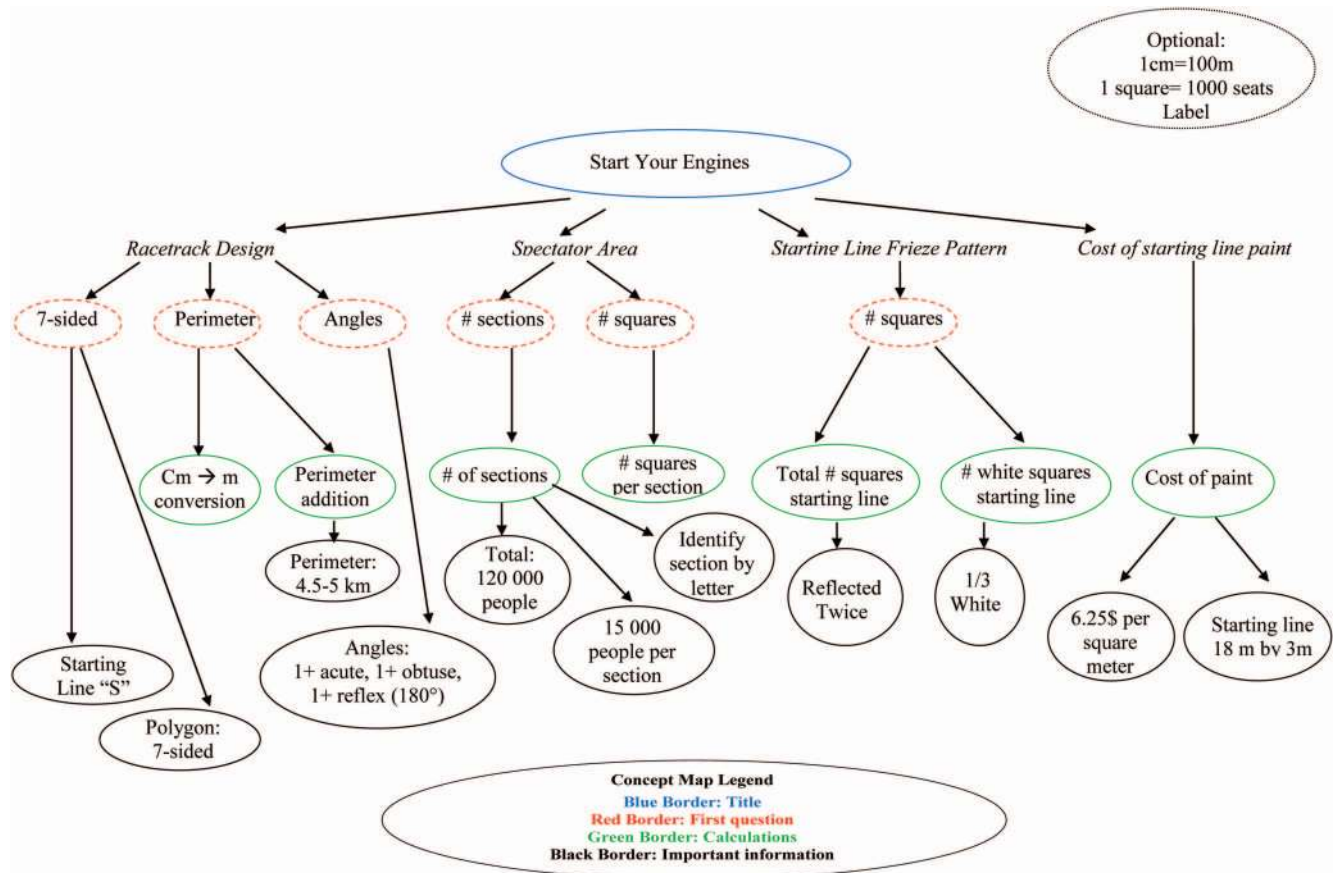


Figure 2. Expert concept map. See the online article for the color version of this figure.

rater agreement, Cynthia Psaradellis then coded the remaining concept maps. Coding was blinded to ensure no bias.

Self-regulatory processes. A think-aloud protocol (Type I protocol; see Ericsson and Simon, 1998) was used to capture students' self-regulatory processes as they read the problem, developed their concept map, and solved the complex mathematics problem. Students were instructed to state out loud whatever came to mind. According to Ericsson and Simon (1998), a concurrent Type I think-aloud protocol, which involves thinking out loud while completing a task, does not change the sequence of thoughts and does not affect performance. As such, think-aloud protocols provide a more accurate assessment of individuals' self-regulatory processes as they occur, compared with retrospective self-reports of strategies used during problem solving (see Winne, Jamieson-Noel, & Muis, 2002).

Coding of self-regulatory processes. Students' think-aloud protocols were transcribed verbatim. Think-aloud protocols ranged from 90 min to 4.5 hr, which resulted in 1,288 single-spaced pages of text. Greene and Azevedo's (2009) self-regulatory processes coding scheme, as well as Schoenfeld's (1982) and Muis's (2007) theoretical models were used as guides to develop a micro-macro-level coding scheme specific to mathematics problem solving. Macrolevel codes are general self-regulatory processes whereas microlevel codes are specific self-regulatory processes. Six macrolevel processes considered from Schoenfeld's (1982) model in-

cluded reading, analyzing, exploring, planning, implement, and verifying. Greene and Azevedo's (2009) model included five macrolevel and 35 microlevel processes. The macrolevel processes included planning, monitoring, strategy use, as well as handling of task difficulty and demands, and interest. From Muis's (2007) model, we considered four macrolevel processes, which reflected the four phases of her model: task definition, planning and goal setting, enactment, and evaluation.

Based on the codes developed from these three models, five trained research assistants and Krista R. Muis then coded two transcripts (102 pages) to further refine the coding scheme and establish interrater agreement. Interrater agreement was established at 82%, and disagreements were resolved through discussion. Four more weeks were spent refining the coding scheme, and 20 microlevel codes emerged and were categorized along four macrolevel codes based on Muis's (2007) model: task definition, planning and goal setting, enactment, and monitoring and evaluation. Definitions of each of the macrolevel and microlevel codes, along with examples drawn from the transcripts are presented in Table 2.

Once these codes were established, Krista R. Muis then selected three of the longest transcripts from each of the four classes, plus one of the shorter transcripts from each class to ensure comparability across length. The two original protocols used to develop the coding scheme were included in the 16 transcripts chosen, which

Table 2
Self-Regulated Learning Coding Scheme for Mathematics Problem Solving

Level (macro)/micro	Definition	Examples
Level 1: Task definition	A learner generates a perception about the task, context, and the self in relation to the task. External and internal conditions play a major role.	Prior knowledge activation, beliefs, motivation, and knowledge of strategies are activated during this level.
Prior knowledge activation	Searching for or explicitly recalling relevant prior knowledge.	"A reflex angle. That's more than 180 degrees." "5 km, which is short for kilometer."
Identifying important information	Recognizing the usefulness of information.	"So that will help us figure out how many people are in each row . . ."
Level 2: Planning and goal setting	The learner begins to devise a plan to solve the problem and sets goals.	e.g., planning to use means-ends analysis, trying trial and error, identifying which part of the problem to solve first, solving it within a specific amount of time.
Making/restating a plan	Stating what approach will be taken, what strategy will be used to solve the problem, or what part of the problem will be solved in some sequence. This includes restating plans.	"First, I have to figure out how many are in each row, then I can figure out how many people fit in each row to fit 120,000 people." "Lets just do trial and error."
Setting/restating a goal	A goal is modeled as a multifaceted profile of information, and each standard in the profile is used as a basis to compare the products created when engaged in the activity. This includes restating goals.	"We have to have an acute angle, obtuse angle, and one reflex angle." "I don't want to spend too much time figuring out the track." [time goal] "I want to make sure my calculations are neat."
Level 3: Enactment	Enactment occurs when the learner begins to work on the task by applying tactics or strategies chosen for the task.	
Hypothesizing	Making predictions.	"The next one is probably going to tell us the information about the design."
Summarizing	Summarizing what was just read in the problem statement.	"Next, the spectator seating area, must be divided into sections each section must have seats for 15,000 people. So there, each section has 15,000 people." "The starting line must be painted with a frieze pattern, this pattern is a rectangular design that has to be, that has been reflected twice, so it has to be reflected twice."
Help seeking	Asking for help from a teacher, peer, or other source. Help seeking for information VERSUS help seeking for evaluation.	[turns to teacher and asks a question] "But what if my track isn't exactly 5 km?" "So we're supposed to do something like this?" "Is this correct?"
Coordinating informational sources	Using other sources of information to help solve the problem.	"Let's go back to our popplet." [Popplet includes the concept map, and learner is going back to the concept map he created to help solve the problem].
Level 3: Enactment continued	Enactment occurs when the learner begins to work on the task by applying tactics or strategies chosen for the task.	
Highlighting/labeling /coloring/ drawing/writing	Highlighting information, labeling information as part of the problem-solving process, or taking notes in reference to the problem. Making a drawing to assist learning or as part of solving the problem.	"We can put the starting line just like right there." [labeling] [you can hear the learner's pencil] "So its two sides, 2 sides, 3, kind of look like a good drawing [evaluating quality of drawing], 4." "This is a reflex angle."
Calculating/measuring	Solving equations, measuring, or other similar features.	[adding up the sides] "10 so that's like 1 km plus 1 km and 400 meters . . ." "4.4 plus 3.1 plus . . . equals . . ." "I'm just going to re-read this . . ."
Re-reading	Re-reading a section of the problem, word for word. Important that it is word for word, otherwise it is summarizing.	
Making inferences	Making inferences based on information read or products created from solving the problem. (self-explanation) Explaining why something was done. Key word is <i>because</i> .	"So it doesn't say it has to be irregular or regular." "I'm just, I'm multiplying 18 by 6.25 [calculating] because there are 6.25 per white squares." [self-explanation]

(table continues)

Table 2 (continued)

Level (macro)/micro	Definition	Examples
Goal-directed search Level 4: Monitoring and evaluation	Various types of reactions and reflections are carried out to evaluate the successes or failures of each level or products created for the task, or perceptions about the self or context. Reaction and reflection also includes judgments and evaluations of performance on a task as well as the attributions for success or failure.	"I'm looking for another thing that might be useful." Products created are compared to the standards set via metacognitive monitoring. Monitoring and evaluation can include any facet listed above (e.g., progress, motivation, plans, goals, strategies, products like answers or drawings made).
Self-questioning	Posing a question.	"But how much is that?" "What is the most important thing?" "So how do we turn meters into km?"
Monitoring	Monitoring something relative to goals.	"I'm not sure there is a reflex angle in my drawing. Let me check." "I might forget that 'each section must have seats for 15,000 people.'" [learner is counting the number of sides for the polygon] "So we have 1 side, 2 side, 3 side, 4 side, 5 side, 6 side, 7 sides."
Judgment of learning	Learner is aware that something is unknown, not fully understood, or difficult to do.	"That would be an acute angle, which is kind of hard to draw, this is hard to draw." "I don't really understand this." "I'm not sure." "This is going to be very hard to figure out."
Level 4: Monitoring and evaluation continued	Various types of reactions and reflections are carried out to evaluate the successes or failures of each level or products created for the task, or perceptions about the self or context. Reaction and reflection also includes judgments and evaluations of performance on a task as well as the attributions for success or failure.	Products created are compared with the standards set via metacognitive monitoring. Monitoring and evaluation can include any facet listed above (e.g., progress, motivation, plans, goals, strategies, products like answers or drawings made).
Self-correcting	Correcting one's mistakes.	"Here are 4 km. Not 4 km. Sorry, 400 meters." "So the first thing was the track had to be has to be a 4-sided [summarizing], not a 4-sided sorry a 7-sided polygon [self-correcting]." "Oops, that was actually an obtuse angle."
Evaluation	Judging whether goals have been met, whether a particular strategy is working, whether the answer is correct, whether the work is neat, etc. Judgment of all facets that fall under monitoring.	After counting the number of sides of the polygon, the learner states, "Yes, I have 7 sides. Okay, we're good." "I measured the wrong thing by accident." "That's not very neat."
Control	Changing strategy when monitoring or evaluation results in a determination that goal has not been met.	[after judging that polygon was not 7-sided] "I'm just going to erase this. It has to be a 7-sided polygon so lets do a different one."
Task difficulty	Statements reflecting the difficulty or easiness of a task.	"This is difficult." "This is easy."

resulted in a total of 472 pages of transcripts. Krista R. Muis and five research assistants then spent an additional 8 weeks working together to establish and modify the coding scheme, coding, and then recoding the transcripts until an acceptable level of interrater reliability was achieved. At the end of this process, interrater agreement was 85% for the 16 transcripts ($\kappa = .68$ overall, with similar values for each of the macro categories, although planning and goal setting was the highest at .75).

The frequency with which the original four macrolevel and 20 microlevel codes occurred was then calculated. Following Greene and Azevedo (2009), codes with low frequencies were removed (e.g., averages less than three over a 4-hr period). The modified coding scheme consisted of three macrolevel and 12 microlevel codes. The first macrolevel code was labeled planning and goal setting in which two microlevel codes, plans and goals, were

included. The second macrolevel code was cognitive processes, which included summarizing, help seeking for information, help seeking for evaluation, calculating, coloring, and rereading. The third macrolevel code was metacognition, which included monitoring, judgment of learning, evaluation, and control. The frequency with which students used each of the microlevel codes were then summed within each of the three macrolevel codes, which were used in subsequent analyses. Coding was blinded to ensure no bias.

Procedure

Students at each school were randomly assigned to one of the two conditions: learning by preparing to teach or the control condition (learning for learning). One day prior to solving the

mathematics problem, Krista R. Muis trained students to think out loud (a script was used to ensure identical training for all students). Thinking out loud was described as, “*saying out loud everything that you say to yourself silently.*” The students then heard a practice think-aloud audio file that modeled what not to do followed by an appropriate think out loud example. That is, the second attempt included intermediate steps and spontaneous thoughts. Students then spent 15 min practicing thinking out loud with a short problem.

The next day, students were given the problem to solve. To ensure that all students were thinking out loud, five trained research assistants and Krista R. Muis were present in the classroom to prompt students to continue to think out loud if they were silent for more than 5 seconds (a ratio of approximately four supervised students per prompter, with students sitting at round tables). Students were told that the problem was to be treated as if it were an exam, and they were not allowed to work together or copy each other’s work during problem solving. Barriers were used to minimize noise levels and cheating (which is normal practice for tests done in class). Earpods with microphones were used to capture students’ thought processes, which were recorded on the tablets using the application *Evernote*. Microphones were placed close to students’ mouths to ensure high quality recordings.

In the learning by preparing to teach condition, students were instructed to first read the problem, create a concept map of the problem, solve the problem, and then create a video, using *Doodle Cast*³ (an application designed for teaching a lesson), to teach other students how to solve the problem. Students in this condition were told that when they developed their teaching video they needed to explain all steps involved, explain how they solved each step, and were told they could use all materials they created to solve the problem (e.g., their concept maps, calculations). In the control condition, students were instructed to first read the problem, create a concept map of the problem, and then solve the problem. Students in both conditions were told on many occasions not to tell other students about their task and were seated in different classrooms to ensure no confounding. Both conditions were conducted each day consecutively and counterbalanced across times such as late morning or early afternoon.

Once students were told the explicit instructions for their condition, audio recording began, and students read the problem out loud. In both conditions, students then completed the concept map using *Popplet*, a concept-mapping application, to construct their concept maps, and then solved the problem. All work was done during regular class time, and time spent on task for both learning conditions was equivalent (approximately 1.0 to 1.5 hours each day over 3 to 4 days). Students took as much time as they needed to solve the problem but to ensure equivalent time on task across the two conditions, time on task was measured for each student (as a function of the length of their think-aloud recordings). During problem solving, students recorded their calculations and notes in *Noteshelf*, and were provided several copies of the racetrack design on which to draw their work. Once students solved the problem, they were asked to submit all work, which was then scored for correctness. That is, all students’ work was scored and coded after completion of the learning phase. No materials were collected after this phase (i.e., videos were not analyzed for correctness). To thank students for their participation, each student received a \$15 iTunes card.

Table 3
Zero-Order Correlations

Variables	1	2	3	4	5
1. Prior knowledge					
2. Planning and goal setting	.13				
3. Cognitive strategies	.40**	.39**			
4. Metacognitive strategies	.26**	.48**	.58**		
5. Concept map	.29*	.14	.37**	.40**	
6. Mathematics achievement	.46**	.21	.45**	.46**	.33**

* $p < .05$. ** $p < .01$.

After submission of their work, students in the learning by preparing to teach condition developed their teaching video with *Doodle Cast* using the same materials they developed for solving the problem. The researchers did not collect the videos given that the focus was on learning by preparing to teach; however, the teachers used them for pedagogical purposes.

Results

Preliminary Analyses

Skewness and kurtosis values were examined for normality for all variables. For kurtosis, all variables were within an acceptable range (using *Tabachnick & Fidell’s*, 2013, criteria of < 13). For skewness, with the exception of planning and goal setting (5.29), variables were within an acceptable range. Because the measurement of plans and goals was on a ratio scale, scores were not transformed (see *Tabachnick & Fidell*, 2013).

We then examined whether there were differences across schools on each of the variables, and ICCs were calculated. No differences were found on any of the variables (all $p > .05$), and all ICCs were below .05. As such, the two schools were combined into one overall sample. Group differences as a function of learning condition were then examined for time on task and prior knowledge. As expected, there were no differences between the learning by preparing to teach intervention group compared with the control group on time on task (average time on task for both groups was slightly less than 2 hr, $p = .78$) or on prior knowledge ($p = 1.00$). However, as expected, gender differences were found for prior knowledge, $F(1, 76) = 4.61$, $p < .05$, $\eta^2 = .06$, whereby females had a higher level of prior knowledge than males. As such, prior knowledge was used as a covariate in all subsequent analyses.

Correlations Among Variables

Table 3 presents the zero-order correlations for all variables: prior knowledge, planning and goal setting, cognitive strategies, metacognitive strategies, concept map score, and mathematics problem solving achievement score. Prior knowledge was significantly positively related to concept map score, $r(69) = .29$, $p < .05$; cognitive strategies, $r(78) = .40$, $p < .001$; metacognitive

³ All 78 students had been trained to use *Doodle Cast* 3 weeks prior to the study wherein they developed a teaching video using different content (i.e., social studies). This was done with teachers and was not associated with this study in particular.

Table 4
Means and Standard Deviations for Each Variable as a Function of Learning Condition by Gender (Raw Frequencies Shown Unless Otherwise Noted)

Variables	Learning by preparing to teach		Learning for learning	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Planning and goal setting				
Females	74.59	47.68	80.38	55.06
Males	30.41	22.40	29.77	25.02
Total	49.67	41.56	51.08	47.18
Cognitive strategies				
Females	114.65	41.99	115.50	56.21
Males	76.32	56.31	61.00	31.27
Total	93.03	53.53	83.95	50.79
Metacognitive strategies				
Females	81.94	35.64	60.25	26.04
Males	51.50	28.13	42.14	23.46
Total	64.77	34.73	49.76	25.88
Concept map (in %)				
Females	49.13	21.76	43.50	16.09
Males	40.55	14.53	30.00	14.22
Total	44.23	18.21	36.14	16.35
Mathematics achievement (in %)				
Females	91.72	6.22	86.41	11.51
Males	85.89	10.96	76.55	11.76
Total	88.43	9.55	80.70	12.51

strategies, $r(78) = .26, p < .01$; and mathematics achievement, $r(78) = .46, p < .001$. Cognitive strategies were significantly positively related to concept map score, $r(67) = 0.37, p < .01$, as well as with mathematics achievement, $r(76) = 0.45, p < .01$. Metacognitive strategies were significantly positively related to concept map score, $r(67) = 0.40, p < .01$, and mathematics achievement, $r(76) = 0.46, p < .01$. Concept map score was significantly positively related to mathematics achievement, $r(69) = 0.33, p < .01$. These results suggest that students who used more cognitive and metacognitive strategies developed better concept maps and achieved higher score in mathematics problem solving. Moreover, students who developed better concept maps had higher levels of achievement in mathematics problem solving.

Main Analyses Overview

To answer the first research question, whether students' task definitions differ when learning by preparing to teach versus learning for learning, a two-way analysis of covariance (ANCOVA) was used. The two independent variables included the learning condition (learning by preparing to teach vs. control condition) and gender (male vs. female). The dependent variable was students' score on the concept map, which was scored for both content and organization of concepts. For the second research question, whether there are differences in the frequency of self-regulatory processes when learning by preparing to teach versus when learning for learning, a two-way multivariate analysis of covariance was used. The two independent variables were learning condition (learning by preparing to teach vs. control condition) and gender (male vs. female). The dependent variables were planning and goal setting, cognitive learning strategies, and metacognitive strategies. For the third research question, whether students in the learning by preparing to teach condition attained higher levels of mathematics problem

solving achievement compared with students in the control condition, a two-way ANCOVA was used. The two independent variables were learning condition (learning by preparing to teach vs. control condition) and gender (male vs. female). The dependent variable was students' total score on the mathematics problem. For all analyses, the covariate was students' prior knowledge and, with the exception of the analysis for differences in planning and goal setting (power was only .47), all analyses achieved a power of .72 or higher.

Students' Understanding (Task Definitions) of the Problem

Table 4 presents the means and standard deviations for both conditions as a function of gender for all variables. The first research question addressed whether there were differences in students' understanding of the problem as a function of learning condition, controlling for prior knowledge. To assess this, students developed a concept map of the problem that was coded for both content and organization of that content. Examples of students' concept maps are presented in Figures 3 and 4. As shown in Figure 3, this particular student included the title and all important information, but failed to include the first step needed to solve the problem (red popple) and several popples were colored black when they should be colored green. Despite the missing aspects of the concept map, the organizational structure of the problem was accurate. In Figure 4, this particular student included five of the seven calculation popples (in green), the first step (red), and all important information. However, the concept map was not perfectly structured.

Analyses of students' concept map scores revealed a main effect of learning condition, $F(1, 63) = 4.30, p = .042, \eta^2 = .064$, a main effect of gender, $F(1, 63) = 6.03, p = .017, \eta^2 = .087$, but

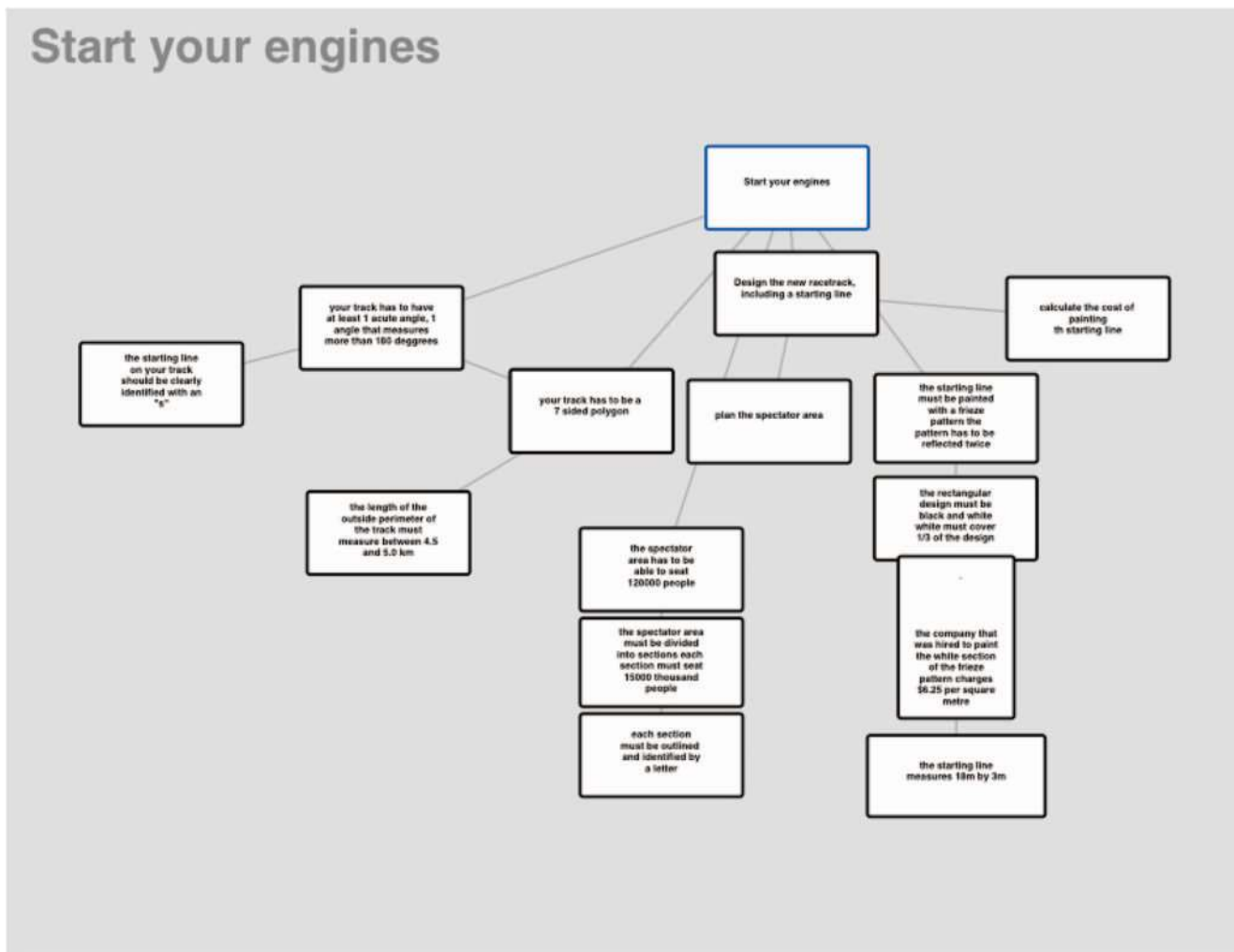


Figure 3. Sample student concept map. See the online article for the color version of this figure.

no condition by gender interaction $F(1, 63) = .30, p > .05$. Specifically, consistent with our hypotheses, students in the learning by preparing to teach condition developed a more detailed and better conceptually organized concept map compared with students in the control condition, and females developed a better concept map than males.

Self-Regulatory Processes

The second research question addressed whether there are differences in the frequency of self-regulatory processes when solving a complex mathematics problem as a function of learning condition. The omnibus multivariate test was significant, $F(3, 72) = 54.19, p < .001$, Wilk's $\Lambda = .30, \eta^2 = .70$. Analyses of students' planning and goal setting revealed a main effect for gender, $F(1, 72) = 26.53, p < .001, \eta^2 = .27$, but no effect for learning condition, $F(1, 72) = .12, p > .05$, and no condition by gender interaction, $F(1, 72) = .060, p > .05$. Specifically, females had higher frequencies of planning and goal setting than males.

Analyses of students' cognitive learning strategies revealed a main effect of gender, $F(1, 72) = 16.03, p < .001, \eta^2 = .18$, but

no effect of condition, $F(1, 72) = .43, p > .05$, and no condition by gender interaction, $F(1, 72) = .47, p > .05$.⁴ Specifically, females had higher frequencies of cognitive learning strategies than males. Finally, analyses of students' metacognitive strategies revealed a main effect of learning condition, $F(1, 72) = 5.93, p = .017, \eta^2 = .076$, a main effect of gender, $F(1, 72) = 11.46, p < .001, \eta^2 = .14$, but no condition by gender interaction, $F(1, 72) = 1.16, p > .05$. Specifically, as hypothesized, students in the learning by preparing to teach condition engaged in more metacognitive processing strategies compared with students in the control condition, and females had higher frequencies of metacognitive strategies than males.

⁴ We also assessed differences across groups for each cognitive strategy to evaluate whether students in the intervention condition used different kinds of strategies compared with students in the control condition. No differences were found.

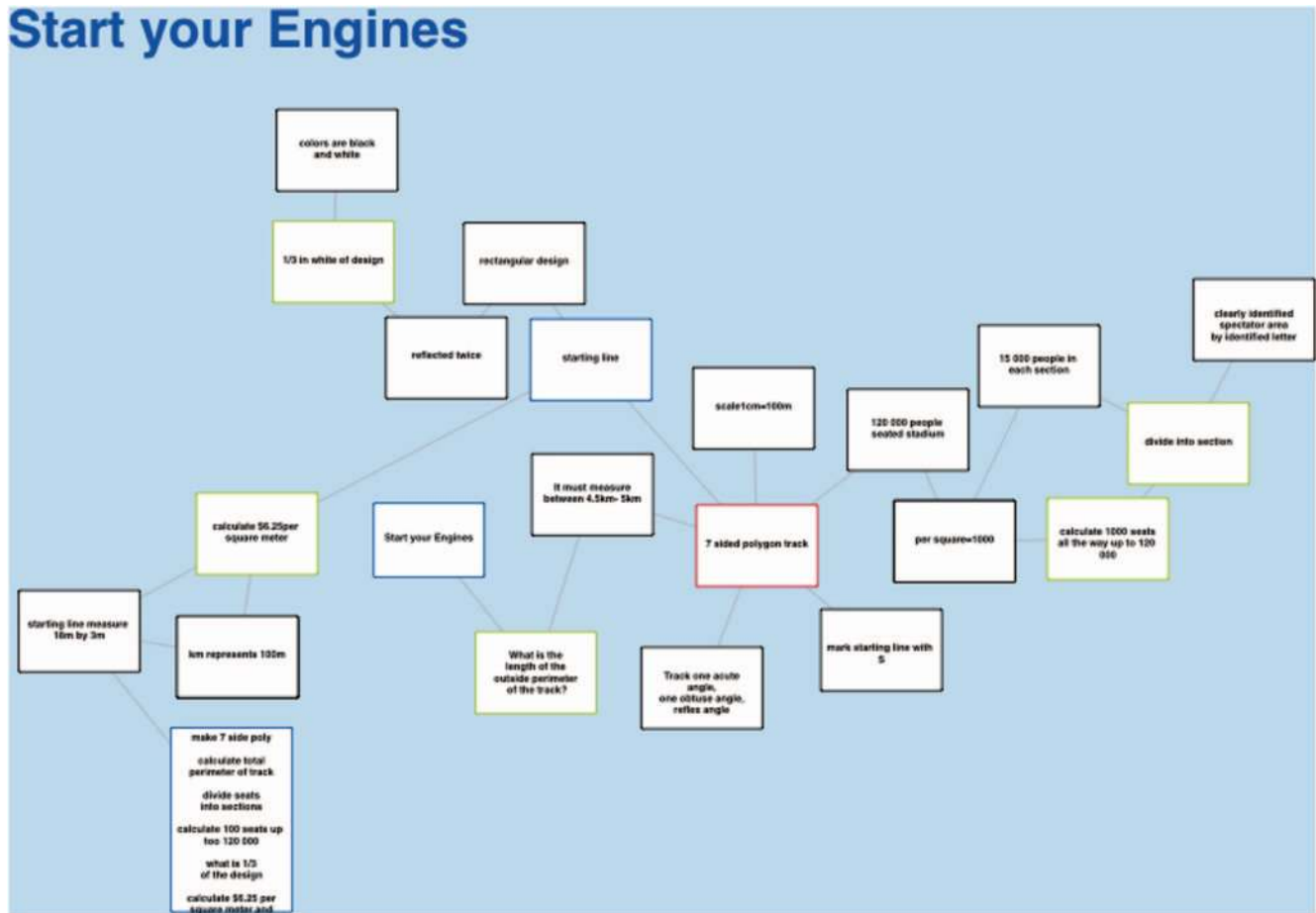


Figure 4. Sample student concept map. See the online article for the color version of this figure.

Mathematics Problem-Solving Achievement

The third question addressed whether students in the learning by preparing to teach condition obtained higher levels of mathematics problem-solving achievement compared with students in the control condition. Analyses of students' scores revealed a main effect of learning condition, $F(1, 72) = 9.37, p = .003, \eta^2 = .12$, a main effect of gender, $F(1, 72) = 8.73, p = .004, \eta^2 = .11$, but no condition by gender interaction, $F(1, 72) = .60, p > .05$. Specifically, as hypothesized, students in the learning by preparing to teach group had higher levels of achievement compared with students in the control group, and females had higher levels of achievement than males. We discuss the theoretical and practical implications of these results next.

Discussion

We developed an intervention grounded in the learning by teaching paradigm in the context of complex mathematics problem solving within a technology-rich learning environment. Our goals were twofold. First, we assessed whether the intervention would foster a greater understanding of the problem and increase self-regulatory processes and learning outcomes compared with a control condition. Second, we empirically evaluated the specific

mechanisms theoretically responsible for improvement in learning during the first phase of learning by teaching, that is, learning by preparing to teach. We hypothesized that students in the learning by preparing to teach intervention group would develop a better conceptualization of the problem via the concept maps they created during the initial phase of problem solving compared with students in the control group. We also hypothesized that students in the learning by preparing to teach group would engage in more planning and goal setting, and use more cognitive and metacognitive strategies during problem solving compared with students in the control group.

Results revealed that the intervention designed within the learning by teaching paradigm effectively increased students' understanding of the problem, metacognitive processes, and learning outcomes during complex mathematics problem solving, but no differences were found in students' planning and goal setting, or in the frequency or type of strategies used to solve the problem. In addition, females developed a more complex concept map, had higher frequencies of planning and goal setting, cognitive learning strategies, and metacognitive strategies, as well as higher levels of achievement compared with males. We discuss each of these results in the context of theoretical and educational implications next.

Theoretical Implications

Understanding of the problem. To date, the majority of previous research within the learning by teaching paradigm has focused on the effects of learning by teaching on learning outcomes (Biswas et al., 2010; Palinscar & Brown, 1984; Peets et al., 2009; Roscoe & Chi, 2007). Although some studies have explored whether learners organize and structure the content differently when learning by teaching, interpretations of these differences were based on learners' retrospective self-reports, on transfer tasks that required a better organizational structure of the content for greater performance, or on delayed performance tests (see Fiorella & Mayer, 2013). To better understand why learning by preparing to teach can benefit learners, a direct test of learners' organizational structure of the content was necessary. Our research addressed this gap in the literature. Importantly, results from this research provide evidence that, when asked to develop a concept map of the problem, learners do indeed organize content better when learning by preparing to teach versus when solely learning for learning.

That is, in our study, students in the teaching condition included substantially more information about the problem (given the large effect size; see Lipsey et al., 2012) and linked concepts better than students in the control condition. Specifically, the mean for the intervention group was higher than 69.15% of the scores in the control group given the overall 8% difference between the two group means. These results are consistent with Bargh and Schul's (1980) initial hypothesis that individuals must develop a good understanding of the problem and then structure that problem in a way that can be presented to others. Given the positive correlation between the quality of the concept map and problem solving achievement, the development of a good understanding of the problem may have been one critical factor for successful problem solving. It may also be the case that the concept maps served as a guideline for solving the problem. Indeed, students frequently referred back to the concept map when solving the problem (coordinating informational sources in our coding scheme) to assess which step they had completed and what they needed to solve next. Although there was no difference in the frequency with which students referred back to the concept map between groups, a higher quality concept map may have benefitted students in the teaching condition.

Planning, goal setting, and cognitive strategies. Given differences in task definitions between the two groups, we also expected that students in the teaching condition would engage in more planning and goal setting during learning compared with students in the control group. Specifically, we hypothesized that students in the teaching condition would plan to use more metacognitive strategies to ensure progress on understanding and completing the task, and set more goals for completing the problem compared with students in the control condition. Counter to our hypotheses, frequency of planning and goal setting did not differ between the two groups. This lack of difference is consistent with De Backer et al.'s (2012) study, wherein they found no differences in undergraduate students' use of planning strategies. In De Backer's study, planning occurred infrequently, whereas in our study, students engaged in nearly as much planning and goal setting as they did in use of cognitive and metacognitive strategies. We interpret

these results to suggest that perhaps due to the complexity of the problem, all students engaged in a high level of planning and goal setting to help them sort out the complexities of the problem (e.g., "I'll do this step first, and then try that one").

These results may also suggest that being given the task of teaching content to others does not affect the planning and goal setting stage of self-regulated learning during complex problem solving. We do not believe this to be the case. Rather, it could be that students set different kinds of goals or make different kinds of plans, depending on their learning condition. In this regard, frequency of planning and goal setting may not differ; differences may lie in the quality of plans and goals set. Future research should explore this possibility. Alternatively, perhaps due to the increase in the cognitive demands of the task, and the increase in metacognitive processes that occurred in the teaching group, students did not have the cognitive resources needed to be more planful during problem solving. Given that cognitive load was not measured during learning, our interpretations are highly speculative and require future research to explore these possibilities empirically. Additionally, it is important to note that power to detect a difference for this particular variable was less than optimal. However, we also note that frequency of planning and goal setting was actually slightly higher for learners in the control condition compared with the intervention condition. Certainly, larger sample sizes should be considered in future work in addition to an exploration of the types of plans that are made.

Similarly, we found no differences in the frequency of cognitive strategy use between the two learning conditions. From a theoretical perspective, this may suggest that the task of teaching others how to solve a complex problem does not change students' approaches to solving that problem, at least not for this kind of task. Alternatively, perhaps elementary students do not yet have a rich repertoire of deeper processing cognitive strategies that they could implement to understand and solve the complex problem (Butler & Winne, 1995; Zimmerman & Martinez-Pons, 1990). Rather, it may be the case that students in the teaching condition set higher standards for understanding and solving the problem, which would result in checking answers to ensure correct and complete solutions more frequently than students in the control condition. Unfortunately, we did not directly measure students' standards for problem solving. We suggest that future research explicitly ask students to indicate what their standards entail for solving the problem to assess whether differences arise between these learning conditions. It may be the case that a more fine-grained analysis of various components of self-regulated learning within each phase is necessary to paint a clearer empirical picture of the precise mechanisms involved.

Metacognitive processes and mathematics achievement. Previous research has also theorized that when learning by preparing to teach, individuals may engage in more metacognitive processes to ensure a deep understanding of the content to be able to later explain that content to others (Bargh & Schul, 1980; Fiorella & Mayer, 2013, 2014; Roscoe, 2014; Roscoe & Chi, 2007). Consistent with theoretical predictions, results from this study revealed that students in the learning by preparing to teach intervention engaged in substantially more metacognitive processes than students in the control condition (given the large effect size found for the difference between groups; see Lipsey et al., 2012). That is, the mean for metacognitive processes used for the

intervention group was higher than 71.90% of the scores in the control group; a difference that reflects 1.3 times more metacognitive strategy use in the intervention group compared with the control group.

As such, this study provides empirical evidence that the expectation of teaching changes the way individuals actually engage in learning at the metacognitive level (Bargh & Schul, 1980; Roscoe & Chi, 2007), which is essential for mathematics learning (de Corte et al., 2000; Muis, 2004; Schoenfeld, 1985; Zimmerman & Labuhn, 2012). Indeed, coupled with a better conceptualization of the problem, a higher frequency of metacognitive processes was predictive of better problem solving achievement. In fact, given the large effect size (see Lipsey et al., 2012) associated with the differences between the two groups on mathematics achievement, we infer that our intervention is very promising for improving students' learning processes and outcomes in this specific context. Indeed, the average achievement score for the intervention group was higher than 73.24% of the scores in the control group given the overall 7.73% difference between the group means. It may be that structuring a problem prior to solving it is critical for better problem solving outcomes within a learning by teaching paradigm.

This research also adds to the current literature on self-regulated learning. Learners in the learning by teaching group engaged in more frequent metacognitive processes, likely due to the differences in task definitions or the standard that students set for learning. As Muis (2007) suggests, differences in task definitions and standards for learning lead to differences in the way that individuals approach a learning task. Like previous research (Chevrier et al., 2015), when students are given different learning tasks, like preparing for a multiple-choice test that requires recognition of information to successfully complete the task, versus an inference verification task that requires learners to deeply understand the content to make correct inferences, the standards they set for learning differ, as do the self-regulatory strategies they use to prepare for those tasks. In our study, although we did not directly measure the standards that learners set for understanding and completing the problem, it may be the case that students in the teaching condition set higher standards, which resulted in more metacognitive processes.

Results from this research are also consistent with previous research (Zimmerman & Martinez-Pons, 1990) with regard to differences found between females and males in their use of learning strategies, and on mathematics problem solving performance (Hyde et al., 1990; Voyer & Voyer, 2014). Specifically, females used more learning strategies compared with males, and outperformed males on understanding the problem and on achievement outcomes. It is interesting to note that gains for females and males in the learning by preparing to teach condition were equivalent, which suggests that a gender bias was not at play in fostering improved learning outcomes. This has important educational implications.

Educational Implications

From an instructional perspective, teachers can readily integrate this approach into their daily classrooms. Specifically, we developed our intervention with teachers who recently integrated tech-

nology into their classrooms, or were in the process of doing so. Although spending two hours to solve a complex problem may seem like a luxury, in the province of Quebec this is standard practice. Students must complete three provincially mandated situational problems each year, and teachers spend considerable classroom time having students practice these complex problems, spread out over several days. Practicing these complex problems within a learning by teaching paradigm may help to improve students' ability to solve these complex problems, which mirror the kinds of complex problems people face outside of the school context. Moreover, students as young as 10 to 11 years of age were capable of using the various tablet applications to create concept maps of the problem, solve the problem, and subsequently develop a teaching video. With use of videos, teachers can further assess students' understanding of the problem, and identify gaps in students' understanding or misconceptions that can be subsequently addressed. We speculate that further learning gains can be obtained from actually developing the video, and it is likely the case that students continue to monitor and correct errors during the second phase of learning by teaching. Future research that we plan to conduct will assess this possibility.

Recommendations for Future Research

Further research is also needed to test the additional effects of actually teaching on student learning. As discussed earlier, four lines of research have been conducted within the learning by teaching paradigm, including learning by preparing to teach (e.g., Annis, 1983; Bargh & Schul, 1980; Fiorella & Mayer, 2013; Renkl, 1995), learning by teaching (Annis, 1983; Fiorella & Mayer, 2013, 2014), learning through peer tutoring (Chi et al., 2001; Roscoe, 2014; Roscoe & Chi, 2007), and teachable agents (Biswas et al., 2005, 2010; Roscoe et al., 2013). Although this study only focused on learning by preparing to teach, other studies in the learning by teaching literature have systematically explored the effects of learning by preparing to teach versus actually teaching (Annis, 1983; Fiorella & Mayer, 2013, 2014). Actual teaching includes both components of teaching which are preparing to teach and explaining content to others. These studies found that actually teaching provided additional benefits with regard to learning gains, but more work is necessary to systemically explore why these gains are achieved. Future research should include a learning by teaching group, which will allow students to develop additional teachable self-regulatory skills and an even deeper understanding of the content (Fiorella & Mayer, 2013, 2014). This is important because previous research has shown that elementary school students are poor self-regulated learners (Schoenfeld, 1994; Schunk & Zimmerman, 1997).

Future studies should also assess whether there are motivational and emotional implications in the learning by teaching condition. For example, according to Pekrun's (2006) control-value theory of achievement emotions, high value for a specific task and high perceived control will lead to enjoyment and increased learning (Johnson & Sinatra, 2013; Pekrun, 2006). On the other hand, low task value and low control will lead to boredom and decreased learning (Pekrun, 2006). Thus, future studies should assess students' emotions during learning when preparing to teach, and assess whether they value learning more when they are expected to teach the content. If the expectation

of teaching increases value, this may have additional motivational benefits (Pekrun, 2006).

Conclusion

The purpose of this research was to explore whether learners' task definitions and self-regulatory processes differed when learning by teaching versus learning for learning during complex mathematics problem solving. Previous research has focused on the learning outcomes of learning by teaching. This research advances understanding of how learning by preparing to teach improves learning outcomes by exploring the specific mechanisms involved. These results provide empirical evidence of theoretical predictions that learners who learn by teaching engage in more frequent metacognitive processes, likely due to the differences in task definitions or standards students set for learning (Muis, 2007). Given the authentic nature of our research, carried out in actual classrooms in collaboration with teachers, we believe that this work has important education implications. With the availability of technology increasing at a rapid rate, it is imperative that interventions are developed and empirically assessed to reap the benefits that technology may have to offer.

References

- Annis, L. F. (1983). The processes and effects of peer tutoring. *Human Learning: Journal of Practical Research & Applications*, 2, 39–47. <http://psycnet.apa.org/psycINFO/1984-21611-001>
- Azevedo, R. (2005). Computers as metacognitive tools for enhancing learning. *Educational Psychologist*, 40, 193–197. http://dx.doi.org/10.1207/s15326985sep4004_1
- Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology*, 72, 593–604. <http://dx.doi.org/10.1037/0022-0663.72.5.593>
- Bédard, J., & Chi, M. T. (1992). Expertise. *Current Directions in Psychological Science*, 1, 135–139. <http://dx.doi.org/10.1111/1467-8721.ep10769799>
- Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal*, 21, 755–765. <http://dx.doi.org/10.3102/00028312021004755>
- Biswas, G., Jeong, H., Kinnebrew, J. S., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning*, 5, 123–152. <http://dx.doi.org/10.1142/S1793206810000839>
- Biswas, G., Leelawong, K., Schwartz, D., & Vye, N., & The Teachable Agents Group at Vanderbilt. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence*, 19, 363–392. <http://dx.doi.org/10.1080/08839510590910200>
- Biswas, G., Schwartz, D., & Bransford, J. (2001). Technology support for complex problem solving: From SAD environments to AI. In K. D. Forbus & P. J. Feltoich (Eds.), *Smart machines in education* (pp. 71–98). Menlo Park, CA: AAAI Press.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65, 245–281. <http://dx.doi.org/10.3102/00346543065003245>
- Chevrier, M., Muis, K. R., & Di Leo, I. (2015, April). *Interactions between assessment and epistemic beliefs: Effects on self-regulated learning and outcomes*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Chi, M. T., De Leeuw, N., Chiu, M. H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439–477.
- Chi, M. T., Feltoich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152. http://dx.doi.org/10.1207/s15516709cog0502_2
- Chi, M., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science*, 25, 471–533. http://dx.doi.org/10.1207/s15516709cog2504_1
- Cloward, R. (1967). Studies in tutoring. *Journal of Experimental Education*, 36, 14–25. <http://dx.doi.org/10.1080/00220973.1967.11011022>
- De Backer, L., Van Keer, H., & Valcke, M. (2012). Exploring the potential impact of reciprocal peer tutoring on higher education students' metacognitive knowledge and regulation. *Instructional Science*, 40, 559–588. <http://dx.doi.org/10.1007/s11251-011-9190-5>
- de Corte, E., Verschaffel, L., & Op't Eynde, P. (2000). Self-regulation: A characteristic and a goal of mathematics education. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 687–726). San Diego, CA: Academic Press. <http://dx.doi.org/10.1016/B978-012109890-2/50050-0>
- Durnin, J. H., Perrone, A. E., & MacKay, L. (1997). Teaching problem solving processes in elementary school mathematics. *Journal of Structural Learning & Intelligent Systems*, 13, 53–69.
- Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, 13, 277–287. <http://dx.doi.org/10.1027/1016-9040.13.4.277>
- Ehly, S., Keith, T. Z., & Bratton, B. (1987). The benefits of tutoring: An exploration of expectancy and outcomes. *Contemporary Educational Psychology*, 12, 131–134. [http://dx.doi.org/10.1016/S0361-476X\(87\)80046-2](http://dx.doi.org/10.1016/S0361-476X(87)80046-2)
- Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, 5, 178–186. http://dx.doi.org/10.1207/s15327884mca0503_3
- Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, 38, 281–288. <http://dx.doi.org/10.1016/j.cedpsych.2013.06.001>
- Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, 39, 75–85. <http://dx.doi.org/10.1016/j.cedpsych.2014.01.001>
- Fuchs, L. S., Fuchs, D., Finelli, R., Courey, S. J., Hamlett, C. L., Sones, E. M., & Hope, S. K. (2006). Teaching third graders about real-life mathematical problem solving: A randomized controlled study. *The Elementary School Journal*, 106, 293–311. <http://dx.doi.org/10.1086/503633>
- Fuchs, L., Fuchs, D., Karns, K., Hamlett, C., Dutka, S., & Katzaroff, M. (1996). The relation between student ability and the quality and effectiveness of explanations. *American Educational Research Journal*, 33, 631–664. <http://dx.doi.org/10.3102/00028312033003631>
- Galbraith, J., & Winterbottom, M. (2011). Peer tutoring: What's in it for the tutor? *Educational Studies*, 37, 321–332. <http://dx.doi.org/10.1080/03055698.2010.506330>
- Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, 34, 18–29. <http://dx.doi.org/10.1016/j.cedpsych.2008.05.006>
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139–155. <http://dx.doi.org/10.1037/0033-2909.107.2.139>
- Jacobse, A. E., & Harskamp, E. G. (2012). Towards efficient measurement of metacognition in mathematical problem solving. *Metacognition and Learning*, 7, 133–149. <http://dx.doi.org/10.1007/s11409-012-9088-x>

- Jacobson, J., Thrope, L., Fisher, D., Lapp, D., Frey, N., & Flood, J. (2001). Cross-age tutoring: A literacy improvement approach for struggling adolescent readers. *Journal of Adolescent & Adult Literacy*, 44, 528–536.
- Johnson, M. L., & Sinatra, G. M. (2013). Use of task value instructional inductions for facilitating engagement and conceptual change. *Contemporary Educational Psychology*, 38, 51–63. <http://dx.doi.org/10.1016/j.cedpsych.2012.09.003>
- King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Review*, 10, 57–74. <http://dx.doi.org/10.1023/A:1022858115001>
- King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology*, 90, 134–152. <http://dx.doi.org/10.1037/0022-0663.90.1.134>
- Kwon, K., & Jonassen, D. H. (2011). The influence of reflective self-explanations on problem-solving performance. *Journal of Educational Computing Research*, 44, 247–263. <http://dx.doi.org/10.2190/EC.44.3.a>
- Laturno, J. (1994). The validity of concept maps as a research tool in remedial college mathematics. In D. Kirshner (Ed.), *Proceedings of the sixteenth annual meeting of the North American chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 60–66). Baton Rouge, LA: Louisiana State University.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., . . . Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U. S. Department of Education.
- Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, 104, 1–21. <http://dx.doi.org/10.1016/j.jecp.2008.08.004>
- Ministère de l'éducation, du loisir et du sport. (2009). *Start your engines! Situational problem for Elementary Cycle Three—Year One* [Reference document and teacher's guide]. Quebec, Canada: Author.
- Morgan, R., & Toy, T. (1970). Learning by teaching: A student-to-student compensatory tutoring program in a rural school system and its relevance to the educational cooperative. *The Psychological Record*, 20, 159–169.
- Muis, K. R. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of Educational Research*, 74, 317–377. <http://dx.doi.org/10.3102/00346543074003317>
- Muis, K. R. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist*, 42, 173–190. <http://dx.doi.org/10.1080/00461520701416306>
- Muis, K. R. (2008). Epistemic profiles and self-regulated learning: Examining relations in the context of mathematics problem solving. *Contemporary Educational Psychology*, 33, 177–208. <http://dx.doi.org/10.1016/j.cedpsych.2006.10.012>
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117–175. http://dx.doi.org/10.1207/s1532690xci0102_1
- Peets, A. D., Coderre, S., Wright, B., Jenkins, D., Burak, K., Leskosky, S., & McLaughlin, K. (2009). Involvement in teaching improves learning in medical students: A randomized cross-over study. *BMC Medical Education*, 9, 55. <http://dx.doi.org/10.1186/1472-6920-9-55>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315–341. <http://dx.doi.org/10.1007/s10648-006-9029-9>
- Pintrich, P. R. (2000). An achievement goal theory perspective on issues in motivation terminology, theory, and research. *Contemporary Educational Psychology*, 25, 92–104. <http://dx.doi.org/10.1006/ceps.1999.1017>
- Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63, 167–199. <http://dx.doi.org/10.3102/00346543063002167>
- Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, 88, 144–161. <http://dx.doi.org/10.1037/0022-0663.88.1.144>
- Renkl, A. (1995). Learning for later teaching: An exploration of mediational links between teaching expectancy and learning results. *Learning and Instruction*, 5, 21–36. [http://dx.doi.org/10.1016/0959-4752\(94\)00015-H](http://dx.doi.org/10.1016/0959-4752(94)00015-H)
- Rohrbeck, C., Ginsburg-Block, M., Fantuzzo, J., & Miller, T. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95, 240–257. <http://dx.doi.org/10.1037/0022-0663.95.2.240>
- Roscoe, R. D. (2014). Self-monitoring and knowledge-building in learning by teaching. *Instructional Science*, 42, 327–351. <http://dx.doi.org/10.1007/s11251-013-9283-4>
- Roscoe, R. D., & Chi, M. T. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, 77, 534–574. <http://dx.doi.org/10.3102/0034654307309920>
- Roscoe, R. D., Segedy, J. R., Sulcer, B., Jeong, H., & Biswas, G. (2013). Shallow strategy development in a teachable agent environment designed to support self-regulated learning. *Computers & Education*, 62, 286–297. <http://dx.doi.org/10.1016/j.compedu.2012.11.008>
- Schoenfeld, A. H. (1982). Measures of problem-solving performance and of problem-solving instruction. *Journal for Research in Mathematics Education*, 13, 1–34. <http://dx.doi.org/10.2307/748435>
- Schoenfeld, A. H. (1994). Reflections on doing and teaching mathematics. In A. H. Schoenfeld (Ed.), *Mathematical thinking and problem solving* (pp. 53–75). Hillsdale, NJ: Erlbaum.
- Schunk, D. H., & Ertmer, P. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631–649). San Diego, CA: Academic Press. <http://dx.doi.org/10.1016/B978-012109890-2/50048-2>
- Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist*, 32, 195–208. http://dx.doi.org/10.1207/s15326985ep3204_1
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Allyn Bacon.
- van der Stel, M., & Veenman, M. V. (2010). Development of metacognitive skillfulness: A longitudinal study. *Learning and Individual Differences*, 20, 220–224. <http://dx.doi.org/10.1016/j.lindif.2009.11.005>
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140, 1174–1204. <http://dx.doi.org/10.1037/a0036620>
- Williams, C. G. (1998). Using concept maps to assess conceptual knowledge of function. *Journal for Research in Mathematics Education*, 29, 414–421. <http://dx.doi.org/10.2307/749858>
- Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Mahwah, NJ: Erlbaum.
- Winne, P. H., Jamieson-Noel, D. L., & Muis, K. R. (2002). Methodological issues and advances in researching tactics, strategies, and self-regulated learning. In P. R. Pintrich & M. L. Maehr (Eds.), *Advances in motivation and achievement: New directions in measures and methods* (Vol. 12, pp. 121–155). Greenwich, CT: JAI Press.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41, 64–70. http://dx.doi.org/10.1207/s15430421tip4102_2
- Zimmerman, B. J., & Labuhn, A. (2012). Self-regulation of learning: Process

approaches to personal development. In K. Harris, S. Graham, & T. Urdan, (Eds.), *APA Handbook of educational psychology: Vol. 1. Theories, constructs, and critical issues* (pp. 399–425). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/13273-014>

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51–59. <http://dx.doi.org/10.1037/0022-0663.82.1.51>

Appendix A

Quebec Exam in Mathematics (2009) “Start Your Engines” Marking Rubric

	Your Mark	Total Mark
Racetrack design:		
7-sided polygon		7
Perimeter between 4.5 km and 5 km		4
Measures of each line segment (with ruler & label)		4
1 acute angle, 1 obtuse angle, and 1 reflex [included and correctly labeled]		6
Identifies the starting line with an “S”		1
Spectator area:		
8 sections		1
Letter identification for each section (A, B, etc.)		1
15 squares per section		1
Starting line frieze pattern:		
Rectangular design measuring 6 squares by 3 squares, reflected twice		3
1/3 white and 2/3 black		3
Cost of starting line painting:		
Costs \$112.50		1
Calculations (example):		
50 cm represents 5,000 m		3
$6\text{ cm} + 5\text{ cm} + 4\text{ cm} + 9\text{ cm} + 5\text{ cm} + 10\text{ cm} + 10\text{ cm} = \mathbf{49\text{ cm}}$		3
$15,000 \div 1,000 = \mathbf{15}$ squares		3
$120,000 \div 15,000 = \mathbf{8}$ sections		3
6 squares white, 12 squares black, 18 squares total		3
$\$6.25/\text{m}^2 \times 18\text{ m}^2 = \112.50		3
Total:		50
Percent:		

(Appendices continue)

Appendix B

Concept Map Marking Rubric

	Your Mark	Total Mark
Blue Border Start Your Engines Title [1]		1
Red Border (multiple answers 1 out of 6) [1]		1
7-sided, perimeter, angles, # sections, # squares spectator, # squares line <i>Please write red choice: _____</i>		
Racetrack design: [6]		
Green borders		
Centimeters to meters conversion		1
Perimeter addition		1
Black borders		
Identifies the starting line with an "S"		1
7-sided polygon		1
Perimeter between 4.5 km and 5 km		1
At least 1 acute angle, 1 obtuse angle, and 1 reflex		1
Spectator area: [5]		
Green borders		
Number of sections		1
Number squares per section		1
Black borders		
Total: 120,000 people		1
15,000 people per section		1
Letter identification for each section		1
Starting line frieze pattern: [4]		
Green borders		
Total number squares starting line		1
Number white squares starting line		1
Black borders		
Reflected twice		1
One third white		1
Cost of starting line painting: [3]		
Green borders		
Cost of paint		1
Black borders		
\$6.25 per square meter		1
Starting line 18 m by 3 m		1
Subtotal:		20
<i>-0.5 for each bubble wrong color (___ × -0.5)</i>		-__
<i>+0.5 for correct links between concepts _____</i>		+__
Final New Total:		20
Percent:		

Received November 9, 2014
Revision received June 22, 2015
Accepted June 22, 2015 ■