6-2014

Enhancing Self-monitoring and Self-reflection through a Self-regulatory Skills Intervention Embedded in a Middle School Mathematics Curriculum

Gregory DiGiacomo

Graduate Center, City University of New York

Enhancing Self-monitoring and Self-reflection through a Self-regulatory Skills Intervention Embedded in a Middle School Mathematics Curriculum.

by

Gregory DiGiacomo

A dissertation submitted to the Graduate Faculty in Educational Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2014

This manuscript has been read and accepted for the Graduate Faculty in Educational Psychology:School Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Peggy P. Chen, Ph.D.

Date Chair of Examining Committee

Alpana Bhattacharya, Ph.D

Date Executive Officer

Barry J. Zimmerman, Ph.D.

Georgiana Tryon, Ph.D.

Hefer Bembenutty, Ph.D.

Timothy Cleary, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Enhancing Self-monitoring and Self-reflection through a Self-regulatory Skills Intervention Embedded in a Middle School Mathematics Curriculum.

By

Gregory DiGiacomo

Advisor: Peggy P. Chen, Ph.D.

The purpose of this study was to investigate the effects of a self-regulatory strategy intervention designed to improve participants' calibration accuracy, self-regulatory skills, and math achievement. Monitoring and self-reflection processes were the main focus of this intervention as they are key processes in many well-validated models of self-regulated learning and have been found to impact academic achievement and overall self-regulatory skill (Bol et al., 2010; Dunlosky & Rawson, 2011; Hacker et al., 2008; Nietfeld et al., 2005). The participants were 30 sixth and seventh grade students who were learning about probability as part of their normal math curriculum during the study. They were randomly assigned to a treatment group or a control group. The treatment group received an intervention that was built upon previously successful monitoring and self-regulation interventions.

Results show that participants who received the intervention had higher predictive and postdictive calibration accuracy and higher math performance as compared to the control group, but did not report using more self-regulatory and metacognitive strategy use. Qualitative data suggest that participants use different sources for their calibration judgments depending on how accurate their calibration judgments were and fell largely in line with previous theoretical understandings. The educational implications of the findings for school psychologists and educators were considered.

ACKNOWLEDGEMENTS

I would like to that the many people that helped me complete this journey. First and foremost I would like to thank my advisor Dr. Peggy Chen for spending many hours coaching me through the process and setting high standards for my work. She provided keen insight during the design, implementation and analysis of the study all while maintaining a sense of humor and helping me keep perspective. I am honored to have started my graduate career working with one of the world's preeminent educational psychologists, Dr. Barry Zimmerman, who has provided me valuable input during all stages of this project. I am indebted to him for exposing me to the powerful self-regulated learning paradigm that has influenced my research and professional practice. Dr. Georgiana Tryon also deserves special acknowledgment for the tremendous support, encouragement and advocacy she provided throughout the years. She went above and beyond her job duties by establishing and committing her time to our programs dissertation support group that has helped myself and many of my colleagues meet our goals.

I also want to express my warm gratitude to Dr. Hefer Bembenutty for taking me a special interest in me when I started adjuncting at Queens College and mentoring me through the trials and tribulations of graduate school. He has deepened my understanding of self-regulated learning and psychological research through passionate discussion and a friendly helping hand. I also want to thank Dr. Timothy Cleary for serving on my committee and providing inspiration for this project and guidance on the measurements used in this study.

I would not have been able to get to this point without the friendship, empathy and logistical support my classmates provided me while designing and implementing this study. Your assistance coding my qualitative data and editing my dissertation were invaluable. I am also grateful to Dr. Julie Mayring and Dr. Charles Fasano who encouraged me to conduct this study at their school in the true spirit of the scientist-practitioner traditions of our profession. I also greatly appreciate the students who participated and gave up their lunch to work with me and their parents who allowed their children to participate.

Finally, I would like to thank my family for providing me emotional support when things didn’t go my way and celebrating my accomplishments and progress throughout this process.

TABLE OF CONTENTS

LIST OF TABLES/FIGURES

CHAPTER I

Introduction

Students who skillfully regulate their own learning processes are more likely to succeed academically and to develop a deeper understanding of content and how it relates to the real world. Models of self-regulated learning (SRL) explore how learners activate and sustain cognition, behavior, and affect that are systematically oriented toward attaining their goals (Schunk, Pintrich, & Meece, 2008). Monitoring and self-reflection are integral SRL processes theorized to underlie academic success and are the main focus of this dissertation (Bol, Riggs, Hacker, Dickerson, & Nunnery, 2010; Chen, 2003; Hacker, Bol, Horgan, & Rakow, 2000). Monitoring allows individuals to assess changing task demands, focus awareness on their mistakes, and generate internal feedback, while self-reflection helps individuals interpret feedback, learn from their mistakes and make decisions that enhance subsequent learning and performance (Zimmerman, 2000). However, research has shown that many students who do not have adequate monitoring and reflection skills hinder their ability to regulate themselves and make adaptive decisions during academic pursuits (Dunlosky & Rawson, 2011; Hacker, Bol, & Keener, 2008). This problem is likely compounded by current school contexts that provide little support for the development of these skills, as evidenced by low-achieving students' lack of improvement of monitoring accuracy in naturalistic studies (Bol et al., 2010; Dunlosky & Lipko, 2007; Nietfeld, Cao, & Osborne, 2005). Recent monitoring and self-reflection research focuses on moving beyond building theoretical understandings by designing and implementing interventions that develop these vital skills in students' learning settings such as classrooms (Bol et al., 2010; Hacker, Dunlosky, & Graesser, 1998; Nietfeld et al., 2005). Building on this literature, this current study investigated the effects of an intervention to improve middle-school students self-monitoring and reflection skills while solving mathematical problems.

Two theoretical frameworks guide this proposal: Nelson and Narens' (1990) model of metacognition and Zimmerman's (2000) model of academic self-regulation. *Metacognition* has been defined as the monitoring and control of a lower level of thought by a higher level of thought (Hacker, Bol, & Keener, 2008). Nelson and Narens' (1990) model of metacognition provides a basic theoretical framework to understand how monitoring, reflection, and regulation are related. The model is divided into two levels: the *object-level* is conceptualized as cognition about a given object or event (e.g., thoughts, feelings, procedural knowledge) whereas the *meta-level* is conceptualized as more reflective, higher-order thinking about the object. The two levels reciprocally influence one another through the processes of monitoring and control. Monitoring consists of metacognitively interpreting the status of knowledge or strategies at the cognitive or object-level. *Control,* or regulation, refers to using one's metacognitive knowledge to reflect on and regulate thought and action at the cognitive level (Hacker, Bol, & Keener, 2008). For students to self-regulate their learning effectively, their monitoring processes must be well-*calibrated*, i.e., students' judgments of their current knowledge and skill levels on a particular task and must closely match the actual performance on the task. Accurate calibration underlies effective self-regulation because monitoring generates the internal feedback that students use to adjust and control their learning and performance (Butler & Winne, 1995; Nietfeld et al., 2005). If this internal feedback is inaccurate, attempts to regulate behavior will likely be unsuccessful because students may withdraw their effort, inefficiently allocate their intentional resources, or use inappropriate strategies (Dunlosky & Rawson, 2011; Winne, 2004).

Although the construct of metacognition initially emerged from laboratory-based cognitive psychology research, social-cognitive models of SRL have applied this construct in educational contexts. Zimmerman (2000) put forth a well researched model of SRL that incorporates the processes of monitoring and self-reflection into a three-phase dynamic feedback loop. This model is divided into three sequential phases: (a) forethought, when learners analyze a task and prepare themselves for action; (b) performance, when learners engage with the task; and (c) self-reflection, when learners judge their performance and react to these judgments. The feedback loop of this model indicates that learners gain and use information from one phase to adjust their plans and behavior during the next phases in the learning sequence. Monitoring is a key element in the performance phase that allows learners to judge and assess their understanding of ongoing cognitive activity (Zimmerman, 2000). Strong metacognitive monitoring skills produce more accurate calibration and facilitate the effective regulation of learning by enabling students to gauge progress toward pre-specified goals through internal feedback (Dunlosky & Rawson, 2011; Stahl, Pieschl, & Bromme, 2006; Thiede, Anderson, & Therriault, 2003). If accurate, this feedback improves self-reflection because learners use the internal feedback generated during monitoring to decide if their current approach was effective or needs to be modified. Since monitoring, calibration, and self-reflection skills greatly facilitate the learning process they make excellent targets for intervention, especially considering that research shows that most students need explicit instruction in these skills before they can effectively use them to regulate their own learning (Schunk & Hanson, 1985; Schunk & Zimmerman, 2007).

Although SRL research has generated many successful educational interventions that target regulation of behavior, research on classroom-based calibration interventions has produced mixed results (Cleary & Zimmerman, 2004; Harris, Santangelo, & Graham, 2008; Schunk & Zimmerman, 1998). A number of interventions show that students explicitly trained to monitor their progress metacognitively and reflect on their strategy use showed more accurate calibration as compared to their non-trained peers (Dimmitt & McCormick, 2012; Graham & Harris, 2003; Zimmerman, Moylan, Hudesman, White & Flugman, 2011). However, interventions that only provide practice and feedback have had little success in creating changes in students' calibration accuracy (Bol & Hacker, 2001; Bol et al., 2005; Nietfeld et al., 2005). Therefore, more research is needed to evaluate new and existing monitoring and calibration interventions, and to understand the mechanisms underlying their effects. Successful interventions that target these skills and consume little instructional time are also necessary so that they place minimal demands on today's heavily-burdened students and teachers (Huff & Nietfeld, 2009; Nietfeld, Cao, & Osborne, 2006).

This study tested an intervention designed to promote metacognition and self-regulatory strategy use in middle school students. To strengthen ecological validity, this intervention was designed to improve students' calibration by incorporating SRL into their daily learning of math over time. The main goal of the study is to explore the effects of the intervention, which focused on developing these students' monitoring, reflection, and self-regulation skills during mathematical problem-solving. The study built upon successful monitoring and self-regulation interventions by incorporating their effective elements into one curriculum. A key component of the study was adapted from the structured monitoring and reflection exercises that Nietfeld et al. (2006) and Zimmerman et al. (2011) successfully used to improve calibration accuracy and achievement. Nietfeld et al. (2006) distributed 11 brief weekly monitoring worksheets over the course of a college semester which prompted students to make calibration judgments (administration took approximately 5-10 minutes each week). These worksheets were then reviewed with the class and students were encouraged to reflect on their calibration accuracy. Zimmerman et al. (2011) took a similar approach and provided students with optional monitoring and reflection opportunities for each quiz question that they answered incorrectly in a college-level remedial math class. The current study adapted the monitoring and reflection exercises used in these two studies to a middle school mathematics curriculum in the hopes of improving students' abilities to use monitoring to inform their meta-level understandings of the situation, and ultimately fostering more adaptive academic behavior. Participating students made calibration judgments on a number of math review questions during the course of five training sessions. Graphs of their calibration accuracy over the course of the intervention were provided to give visual feedback about the discrepancy between their judgments and their actual performance (Kitsantas & Zimmerman, 2006; Labuhn et al., 2010). In addition, students learned regulatory strategies from all three phases of Zimmerman's self-regulated learning model (Cleary, Platten, & Nelson, 2008; Perels, Dignath, & Schmitz, 2009; Zimmerman, Bonner, & Kovach, 1996). The end of each session was dedicated to completing worksheets intended to foster reflection about their approach to these problems as well as what strategies they can use to enhance their understanding of the content in their math classes (Zimmerman et al., 2011). These methods were hypothesized to facilitate more adaptive monitoring and reflective processes and enable students to take appropriate regulatory action to correct any inaccuracies in their calibration judgments (Bangert-Drowns et al., 1991; Huff & Nietfeld, 2009; Perels et al., 2009).

The intervention was implemented within the context of a naturally occurring unit of mathematics instruction. Because training occurred over multiple sessions, students had ample opportunities to engage in many cycles of self-regulation and improve their monitoring skills and calibration while using internal and external feedback to fine-tune these processes. Specifically, during the intervention students estimated their confidence about solving math problems correctly both before and after attempting to solve the problems. To the author's knowledge, most calibration interventions have not prompted students to make both pre- and postdictive calibration judgments. It was hypothesized that asking students to estimate their confidence at both times may stimulate superior metacognitive monitoring during the performance phase, leading to more productive reflection processes. This approach may also increase understanding of metacognitive and self-regulatory processes and address the limitations of many interventions that occur during one session or in lab settings that do not provide externally valid contexts for calibration (e.g. Dunlosky and Rawson, 2011; Lin, Moore & Zamrucky, 2001; Lundeberg, & Fox 1991; Hacker, Bol & Keener, 2008; Ramdass & Zimmerman, 2008). In addition, the intervention can be readily adapted to other classrooms and content areas. Although this intervention was implemented during one math unit, the procedures could be extended to other units or integrated into a whole curriculum. It is therefore important to explore ways to help all students improve these skills, which are so critical to learning and academic success.

Finally, the study also addressed the mechanisms underlying calibration processes. Further research on how students monitor and evaluate their work will help psychologists better understand why calibration interventions can improve students' performance (Dimmitt & McCormick, 2012; Mevarech & Fridkin, 2006). The field lacks a substantial theoretical understanding of what types of information students use to form calibration judgments and interpret feedback. Preliminary research suggests that people form judgments based on preconceived beliefs about their skills or irrelevant features of the task rather than on pertinent memory traces, which may explain why these judgments resist change (Bol et al., 2005; Hacker, Bol, & Bahababahani, 2008; Hacker et al., 2000). After the three-week intervention, the PI interviewed each participant individually to gain better insight into the sources of information they use to monitor their performance.

To summarize, this study attempted to address the following research questions:

1. Can a self-regulatory strategy intervention embedded into a middle school mathematics unit improve students' calibration accuracy, self-regulatory skills, and math achievement?
2. Will students in the sixth and seventh grade respond differently to the intervention and will they display variations in self-regulation and metacognition?
3. How do students formulate their metacognitive calibration judgments?

CHAPTER II

Literature Review

"True wisdom is knowing what you don't know"

(Confucius, around 400 B.C./1955)

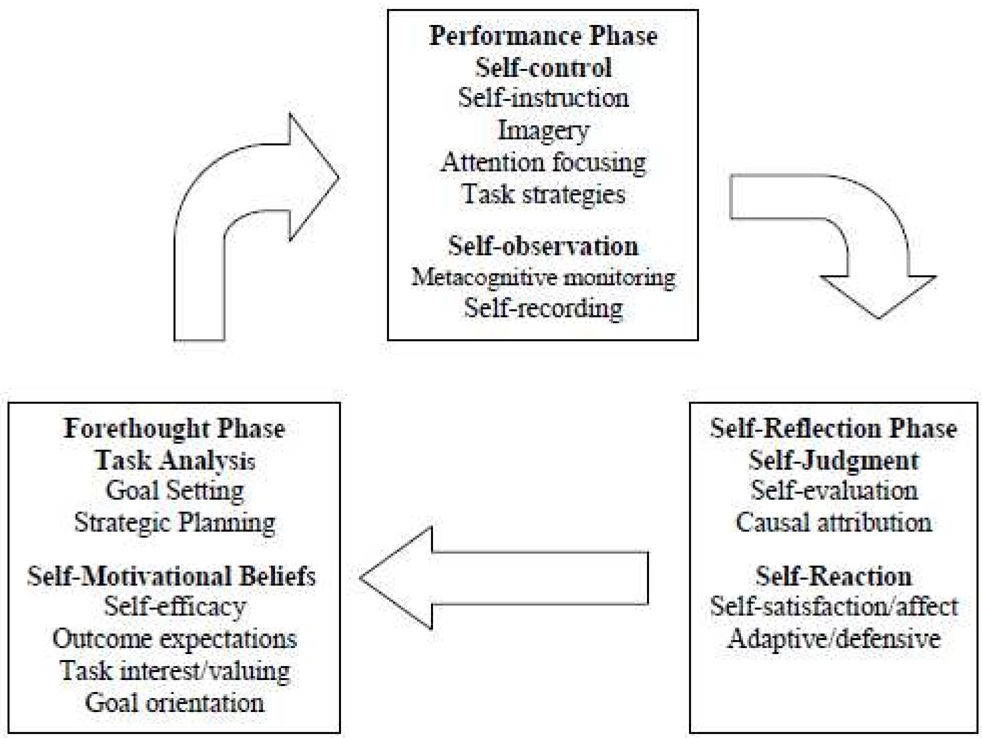
This chapter is divided into three sections. The first section will discuss self-regulated learning (SRL) models and flesh out Zimmerman's three phase model of SRL. The second section will focus on how calibration relates to Zimmerman's model, concentrating on the subprocesses of monitoring and reflection. The third section will review studies with interventions targeting monitoring and reflection. The fourth section will provide a rationale for the current intervention study and conclude with research hypotheses.

Self-Regulated Learning (SRL)

Multiple models of how students self-regulate their learning have been published, and although they propose different mechanisms for how this occurs, they commonly present learning as a cyclic process geared toward goal attainment (Pintrich, 2000a; Winne, 2001; Zimmerman, 2000). Self-regulated learners are typically seen as those who actively control their thoughts, feelings, actions and environment to aid in these pursuits. *Self-regulated learning* (SRL) has been defined as the process whereby learners activate and sustain cognitions, behaviors, and affects that are systematically oriented toward attainment of their goals (Schunk et al., 2008).

Zimmerman (2000) has put forth a prominent, well-tested model of self-regulated learning rooted in social cognitive theory that has produced many successful educational interventions (Cleary et al., 2008; Harris, Santangelo, & Graham, 2008; Schunk & Zimmerman, 1998). The personal feedback loop is an important feature of this model. Learners constantly receive both internal and external feedback about their performance during learning, which can be used to adjust their plans and strategies. The model is divided into three sequential phases that act upon one another in a cyclical manner; the phases are forethought, performance and self-reflection, and will be discussed in order next (See Figure 1).

Figure 1. Cyclical phases and subprocesses of self-regulation



From "Phases and subprocesses of self-regulation. Motivating self-regulated problem solvers", by B. J. Zimmerman and M. Campillo, 2003, p. 239. In J. E. Davidson & R. J. Sternberg (Eds.), *The nature of problem solving*, New York: Cambridge University Press. Copyright by Cambridge University Press.

Forethought phase. The learning cycle begins with the forethought phase when learners analyze a task and prepare themselves for action. Two main components of task-analysis are strategic planning and goal-setting. *Goals* can be defined as objects or aims of an action (Locke & Latham, 2002). Setting goals can facilitate performance because they serve to focus a learner's attention, increase effort and persistence, and can lead to adaptive affective reactions (Zimmerman, 2008). Goals are most beneficial when they are specific, proximal, and challenging because they facilitate strategizing about the best way to accomplish these goals (Locke & Latham, 2002). Organizing goals hierarchically by breaking down long term goals into more readily accomplished short term goals can enhance self-regulation of learning because these short term goals then serve as indicators of progress toward long term goals (Bandura & Schunk, 1981; Cleary & Zimmerman, 2004). For instance, a student who wants to earn high marks on his high school transcript may set short term goals of reviewing his notes nightly, setting aside three nights to study for each test, and calculating his grade on a regular basis to ensure that he is on track to achieve his goal to get into college.

Strategic planning refers to choosing or constructing advantageous learning methods that are appropriate for the task and environmental setting (Zimmerman & Moylan, 2009). This may involve breaking down a learning task into its component parts, developing a plan to complete the task, and selecting a strategy or strategies to enact this plan. By breaking down a task, students can gain a better understanding of what is required to accomplish it successfully, which helps them set more specific, proximal goals. This also allows students to determine the specific steps they need to take and the strategies they need to use to complete the task. Effective strategies can increase achievement by allowing students to accomplish tasks more efficiently and improve their performance. For example, students who want to improve their comprehension of their assigned reading can use concept maps to help them understand how the main ideas in the reading are related (Redford, Thiede, Wiley & Griffin, 2012). With practice, students often internalize strategies and can use them automatically. As this occurs, they may plan to redirect their efforts by using a new strategy that can help them further their mastery of a task. To continue with the reading comprehension example, with repeated practice using concept maps, students may begin to identify the connections between main ideas naturally and instead focus their energies on summarizing the text while they read to deepen their understanding (King, Biggs, & Lipsky, 1984).

The motivational beliefs that individuals hold shape the goals and plans they develop during this phase (Zimmerman, 2000). The four beliefs outlined in Zimmerman's model are self-efficacy, outcome expectations, task interest/valuing and goal orientations, which will be discussed sequentially. The most powerful of these motivational beliefs is *self-efficacy*, which refers to an individual's perceived capability to perform actions at designated levels. This belief strongly predicts the quality of a learner's self-regulation (Schunk & Swartz, 1993) and governs learner effort, persistence, achievement, motivation, strategy use, and adaptive functioning (Linnenbrink & Pintrich, 2002; Pintrich & de Groot, 2001).

Outcome expectations refer to beliefs about the ultimate ends of performance, which also have a powerful influence on one's motivation to enact a given task (Bandura, 1997). Examples include expectations of receiving monetary compensation for opening up a business or getting into a good school after studying hard for the Scholastic Aptitude Test (SAT). An individual's level of self-efficacy about accomplishing a specific task also determines how motivating the corresponding outcome expectations will be. For example, although students generally acknowledge that getting high SAT scores will improve their chances of getting into colleges, students who believe that they can improve their score by studying are far more likely to expend effort studying than students who believe that their score will not increase regardless of how much they study.

Task interest or valuing can be defined as how much one likes or dislikes a task because of its inherent qualities rather than for its instrumental qualities (Zimmerman & Moylan, 2009). Interest has been found to promote effort and persistence (Prenzel, 1992), achievement (Naceur & Schiefele, 2005) and can influence students choice of learning strategies and achievement goals, making it an important motivational belief (Ainley, Corrigan, & Richardson, 2005).

Finally *goal orientation* refers to the general pattern of beliefs that an individual holds regarding the purposes for engaging in a given task, as well as the general standards for self-evaluating learning or performance (Pintrich, 2000a; 2000b). The two main recognized goal orientations are learning and performance. A *learning orientation* is defined by goals aimed at improving mastery for the sake of improving ones abilities whereas a *performance orientation* is defined by goals aimed at enhancing or protecting one's standing in the eyes of others. Although originally conceptualized as a dichotomy, it is now widely recognized that students can hold both or neither orientation, as well as be predominately learning or performance focused (Harackiewicz, & Linnenbrink, 2005). Current research provides a further distinction among performance goals, distinguishing between approach (aimed at improving one's status), and avoidance (aimed at protecting one's status from harm). Although the goal orientation literature is complex and difficult to summarize succinctly, learning orientations generally produce adaptive academic behaviors, including seeking more challenging tasks, increased persistence in the face of failure, and increased strategy use (Harackiewicz, Barron, & Elliot, 1998; Linnenbrink & Pintrich, 2000). Furthermore, performance goals can lead to positive outcomes in the face of success, but can negatively impact motivation and performance when students experience failure (Grant & Dweck, 2003).

Performance phase. During a learning cycle, the performance phase begins when an individual initiates a task. The two self-regulatory processes that occur during this phase are *self-observation*, or attending to one's behaviors, and *self-control*, which includes a wide variety of behaviors and cognition that help students focus on the task and optimize their effort. Self-control processes can be task-specific or general. *Task-specific strategies* are systematic processes for addressing specific components of a task (Zimmerman & Moylan, 2009). Examples of task-specific strategies within the context of mathematics include reading the problem, paraphrasing, visualizing, hypothesizing, estimating, computing, and checking a problem (Montague & Bos, 1990). General strategies can include but are not limited to self-instruction, imagery and attention focusing. *Self-instruction* refers to overt or covert descriptions of how to proceed as one executes a task and has been found to improve students' learning if used properly (Schunk, 1982). *Imagery* refers to forming mental representations of information to improve understanding and memory (Pressley, 1977; Pressley & Levin, 1977). *Attention focusing* refers to methods used to improve one's concentration by screening out other covert processes or external events (Zimmerman, 2000).

Self-observation, which consists of monitoring and self-recording, is a lynch-pin of the feedback loop because students' regulatory behavior must be informed by current outcomes in order to be effective. *Monitoring* is defined as informal mental tracking of one's performance processes and outcomes (Zimmerman & Moylan, 2009). *Self-recording* occurs when students explicitly track their learning processes and outcomes with formal records (Schunk & Ertmer, 2000). Monitoring and self-recording are metacognitive processes because learners attend to processing a task at the object level as well as cues about their comprehension and learning processes at the meta-level (Thiede et al., 2009). The self-generated feedback derived from these processes can be used, along with learners' prior *metacognitive knowledge*, which is their abstract knowledge about cognitive and regulatory strategies, to control or regulate behavior (Schraw, 2001).

During the forethought phase, learners use their knowledge and beliefs to construct an interpretation of a task's demands. While engaged in the performance phase, learners generate mental (e.g., realization of progress or predictions of performance) and behavioral feedback (e.g., fatigue). Monitoring and control allow learners to use this feedback to update and possibly revise their initial interpretation of the task during self-reflection.

Self-reflection phase. After the task is completed, a student engages in reflection, during which he judges his performance and reacts to these judgments and outcomes. These judgments and reactions complete a learning cycle and influence future forethought processes (Zimmerman, 2008). In a process called *self-evaluation* the student compares her goals set during the forethought phase to the actual outcome. When learners are able to observe their gradual progress, they are likely to feel a greater sense of control and self-efficacy during their next forethought phase when approaching a similar task (Schunk, 1983). Learners who feel they are meeting their goals are expected to experience self-satisfaction and pleasant cognitive and emotional reactions, whereas those seeing themselves falling short are likely to develop unpleasant reactions (Zimmerman, 2000).

Accompanying these reactions are attributions and adaptive/defensive decisions. *Attributions* are defined as the personally constructed causal explanations about why a certain outcome or consequence occurred, and they are categorized by their locus (internal or external to the person), controllability (controllable or uncontrollable) and stability (stable or unstable over time) (Weiner, 1986). Common attributions include ability, effort/use of strategies, task difficulty and luck (Schunk & Zimmerman, 2006). Internal, controllable or unstable attributions (e.g., effort or use of strategies) typically lead to adaptive decisions such as continuing to engage in a task, even in the face of failure, and continuing to use a strategy or attempting to modify one's approach. On the other hand, external, uncontrollable or stable attributions (e.g., ability or bad luck) typically lead to defensive decisions such as withdrawing effort or lowering one's goals to prevent further unpleasant reactions (Schunk & Ertmer, 2000). Self-regulated students who are guided by self-chosen goals and strategies during the forethought phase are more likely to attribute failure to these strategies or to insufficient effort. Since ineffective strategies are typically interpreted as controllable, these students are likely to have adaptive self-reactions, including using a different strategy or applying more effort. On the other hand, students who do not spend time planning their approach during the forethought phase lack goals to which they can compare their performance. As a result, they are more likely to use the performance of their peers as a standard for evaluation, resulting in attribution of failure to uncontrollable causes such as lack of ability, which then produces withdrawal and can damage self-efficacy. Through these mechanisms, reflective processes feed forward into future forethought phases and the learning cycle begins again (Zimmerman & Labuhn, 2012).

Calibration and SRL

Monitoring is of particular interest in the current study because effective monitoring processes are implicated in enhanced self-regulatory skills and performance (Kitsantas, 2002). In particular, calibration, which is one type of monitoring judgment, is a critical component of the current intervention because the link between calibration and self-regulated learning has been well established. *Calibration* is defined as the degree to which a person's perception of performance (confidence judgments) corresponds with his or her actual performance (Hacker, Bol & Keener, 2008) and has a well-documented, positive relationship with academic performance (Bol et al., 2010; Chen, 2003; Pajares, & Graham, 1999; Pajares and Miller, 1997; Ramdass, 2008; Thiede, Anderson & Therriault, 2003). These judgments, along with other feedback generated by monitoring, provide a “bridge” between past performance and the next learning cycle (Butler & Winne, 1995). Calibration plays a critical role in self-regulation because accurate perceptions of performance can trigger appropriate control strategies, whether this involves continuing to use an effective strategy, putting forth more effort, or retooling an approach that didn’t work (Winne, 2004). On the other hand, inaccurate calibration can prevent students from effectively reevaluating their approach on a task, even if they perform poorly. Confidence judgments can occur before a task or after a task is attempted and can influence the entire learning cycle. If a learner makes a calibration judgment before a task, it is called a *predictive* judgment (akin to self-efficacy) and is likely to influence the forethought phase (but can also influence self-reflection). Confidence judgments made after a task are called a *postdictive* judgment or self-evaluative judgments and are likely to influence the reflection phase.

Dunlosky and Rawson (2011) designed two experiments to isolate the effects that calibration accuracy has on regulatory strategy use. To do this they made sure that each participant used the same regulatory strategy, and ensured that its use was dependent upon participants’ naturally occurring monitoring accuracy. During the experiments, participants studied key-term definitions and rated their understanding of each definition. After participants judged their response as correct for any given individual definition three times, the item was removed from the study pool, thereby ensuring that all participants used the same regulatory strategy. The experiments revealed a strong positive relationship between calibration, or judgment accuracy, and long-term retention. Furthermore, students who were overconfident in their monitoring, meaning they often believed they had accurately retrieved the correct definition of the key terms when they were actually incorrect, prematurely terminated their studying of these terms. Although the students subjectively believed these definitions were well learned, they had poorer learning during practice and lower levels of retention on a post-test.

Overconfidence is likely to lead to under preparation by preventing learners from making appropriate reflections and adaptive regulatory decisions (Lin, & Zabrucky, 1998). Inaccurate calibration and monitoring may also provide a learner with a false sense of strategy effectiveness, which would produce similar maladaptive effects (Hacker, Bol & Keener, 2008). Poor calibration can also produce underconfidence, where learners judge their ability level as lower than they can actually perform. Underconfident learners can misallocate study time on material they have already mastered, while not spending enough time on other academic content or other important functions like sleep. This can have negative impacts on performance, as can problems related to anxiety or motivation that may arise as a learner spends too much time preparing for something they will not feel ready for (Hacker, Bol & Keener, 2008). Thomas and McDaniel (2007) call this phenomenon “negative cascade” because inaccurate monitoring not only prevents ideal performance on a current task due to poor understanding about one’s current knowledge and skill level, it also impairs future control processes (e.g., study time allocation) later in the learning cycle.

Similarly, Thiede, Anderson and Therriault (2003) conducted an experiment examining calibration accuracy’s impact on the effectiveness of regulation and overall reading comprehension. Building upon their previous study that showed that summarizing texts after a delay, as compared to immediately or not at all, improved monitoring accuracy of these texts, undergraduate students were given six passages and received a similar manipulation to create variations in calibration (Thiede & Anderson, 2003). They were able to create these variations by assigning the participants to three groups, those that generated keywords after reading each text, those that generated keywords after reading all texts (delayed-keyword group) and those who did not generate keywords. These variations in calibration accuracy were then used to investigate how monitoring accuracy impacts regulation (i.e., selecting texts for re-study). Results showed that participants in the delayed-keyword group selected texts for restudy that they found difficult. Participants in the other groups did not meaningfully differentiate between the texts they understood well and the texts they did not, and did not appropriately allocate additional study time to the poorly understood texts. Furthermore, on a comprehension posttest, participants in the delayed-keyword group had higher levels of text comprehension than the other two groups, showing that calibration affected regulatory control decisions, which then subsequently impacted performance.

What defines self-regulated learners is that they can reflect on their initial mistakes through a process of self-evaluation and determine where the problem in their approach lies while taking appropriate steps to correct the issue (Gourgey, 2001). Strategic reflection is just as important to facilitate effective regulation as accurate monitoring and calibration, but none are sufficient by themselves (Davis, 2003). As discussed above, inaccurate calibration judgments can negatively impact strategy use and performance. By the same token, reflection is an integral component of SRL, as accurate calibration judgments do not guarantee appropriate control or regulatory strategies (Chi, Bassok, Lewis, & Reimann, 1989). For instance, Chi et al. (1989)

found that when students identified comprehension failures, appropriate reflective processes geared toward finding out why these comprehension failures occurred differentiated between high and low performers.

Schoenfeld (1985) outlines the complex interplay between monitoring and reflection. He found that students who had mastery of the course material still had poor performance when they hastily selected an inappropriate strategy before defining the task and planning out the best approach. When these students encountered difficulties, they had great difficulty generating alternative approaches or weighing which other approach might be best. Students were found to continue to try to solve the problem with an inappropriate strategy or abandon their plan, and they did not reflect on why their initial approach was not working. Meanwhile, higher performing students monitor their understanding and progress toward goals to evaluate whether to continue with their current approach or develop an alternative solution (Schoenfeld, 1985; Whimbey & Lochhead, 1986).

Davis's (2003) experiment sheds light onto how reflection impacts SRL and provides an excellent example of how monitoring and reflection interact. He examined the way different prompts facilitated reflective processes in middle school science students. Davis found that reflection was linked to success on a complex science project. He also found that students who received generic "stop and think" prompts (e.g., "Right now, we're thinking …") had more adaptive reflections, developed more coherent understandings of the content and had more accurate monitoring (were better able to identify errors) than those who received directed prompts, which provided students with hints about what to think about; (e.g., "To do a good job on this project, we need to…") an effect that contradicted his initial prediction. The author hypothesized that the generic prompts allowed learners to take control of their own reflections whereas the reflections triggered by the directed prompts were likely to create *feelings of familiarity* with the material, making students feel like they know the material even though they do not. This feeling of familiarity can then lead to overconfident calibration judgments, making students less likely to critically analyze information.

Ramdass (2008) conducted two studies and found that training students to reflect on their work can improve their academic performance and calibration accuracy. He investigated how reflection training would impact the effectiveness of a 1-hour strategy training session focused on solving fractions. Further, he found that self-reflection training, which consisted of informing students where their errors were after they solved a problem and asking them what they could do to correct them, produced weak, but consistent effects leading to better math performance as well as more accurate calibration judgments. Ramdass & Zimmerman (2008) also found that reflection training, consisting of learning how to check one's answers after solving division problems, enhanced general strategy training. The students who received both trainings showed significantly higher math performance as well as more accurate and less biased calibration judgments as compared to the group that only received strategy training.

Encouraging students to reflect has also been found to improve other areas of academic achievement. For instance, Duijnhouwer, Prins, and Stokking (2012) combined a self-reflection treatment with a feedback treatment to explore their respective impacts on writing quality in a graduate-level education course. The course required the completion of a final paper that comprised 60% of the course grade. All students received structured feedback after turning in a first draft of the paper. The experimental feedback condition received improvement strategies along with overall feedback about their paper, whereas the control just received the overall feedback. The self-reflection treatment consisted of students answering questions focused on how they intended to use the feedback to improve their paper. Those in the control reflection condition also made a reflection on the feedback, but it was focused on their perceptions of the feedback, not how they intended to use it. Students writing performance significantly improved when students were exposed to either the self-reflection treatment or the feedback treatment, but the combination of both produced no improvement. The authors hypothesize that the interaction of both treatments failed to produce effects because the provision of improvement strategies in the feedback treatment unexpectedly had a negative effect on self-efficacy. The number of strategies provided by the teacher was negatively correlated with self-efficacy of the students. The authors propose that this decrease in self-efficacy may have counteracted the beneficial effects of the improvement strategies as feedback, with the counteraction being worsened by having to reflect on these strategies. During interviews, participants reported that they already knew the strategies provided in the feedback condition, which they interpreted as their teachers underestimating their abilities conveying low confidence about their writing skills. The self-reflection treatment did not impact student self-efficacy, effort, or help seeking. Although the results indicate that only improvement strategy feedback or reflection in isolation were beneficial, the authors recommend teachers tailoring their strategic feedback to each student’s capabilities. This type of modification may create a positive interaction between strategic feedback and reflection that aligns with theoretical conceptions of these constructs.

The research reviewed above shows that monitoring, calibration, and reflection are just as important to completing a task successfully as mastery of the related content, making them ideal targets for intervention. However, this research also brings up as many questions as it answers and suggests that more research is needed to understand how these key self-regulatory processes interact with one another and the complexities of how to improve them.

Formation of metacognitive monitoring judgments. Poor calibration may be the result of students making monitoring judgments using cues that are not valid indicators of performance, such as ease of recall (instead of quality of recall) or feelings of familiarity with the material (Lin, Moore, & Zabrucky, 2001; van Loon, de Bruin, van Gog, & van Merriënboer, 2012). Cue-utilization theory hypothesizes that monitoring judgments may be based on a wide variety of cues such as how easily the task was completed, how successfully information was retrieved, how much learning has occurred, how much learning will be forgotten before the next assessment event, how well one will perform given the characteristics of the assessment (e.g., types of items, difficulty), and familiarity with course content (Koriat, 1997). These cues vary in their usefulness as predictors of accuracy, and research shows that feelings of knowing (FOK) and can lead to high confidence levels, even if the feelings of knowing are unsubstantiated (Glenberg, Wilkinson, & Epstein, 1982; Rawson & Dunlosky, 2007).

Redford, Thiede, Wiley and Griffin (2012) used the cue-utilization framework to test whether training seventh grade students to use concept maps would improve their calibration accuracy. They hypothesized that concept maps would help students focus on cues that relate to item difficulty, not cues associated with unwarranted feelings of knowing. Students who used concept maps had significantly more accurate calibration than the control groups, providing support for this theory. Further support comes from a series of experiments conducted by Maki and Serra (1992) where students read passages and made comprehension judgments about how well they understood them. They found that students used their familiarity with the domain covered in passages to make these comprehension judgments. However, data show that student domain familiarity better accounted for their predictive comprehension judgments as compared to their postdictive comprehension judgments, a finding that is consistent with the calibration literature. Postdictive judgments are typically more accurate than predictive judgments, providing additional support for the cue-utilization theory (Glenberg & Epstein, 1985; Hacker et al. 2000; Maki, 1998). Completing a task is thought to focus learners on the internal feedback they generate while completing the task. This is likely a valid cue of performance (at least more valid than domain familiarity) and is thought to help students revise their postdictive judgments to coincide better with their performance (Maki & Serra, 1992). In sum, without training or guidance, students often use ineffective methods to monitor their learning, resulting in overconfidence and underachievement (Dunlosky & Rawson, 2011).

Research showing that confidence judgments are often resistant to change and are remarkably consistent may explain why practice making confidence judgments without any direct instruction about how to improve these skills does not help improve accuracy (Bol & Hacker, 2001; Bol et al., 2005; Nietfeld et al., 2005). Confidence judgments among different tests have been found to be correlated, regardless of performance on these tests, suggesting that at least some of the factors underlying calibration may be unresponsive to context (Schraw, 1997). Furthermore, reliability of confidence judgments has been found to be higher than the reliabilities of actual performance scores themselves (Schraw, Potenze & Nebelsick-Gullet, 1993). Exploration into this phenomenon reveals that students do not use objective performance feedback to revise their confidence judgments in future learning cycles, suggesting students do not retrieve memories of their knowledge directly. Instead, it appears that students continue to base their current judgments off of prior confidence judgments (Hacker et al., 2000). Some data even shows that students who try to maintain a desirable self-image rate themselves as competent calibrators and often overestimate their comprehension level (Kroll & Ford, 1992; Lin et al., 2001). This body of research has led Hacker, Bol, and their colleagues to hypothesize that learners base their judgments on subjective beliefs about their personal attributes. These authors believe learners may be using attributions to justify any contradictions between their performance and confidence judgments, possibly to protect their self-worth. To begin exploring this hypothesis, Bol et al. (2005) tested whether there was a relationship between calibration and attributional style (primarily making internal vs. external attributions). Their results showed that attributional style was linked to calibration accuracy as well as performance, and suggest that people do not simply make objective rational calibration judgments. More specifically, they found that overconfident predictions were related to external attributions and that underconfident predictions were related to internal attributions. Bol et al. (2005) conclude that global self-concept may shape confidence judgments, which may help explain why it is difficult to improve calibration accuracy. Hacker, Bol, and Bahbabani (2008) found that attributional style did not predict calibration accuracy for high-achieving students beyond their performance on a test, but did predict calibration accuracy for low-achieving students beyond test scores. Specifically, lower-performing students' attributions of inadequate studying behavior, such as how well they studied and how well they felt they knew the material, and external social influences, such as how their teachers talked about their test or their interactions with their peers, strongly contributed to their predictive and postdictive confidence judgments.

Better insight into how confidence judgments are formed will help psychologists design better educational interventions. Dinsmore and Parkinson (2012) conducted an exploratory study on what cues students consider when they make confidence judgments about their reading comprehension. They used qualitative methods to compare the cues that poorly calibrated learners' and highly calibrated learners' use to form their judgments. Students reported that they base their confidence judgments on some combination of prior knowledge, characteristics of the text they read, characteristics of the item they answered, and guessing. Interestingly, poorly calibrated learners reported using a larger number of cues to form their judgments than did highly calibrated learners. Clearly more research is warranted to determine how different sources of judgments are formed between better and poorer calibrated students. In addition to providing students with training about how to accurately calibrate and monitor their performance, the current study further explored the sources behind students' confidence judgments and calibration. Students of all ages have problems monitoring information effectively in naturalistic settings (Bulter & Winne, 1995; Pressley, Ghatala, Woloshyn & Pirie, 1990). Likewise, students are generally overconfident and inaccurately calibrated on complex tasks found in classrooms (Bol et al., 2010; de Bruin & van Gog, 2012; Dunlosky & Lipko, 2007; Nietfeld et a., 2005), a pattern that holds true in math (Ewers and Woods, 1993; Pajares & Graham, 1999; Pajares & Kranzler, 1995; Pajares and Miller, 1997; Vermeer, Boekaerts & Seegers, 2000). De Bruin and Van Gog (2012) suggest that the key to improving monitoring accuracy lays in teaching students appropriate cues that they can use to evaluate and predict their performance. Therefore, the current intervention taught monitoring skills to allow students to better identify meaningful cues (previous performance vs. general self-concept or previous calibration judgments) and provide guided practice reflecting on what these cues mean in an attempt to facilitate regulation and achievement.

The role of feedback in fostering self-regulatory skill. Feedback, both internal and external, plays an integral role in Zimmerman's SRL model. *Feedback* can be defined as information provided by an agent regarding aspects of one's performance or understanding (Hattie & Timperley, 2007). It is inseparable from the learning process, is a key catalyst of the regulatory process and is the medium through which learners and their environments communicate (Bangert-Drowns et al., 1991). Learners use internal feedback produced by monitoring to regulate their learning and enact strategic behavior. Social feedback, such as guidance from a teacher or peer, and environmental feedback, such as from the task itself, can also be used to facilitate self-regulation. Butler and Winne (1995) conceptualize internal feedback as a bridge between past performance and the next learning cycle and posit that external feedback should be most useful at these bridge points. Effective external feedback can help learners understand, use, or develop effective domain specific and self-regulatory strategies.

Meta-analyses show that feedback has the greatest effect when it informs the learner about how to complete a task effectively or achieve a goal (Bangert-Drowns et al., 1991; Hattie & Timperley, 2007). Praise, rewards, and punishments often lack this type of information and have been shown to produce small effect sizes. The distinction between process and outcome feedback is useful to make sense of these findings. Process feedback focuses on *how* the learner is attempting to complete the task, or in other words, the methods and strategies employed during task completion. This type of feedback should help learners focus on their strategies and can prompt self-regulation by cueing learners to monitor their own processes and help them develop a better plan in future learning cycles. Process feedback makes connections between a learner's current attempt and the meta-level, allowing him to see the larger picture. This is contrasted to outcome feedback that focuses learners on their performance and the task itself. While providing information about how well a learner is performing, outcome feedback provides little guidance about how to regulate one's behavior (Labuhn, Zimmerman, & Hasselhorn, 2010; Stone, 2000). Process feedback that provides learners with an understanding of what they did well and builds upon changes they made from previous learning trails should be most useful in improving monitoring, calibration, and reflection skills in students (Hattie & Timperley, 2007). The intervention incorporated process feedback by teaching students how regulatory strategies impact the learning cycle and having students graph and track changes of their math scores and confidence judgments over the course of five sessions.

Metacognition and Self-regulated Strategy Use – Interventions That Work

Prior studies show that metacognitive skills, including monitoring, calibration, and reflection can be taught with proper instruction (Ghatala, Levin, Pressley & Goodwin, 1986; Nietfeld et al., 2006; Ramdass, 2008). The current intervention taught students how to monitor and reflect on their learning and make appropriate calibration judgments in the context of mathematical problem-solving. It emphasized the connections among these processes with the intention of deepening the participating students' metacognitive understanding of these skills, including how and when to use them and how these relate to the cognitive strategies being taught in their classroom.

The current section will briefly build the case to use metacognitive strategy training to improve self-regulation. The next sections will review interventions that have improved students' self-regulatory capacities by targeting monitoring, calibration, reflection, or some combination of these components. Importantly, teaching metacognitive strategies has been found to help all students, including those with low academic ability or a lack of relevant prior knowledge (Alexander, Carr, & Schwanenflugel, 1995; Swanson, 1990). Furthermore, metacognitive skills allow learners to better understand any domain specific cognitive skills they learn, creating more flexible learners and enhancing the probability of generalization to new domains and tasks (Schraw, 2001).

Hattie, Biggs, and Purdie (1996) conducted a meta-analysis of skill training interventions and found that strategy training was most effective when it focused on higher level metacognitive processing. This is in line with findings that metacognition plays a central role in strategy selection and use (Carr, Alexander, & Folds-Bennett, 1994). Furthermore, Hattie et al. (1996) found that skills training taught within content transfers better than skills taught in isolation. This allows the student to better understand why a strategy works and when it should be applied, all of which further the strategies generalizability. Ghatala et al. (1986) compared the effects of different strategy training components to see which ones were most effectively generalized to a novel task. They found that only the students specifically trained to use monitoring to select appropriate strategies were able to adapt the strategies they learned effectively to complete the novel task successfully. Pintrich and de Groot (2001) even found that strategy use without appropriate metacognitive and effort management skills was negatively related to performance, suggesting regulatory skills are key for students to use strategies effectively.

Review of Monitoring and Calibration Interventions

As research deepens our understanding of the regulatory cycle and how monitoring and calibration interact with other regulatory processes, psychologists are using these theoretical advancements to design and implement interventions to cultivate these processes (Hacker et al., 1998). This section will show the general progression of monitoring and calibration interventions that have gradually uncovered how these processes can be used to enhance performance and complement other academic interventions in various content areas. The next section will outline how monitoring and calibration can enhance the effectiveness of multi-component SRL interventions.

Spates and Kanfer (1977) conducted an early study testing the differential effects of self-monitoring, self-evaluation, and self-reinforcement on first graders' abilities to calculate addition problems. They compared five treatment conditions on a pre and posttest: a) Control, b) Self-monitoring, c) Criterion-setting, d) Self-monitoring plus criterion-setting, and e) Self-evaluation plus self-monitoring plus criterion-setting. For this study, the criterion setting component consisted of encouraging students to verbalize the general steps used to solve addition problems (e.g., "first I should add the two numbers on the right"). The self-monitoring training component was defined as encouraging the subjects to verbalize the specific calculations they were completing (e.g., "now I am adding these two numbers"). The self-evaluation training component was defined as encouraging students to check their work (e.g., "when you are done with each problem, look at your work and see if you did the right thing"). The results showed that across groups, the criterion-setting component was the effective element that produced significant improvements in the subjects' addition performance. Encouraging self-monitoring alone produced no effects and furthermore did not increase the effectiveness of the criterion-setting component (although this may be due to a ceiling effect), supporting the idea that monitoring one's performance in and of itself may not be enough to improve achievement.

However, some knowledge, such as amount of progress measured by items solved or words written, is more objective than whether an individual got an answer correct or is enacted a strategy appropriately. Schunk (1983) showed that monitoring task progress, defined by recording the number of pages of math problems students completed each day, improves self-perceptions and motivation. This study examined the impact of progress self-monitoring on third graders' perceptions of self-efficacy for solving subtraction problems during three, 30 minute training sessions. A group of students who had difficulty with subtraction were selected and trained on specific subtraction skills. They were then divided into three groups, those who monitored their own progress, those who had their progress monitored for them by the researcher, and those who received no monitoring intervention. Results show that the act of monitoring itself, not the monitoring agent, was important in sustaining student motivation and performance. Both the self- and external monitoring groups had significantly higher levels of efficacy, skill, and persistence as compared to the no monitoring condition but the two monitoring conditions did not differ significantly on these measures. Schunk hypothesized that the beneficial effects of self-monitoring were due to the student realizations that they were getting better, thereby enhancing their motivation to continue improving their skills. Schunk concluded that if explicit performance standards exist (e.g., objective measures of progress like answer keys), self-monitoring alone can be beneficial because it will cue learners into their progress. Further support for this contention comes from research showing that the simple act of self-recording concrete behavior can change a variety of student behaviors like their time on task and how often they talk out (Broden, Hall, & Mitts, 1971). However, Schunk argued that if explicit performance standards do not exist, (e.g., when a student is trying to decide whether she got a problem correct with no objective answer key), other regulatory processes like goal-setting and self-evaluation may be required for monitoring to have an impact.

Bol, Hacker, O'Shea, and Allen (2005) support Schunk's interpretation that merely being prompted to monitor in the absence of explicit performance standards does not improve performance. They conducted a study with college students to determine the impact that making pre and postdictive confidence judgments on a number of quizzes throughout the semester would have on their calibration accuracy and achievement. Three hundred and sixty-five undergraduates, completing both online and in vivo courses, were randomly assigned to a self-monitoring condition or to the control. Both groups took six quizzes throughout the semester. The only difference between the groups was that those in the self-monitoring condition made pre and postdictive confidence judgments about their performance on each quiz. The findings revealed that making confidence judgments had no impact on students' calibration accuracy or their final exam performance. Nietfeld et al. (2005) also found that simply having undergraduate educational psychology students make confidence judgments alone did not improve achievement or calibration accuracy across an entire semester.

However, Schunk's (1996) later study contradicts his initial interpretation and found that the simple act of monitoring one's capabilities can boost achievement. He studied the impact that goal orientation and self-monitoring had on motivation and achievement using a pre-posttest design. In this experiment on fourth grade students, he crossed a goal condition (learning v. performance) with self-monitoring over a seven session intervention designed to improve fraction completion skills. The self-monitoring condition was defined similarly to the treatment Bol et al. (2005) used, as having students make postdictive confidence judgments about their ability to complete the fraction problems covered during each of the first six sessions. At the end of the intervention, those students who self-monitored their progress had significantly higher levels of self-efficacy, fraction completion skill and persistence.

Overall these studies suggest that having students make confidence judgments about or monitor their performance is not sufficient as a stand-alone academic intervention. The studies summarized next in this section show that guidance on *how* to monitor and calibrate one's learning can enhance the effectiveness of academic interventions.

Delclos and Harrington (1991) were among the first to design an experiment to test whether combining monitoring training with general strategy skills training would have additive effects. Using pre-post test design they compared the effects of strategy skills training alone and with the addition of a monitoring training component on math achievement for fifth- and sixth-grade students. The training was conducted for 1 hour a day for 3 weeks. Strategy training involved two 1-hour sessions of training in general-problem-solving strategies such as reading a problem carefully, clarifying a problem, and thinking about similar problems. The additional monitoring training required students to answer questions before, during and after they completed each problem, forcing them to monitor their approach. Examples of questions from each of the three respective phases included "have you looked at the problem carefully and thought about how to solve it?", "did you break the problem into smaller parts?", and "how many points was your answer worth?". The authors found that adding the monitoring training to the strategy training significantly increased student's math achievement, with the greatest differences seen with more complex problems.

Desoete, Roeyers, and De Clercq (2003) conducted a study comparing five hierarchical treatment conditions administered over five 50-minute sessions in 2 weeks with small groups of third grade students. Each of the five treatment conditions incorporated an additional element so that the authors could compare the unique effects of each added component. In order from simplest to most complex, the treatments were small group instruction in unrelated content (spelling and reading; this group served as the control), practice solving math problems, practice solving motivating math problems, explicit math strategy instruction, and explicit math strategy instruction with a metacognitive calibration component. The math strategies taught focused on developing a better understanding of what numbers represent and included basic number reading, procedural calculation, differentiating different key-words in word problems and developing mental representations of numbers. The metacognitive calibration component included explicit practice with and instruction on predicting task difficulty. Findings support calibration training as an effective adjunct to math strategy instruction, even though the training only consisted of five sessions. The combined metacognitive calibration and strategy treatment produced significant results, with the highest post-test math achievement, more accurate posttest prediction scores and better scores on the follow up measure 6 weeks later as compared to the other treatment groups. No effects generalized to non-trained skills such as evaluation or number sense, which indicates that explicit training of higher order skills like this may be necessary before students can become proficient in them. More research is warranted to explore how monitoring and calibration training impact domain specific strategy interventions.

Synergy between monitoring, calibration, and reflection– comprehensive SRL interventions. The literature summarized up to this point shows that monitoring, calibration or reflection training can be effective in improving self-regulation and performance. The research in this section will review interventions that combine one or more of these three processes, which should theoretically allow students to develop metacognitive awareness of the cyclical influence each process has on the others. By capitalizing on the synergistic relationship between these three related SRL processes, researchers have consistently been able to produce positive effects on self-regulatory skills and performance. Once students understand what these processes are and how they interact, they can better use them to regulate themselves and make adaptive academic decisions.

Metacognitive instruction consistent with SRL. This section will outline interventions with a metacognitive focus, which include some combination of monitoring, calibration, or reflection and are consistent with the SRL framework. Brookhart, Andolina, Zuza, and Furman (2004) worked with classroom teachers to develop a 10 week calibration and reflection intervention to boost third graders multiplication accuracy. Their aim was to transform a rote memory task into an exercise that helped students understand how they learn at a metacognitive level. The study centered on the 5 minute multiplication fact tests that students had to complete on a weekly basis. The intervention consisted of a prediction exercise and a reflection sheet administered during the weekly test. When students received the weekly quiz, they predicted how they would do on the test and graphed the prediction. Then, after completing the test and receiving a score, they graphed their actual score and then predicted their next week's score. Finally, a reflection sheet prompted them to write if they had met their goal from the previous week, what study strategies they used and how well they worked, and what strategies they planned to use for next week's test. Although there was no control group, students' predictions became significantly more accurate with time, which has not been found to happen in the absence of intervention (Hacker, Bol, & Keener, 2008). In addition, after the intervention, students reported a lack of practice as the main reason they did not meet expectations. This is an adaptive reflection because it encourages self-regulation; students are able to control this obstacle by planning more study time for the next quiz.

Mevarach and her colleagues conducted a series of studies to build support for IMPROVE, their metacognitive method of mathematics instruction (Mevarech & Kramarski, 1997; Mevarech & Kramarski, 2003; Mevarech & Fridkin, 2006). IMPROVE is an acronym for the major components of the intervention, which are *I*ntroducing new concepts, *M*etacognitive questioning, *P*racticing, *R*eviewing and reducing difficulties, *O*btaining mastery, *V*erification and *E*nrichment. This sequence is similar to many typical curricula, but it used cooperative learning and encouraged metacognition by having students answer questions focused on three areas: (a) comprehension, or what is in the problem; (b) connection, or what are the differences between the current problem and previous problem(s); and (c) strategy use, or what is the strategy/tactic/principle appropriate for solving the problem. Their first study examined the effectiveness of the curriculum with seventh grade students over 1 year (Mevarech & Kramarski, 1997). This initial study found that classrooms using IMPROVE had significantly higher scores on standardized tests than the traditional instruction control classrooms. In addition, analysis of students mathematical explanations shows that IMPROVE students had more complex, reasoned understandings of the material.

Mevarech and Kramarski (2003) used this method with eighth grade students during a 4 week math unit on time, distance, and speed. They compared the IMPROVE curriculum to a curriculum that provided students worked examples of problems that modeled their solutions. The IMPROVE curriculum was modified for this study. While still providing prompts targeting the same three metacognitive areas covered by their 1997 study, this study also included reflection questions, specifically targeting the difficulties students had while solving problems. All students worked in cooperative groups, which generally led students to reread problems and encouraged mutual reasoning, reflective discussion, and the resolution of cognitive conflicts. However, IMPROVE students were significantly more likely to use metacognition during collective problem solving. In contrast, students who were given worked examples often just tried to apply the strategies modeled in these examples without reflecting on whether they were appropriate for the current problems they were solving. In addition, they were less likely to change their initial approach when they encountered difficulties. Analysis of videotaped problem solving revealed that IMPROVE students' metacognition was of a higher quality than that of the worked examples group, as defined by the presence of metacognitive statements throughout the entire problem solving process. Although the intervention only lasted 4 weeks, statistically significant performance differences emerged on the immediate post-test and were still present one year later. Importantly, lower achieving students benefited more from the metacognitive training than the worked examples and their gains did not come at the expense of higher achievers, who achieved similarly under both conditions. The authors conclude that metacognitive processing during mathematical problem solving creates better comprehension of the topic and that other methods of instruction, such as providing modeled solutions, place too much emphasis on application of algorithms without consideration of when and why these procedures are being used.

Mevarech and Fridkin (2006) conducted another study investigating if the IMPROVE method has similar effects on pre-college mathematics students who failed math in secondary school and required remediation before entering college. The method was applied during a 1 month period, for about 12 hours per week and was compared to a traditional remedial curriculum. Students participating in IMPROVE showed significantly higher achievement and reported more general and domain specific metacognition as well as regulation of cognition on self-report questionnaires. The IMPROVE method focuses on monitoring of the problem type and features of the problem that indicate which type of strategy to use as well as reflecting on the barriers to success as a way to determine a new plan of action and has beneficial effects on students of all ages.

SRL interventions with a metacognitive emphasis. This section will outline interventions that specifically target metacognitive self-regulatory skills and include some combination of monitoring, calibration and reflection. Cleary and Zimmerman (2004) created a comprehensive SRL intervention (the Self-Regulation Empowerment Program—SREP), based on by Zimmerman's three-stage SRL model. The first step in this individualized program is to identify weaknesses in students' self-regulatory beliefs and study strategies. Once these weaknesses are identified, a learning coach helps train the student to use strategies to overcome these weaknesses. Coaches develop students' independence by teaching them strategies using modeling and gradually releasing responsibility to students through guided practice. The program also emphasizes teaching students about the cyclical nature of self-regulation. Students receive training in goal-setting, selecting and monitoring strategy effectiveness, making strategic attributions and adjusting one's goals and strategies based on self-evaluation and reflection. To practice these skills, students work with the learning coach to set goals and develop plans to achieve these goals. After enacting their plan, students self-record performance outcomes as well as the process they used that produced this outcome. Students can then directly compare their pre-specified goals to their current progress and try to determine how to resolve any discrepancies that exist. The focus here is on corrective action and ways to regulate students' own strategy use to reach their goals next time. To empower students to feel that learning is under their control, the learning coach continually connects the strategies the student is using to success or failure through self-recording and graphing. This can help the student to see that failure is often due to use of inappropriate strategies and not fixed personality deficits. Cleary et al. (2008) found that this program significantly boosted students' science achievement, increased use of self-regulatory strategies, and enhanced confidence for learning and regulating one's behavior.

The remainder of this section will review multi-component SRL interventions that successfully improved achievement and regulatory skill in mathematical contexts. Fuchs, Fuchs, Prentice, Burch, Hamlett, Owen and Schroeter (2003) compared how an SRL component affected a curriculum designed to enhance transfer and improve mathematical achievement in third grade classrooms. The curriculum was administered over 32 sessions for 4 months. The transfer curriculum consisted of explicitly teaching students about transfer, practice identifying the difference between a superficial-problem feature and a meaningful one, and reflection on how students transferred what they were currently learning outside of their mathematics class time. Students in the SRL training group were taught self-evaluation and goal-setting. Students evaluated their progress by scoring their classwork and homework using an answer key and charting their performance. Goal-setting consisted of comparing students' previous performance with their current performance and setting new goals that exceeded their highest score. Although the transfer treatment improved performance on a near-transfer task, only the SRL plus transfer treatment produced significant positive effects on a novel task. This naturalistic study found that adding SRL training to a mathematical curriculum can improve the curriculum's effectiveness.

Perels, Gürtler, and Schmitz (2005) conducted a similar study that combined SRL training with mathematical problem solving training in a 2 x 2 control group design and found the combination of trainings worked best to improve self-regulatory competences. The training was conducted with eighth grade students over the course of six 90-minute sessions held after school. The SRL component consisted of teaching the students about a modified version of Zimmerman's three phase model of SRL, the importance of goal setting, and how to reflect upon errors and strategy use. Students also received direct instructions about handling various types of volitional problems (e.g., distraction, procrastination) and strategies to overcome these problems (e.g., stopping and reformulating negative thoughts). In addition, students constructed their own volitional strategies to solve these problems in cooperative groups. The problem solving component consisted of lessons about and practice working forwards and backwards, reading tables, figures and equations, and finding commonalties among different problems. All of the sessions used actual classroom content, reviewed content from prior sessions, and homework to reinforce the lessons. The results show that it is more difficult to train students to use self-regulatory strategies than to use general problem solving strategies. Only the students who underwent the combined trainings improved their self-regulatory skills; as compared to the control group, these students significantly increased self-reflection about errors, motivation/volitional control, and self-efficacy according to self-report questionnaires. Although these effects were significant, the effect sizes were small. All groups exposed to problem solving training improved their performance. Furthermore, the SRL training had a significant positive impact on problem solving, which suggests that these two components may work synergistically with one another. These findings clearly indicate the need for further research investigating the most effective ways to teach students self-regulatory skills.

Stoeger and Zielger (2008) targeted general SRL skills in the context of a math class. They conducted a 5-week intervention targeting fourth grade students' time management and self-reflection skills. The SRL intervention, which was conducted in nine classrooms, consisted of daily journal entries to record time management on homework assignments, as well as self-assessment of performance on daily homework and weekly quizzes. Students were explicitly taught Zimmerman's three stage SRL model and were guided through setting goals based on their previous performance. Teachers helped their students think of and record strategic methods they could use to attain these goals. Students completed worksheets and held discussions to reflect on whether the strategies they were using were helping them achieve their goals. Finally, students compared predictions they made about their performance on their homework to their actual performance, which facilitated discussion about self-evaluation and monitoring accuracy. The SRL training led to significant improvements in math performance, daily homework performance, as well as self-reported time management, self-reflection, self-efficacy, and motivation. The studies reviewed in this subsection indicate that structured reflection may be an effective way to improve monitoring during SRL interventions.

Perels, Dignath, and Schmitz (2009) conducted an intervention in sixth grade classrooms that targeted students' self-regulatory skills. Over the course of nine lessons, students learned SRL strategies embedded within one unit of their math curriculum. The strategies taught were developing a positive attitude toward mathematics and learning, motivation, goal-setting, planning, dealing with distractions, concentration, and handling mistakes. Direct instruction and small group discussion were employed to teach these strategies. Each small group designed a poster for every strategy discussed, and the class collectively taught one another about the strategies using these posters. Students practiced applying these strategies to their current math content, learned where they fell in a three stage SRL model similar to Zimmerman's (the authors renamed the three phases pre-action, action and post-action), and completed worksheets that helped them set goals and monitor their progress toward these goals. When compared to a control group that only learned math problem-solving strategies without any SRL instruction, students who participated in the SRL intervention performed significantly better on a post-test, which measured multiplication and division skill. In addition, the SRL treatment produced significantly higher levels of self-reported SRL strategy use. More specifically, students in the SRL treatment reported higher levels of monitoring, goal-setting, self-efficacy, resource-oriented and volitional strategy use, adaptive attributions and handling mistakes in an adaptive way. There were no self-reported differences on motivation or problem solving strategies. Students in the SRL treatment also showed gains in SRL knowledge as measured by a pre and post test (however, the control group did not receive this assessment so no comparison could be done). Overall these findings add support to research showing that it is possible to improve students' self-regulatory skills and math achievement with a SRL intervention embedded into math curricula.

SRL interventions emphasizing self-monitoring and reflection. The next and last set of studies reviewed in this section consist of comprehensive SRL interventions emphasizing self-monitoring training that significantly improved student self-regulatory skill and achievement. Schmitz and Perels (2011) conducted a study of a 7-week self-monitoring and self-regulated learning intervention with eighth grade math students. Students receiving the SRL intervention were given an overview of the ways that self-monitoring enhances learning and taught how to use a daily homework diary. Questions in the diary prompted students to think about all three phases of self-regulation (i.e., forethought, performance, and reflection) and included measures of planning, self-efficacy, motivation, concentration, effort, dealing with distractions, handling mistakes, self-reflection, and goal-setting/attainment. Students were also required to explain what kind of learning strategies they had used, and if none were used, they were required to explain why they didn’t use any strategies. In addition, students in the SRL treatment group were required to fill out a weekly worksheet where they outlined their long-term goals and their weekly short-term goals. Pre and post tests reveal that the SRL treatment produced small but significant effects on overall self-regulation, self-efficacy, and math performance as compared to the control condition. Furthermore, time-series analyses, shows that the SRL intervention gradually improved SRL over time. Unfortunately, no breakdown of what specific SRL skills the intervention impacted was provided in the analyses.

Huff and Nietfeld (2009) demonstrated that a brief, 14-day comprehension monitoring training intervention could improve the calibration accuracy of fifth grade students during reading comprehension activities. The authors compared two control classrooms to two process-oriented comprehension monitoring training groups, one of whom also received response-oriented monitoring accuracy training. The process-oriented comprehension monitoring training consisted of 12 daily 30-40 minute lessons intended to help students become more aware of how well they comprehend what they read, develop strategies to determine how well they understood what they read, as well as strategies to help them better understand material they realize they do not comprehend. The training consisted of explicitly teaching the students about comprehension monitoring and modeling self-monitoring and the use of “fix-up strategies” that could help students make sense of what they read (e.g., rereading, summarizing, self-questioning, adjusting reading speed, and making connections to the text). Students were provided with prompts to facilitate discussion about how to evaluate one’s own comprehension, and when fix-up strategies can be used.

To encourage monitoring, two asterisks were inserted into each independent reading passage they completed, one in the middle and one at the end. When students encountered these asterisks, they were instructed to answer three Likert scales that measured their understanding of the passage, their use of fix-up strategies, and their judgment of confidence as to how well they could explain the passage to someone else. After independent practice, each session ended with a review of the correct answers to the comprehension questions that accompanied the passages, as well as a review of the purpose of the intervention. The response-oriented monitoring accuracy training for the combined treatment group consisted of the addition of a think aloud and end of session discussion that encouraged students to think about why it is important to consider one’s level of confidence when answering comprehension questions and how confidence judgments can be used as tools to help them monitor their comprehension. Students in this treatment group were instructed to reflect on any differences between their confidence judgments and their actual performance and discussed the ramifications of being under or over confident.

Results show that both treatment groups significantly improved calibration accuracy and had higher confidence on test performance than the two comparison classes, who revealed no change in confidence or judgment accuracy. Unexpectedly, students who underwent the monitoring accuracy training also showed a significant bias towards overconfidence, a finding that the authors could not explain. There was no difference among groups for reading comprehension, which Huff and Nietfeld hypothesize could be due to the use of a standardized reading comprehension test, which may show a lack of generalization of the training, the short duration of the training, or the fact that the training didn't include regulatory strategies that would help students use their calibration judgments to improve their reading comprehension. These results indicate that even brief interventions can impact monitoring accuracy and performance and suggests that future research should be conducted exploring other brief interventions.

Nietfeld, Cao, and Osborn (2006) conducted a study on undergraduate students investigating the effects of a monitoring training administered over the course of one 16-week semester during which the class met once a week. The monitoring training consisted of completing a monitoring worksheet and receiving feedback on calibration and test performance. Each time the class met for lecture, students in the treatment condition were given a monitoring worksheet at the end of class. The worksheet asked students to rate their understanding of the day's content and to list any concepts they found difficult to understand from the day's lecture, as well as any action they would take to improve their understanding of these difficult concepts. The worksheet also contained three multiple-choice review questions about the content covered that day and prompted students to make predictive confidence judgments about their ability to solve each question. Once students completed the worksheet, the entire class went over the review questions and students were encouraged to reflect on any discrepancy between their confidence judgments and their actual performance for the review questions.

All students in the study took four tests and were asked to make predictive confidence judgments about their ability to answer each test item. Each time the students took a test, those in the treatment condition received feedback on their overall calibration and bias regarding their confidence judgments for the test items. This feedback was designed to provide an overall estimate of monitoring accuracy for students to use so that they could appropriately regulate future study time. The control group was only offered the opportunity to self-generate feedback about the discrepancies between their confidence judgments and their actual performance for test items, but received no formal feedback. As compared to the control group, the monitoring treatment produced significant improvements in calibration and performance on both multiple choice tests and an integrative end-of-term project. Although this effect took 4 weeks to establish, it was sustained throughout the remainder of the semester. These findings show that monitoring exercises improved mastery of course content, and importantly, did so while consuming minimal instructional time. The authors did not investigate if the intervention impacted any other SRL skills besides calibration, which provides an area of extension.

Zimmerman, Moylan, Hudesman, White, and Flugman, (2011) conducted a large scale study comparing the effectiveness of a semester long SRL intervention added to the pre-existing curriculum used in remedial math classes at an urban technical college. Teachers in the SRL intervention classrooms used *coping* modeling to teach math content, meaning they intentionally made mistakes to serve as a catalyst for discussion about error detection and shifting of strategies. The treatment classrooms received frequent quizzes every two to three class sessions and students had to make predictive and postdictive self-confidence judgments about the quiz items. Students in the SRL treatment group were also given the opportunity to earn back points that they lost from incorrect answers on their quizzes by correcting their mistakes on self-reflection forms. For each item they wanted to correct, the form required students to compare their pre and postdictive judgments with their outcome on the item, explain why the strategy they used was ineffective, determine a new more effective strategy, and indicate their confidence for solving another similar problem. Students were also required to solve a similar problem and outline the strategies and procedures they used to solve it. Findings show that students in the SRL treatment classes had significant higher math achievement and more accurate calibration as compared to students in the control classes. Furthermore, completing the self-reflection forms was associated with more accurate calibration and increases in math achievement. The SRL treatment also improved students' achievement on a high-stakes exam, boosting the pass rate by 25% in comparison to the control students.

Synthesis. Research on SRL shows that monitoring, calibration, and reflection training all add value to interventions aimed at improving students' regulatory capacities. Monitoring and calibration training can foster an accurate understanding of one's progress (Delclos & Harrington, 1991; Desoete et al., 2003), which can lead to appropriate reflection and forethought processes (Dunlosky & Rawson, 2011). Reflection is a well-established component of SRL interventions that has been shown to improve achievement, as well as other self-regulatory skills like calibration accuracy (Ramdass, 2008; Ramdass & Zimmerman, 2008). Interventions that simultaneously target monitoring, calibration, and reflection have great potential to increase student's self-regulatory capabilities as improvements in each of these processes can theoretically enhance the others. The current study attempted to capitalize on the synergistic relationships between these three processes by using reflective instruction about calibration monitoring that fosters deep metacognitive processing with the hopes of increasing students' regulatory abilities and achievement (Koku & Qureshi, 2004; Maki, Foley, Kajer, Thompson, & Willert, 1990). This intervention extends previous calibration research by providing calibration training that links all three phases of self-regulated learning and provides continuous process feedback to learners as they refine their regulatory skills over the course of multiple learning cycles.

Rationale for Study

More research is needed to evaluate classroom based monitoring and calibration interventions that compliment preexisting curricula as well as to understand the mechanisms underlying their effects, particularly at the K-12 level (Cleary & Zimmerman, 2004; Garavalia & Gredler, 2002; Nietfeld et al., 2005). Significant attention should be given to interventions that can improve the accuracy of student monitoring and their adaptive use of self-reflection while consuming little instructional time (Huff & Nietfeld, 2009; Nietfeld et al., 2006). Although we know that these comprehensive interventions improve broad measures of SRL, more research is needed on how these interventions impact specific SRL skills. This is especially true in the face of increasing use of microanalytic measures and other state based measurements that have proven more valid and theoretically sound than older, static, and trait-based measurements. These measures can reliably capture moment to moment metacognitive processes, thereby providing a better understanding of the dynamic nature of regulation (Zimmerman, 2008). This can help to rectify the criticism levied upon past SRL research that it focused on too large of a grain size, preventing a view of actual regulation as it occurs (Butler & Winne, 1995). For instance, few studies have focused on calibration accuracy as an outcome measure as this study intends to (Ramdass, 2008). Importantly, during the current intervention, students estimated their confidence about solving math problems correctly both before and after they attempted to solve problems. Feedback was given to students about their calibration both before and after performance in an attempt to encourage more accurate metacognitive monitoring and to provide useful information about students' calibration processes.

The current intervention combines a variety of different evidence-based methods from the literature above into one comprehensive program and in doing so, extends previous intervention research. This study attempted to generalize the methods used by Nietfeld et al. (2006) and Zimmerman et al. (2011) to middle school students. Both studies used distributed monitoring exercises, which are structured worksheets intended to facilitate consistent practice with and reflection on the participants' calibration and regulatory behavior, as their primary intervention technique. These exercises were integrated into the pre-existing curriculum where the studies were conducted and took up small amounts of instructional time. Both interventions produced more accurate calibration and higher levels of achievement in college students. The current study adapted the monitoring exercises used in these studies and integrated them into the middle school mathematics curriculum at a small private prep school in an attempt to focus learners' attention on metacognitive monitoring during the performance phase and facilitate more productive reflection processes. This approach is in line with recent meta-analytic research showing that it is important to teach skills within the context where they will be used to enhance the probability of transfer (Hattie, Biggs, & Purdie, 1996). Over the course of five training sessions, participating students answered review questions based on the material covered in their math class and made confidence judgments about their performance. They received feedback about their calibration accuracy and engaged in reflection about their calibration and performance on the review questions.

These exercises were complemented with components of other interventions shown to improve student regulatory capacity. Students were supported in developing calibration skills with feedback about their calibration accuracy over the course of the intervention (Desoete et al., 2003). Students graphed their confidence judgments along with their actual performance, which provided visual feedback about any discrepancies between the two factors and served as a platform for discussion about calibration and reflection. Graphing has been used previously in SRL research and has been shown to enhance the effectiveness of feedback (Kitsantas & Zimmerman, 2006; Labuhn et al., 2010). Students also learned regulatory strategies from all three phases of Zimmerman's self-regulated learning model to help them use any improvements in calibration accuracy they experience to make appropriate regulatory decisions (Cleary et al., 2008; Perels et al., 2009; Zimmerman, Bonner, & Kovach, 1996). Students were taught when and where to apply the strategies they learn, which should enhance the generalizability of the training (Hubner, Nuckles, & Renkl, 2010; Zohar & Peled, 2008). At the end of each session, they were prompted to reflect on their knowledge of the current math curriculum and specific strategies they could use to improve their understanding of the concepts they were learning (Zimmerman et al., 2011). The connection between student strategy use and academic outcomes was emphasized, which should help students see mistakes as learning opportunities that are within their control to fix (Cleary & Zimmerman, 2004). Combining these elements into one package has strong theoretical justification and should produce valuable results.

After the training component ends, each student was interviewed one on one to gain a better insight into the sources of information that students use to monitor their performance.

Currently, calibration research has not adequately identified the types of information students use to form confidence judgments or how students calibrate their understanding of academic content. Past research suggests that people often form confidence judgments based on preconceived beliefs about their abilities (e.g., self concept) or on task characteristics that are poor predictors of performance (e.g., feelings of familiarity) (Bol et al., 2005; Hacker, Bol, & Bahabahani, 2008; Hacker et al., 2000). Meanwhile, data show that focusing students' performance on relevant task features or monitoring processes can improve calibration accuracy (Brookhart et al., 2004; Redford et al., 2012). Firmer understandings of how the calibration process unfolds will allow psychologists to design better interventions and better comprehend why people are typically poorly calibrated (Dimmitt & McCormick, 2012; Mevarech & Fridkin, 2006). To the author's knowledge, Dinsmore and Parkinson (2012) were the first researchers to ask students open-ended questions about how they chose their confidence ratings. The current study expanded on this approach by asking students open-ended questions about their monitoring processes in real time as they solved two math problems, including how they made their confidence judgments. Their responses were interpreted in the framework of Bandura's (1986) model of reciprocal determinism that outlines how an individual's regulatory behavior is shaped by both personal characteristics (e.g., metacognition or personality traits) and environmental sources (e.g., teacher feedback or type of task).

Hypotheses. The current study tested the following hypotheses:

HO 1: Students receiving the intervention will show improved calibration as compared to those on the control group who did not yet receive the intervention.

HO 2: Students receiving the intervention will show improved mathematics performance as compared to those on the control group who did not yet receive the intervention.

HO 3: Students receiving the intervention will show increased self-regulatory strategy use as compared to those on the control group who did not yet receive the intervention.

HO 4: Students receiving the intervention will show increased metacognitive strategy use as compared to those on the control group who did not yet receive the intervention.

HO 5: Students will make confidence judgments using information from both personal and task related factors. The lack of research in this area prevents any testable hypotheses from being made.

CHAPTER III

Method

This chapter presents the methodology that the current study used to determine if the intervention improved middle school students' calibration, performance, and use of self-regulatory strategies while solving math problems. The chapter begins with a description of how participants were recruited and the experimental procedures. Next, the measures used in this study will be reviewed and a rationale for their use will be provided. Finally, the chapter will conclude with an overview of the study's design and methods of data analysis.

Design. This is a mixed methods study. An experimental design was conducted by randomly assigning participants to either a treatment group or a control group. All participants were individually interviewed at the conclusion of the study to collect qualitative data on the sources of information they used to form confidence judgments and their other monitoring processes. The control group was used to see if the experimental treatment was able to increase achievement and self-regulatory skills above and beyond another commonly used intervention.

Participants

The participants were sixth and seventh grade students recruited from a secular private school in a large Northeastern city where the principal investigator (PI) served as a school psychology intern during the 2012-2013 academic year. The school predominately enrolled upper-middle to high socioeconomic status families and provides small, resource intensive classes. The PI invited all sixth and seventh grade students enrolled in the school who were currently taking mainstream mathematics classes to participate in the study. A total of 51 students were invited, including 22 from two sixth grade math classes and 29 from two seventh grade math classes. Out of these, 30 students participated in the study, all of whom were included in data analysis. The head of school emailed a brief introduction, the PI's invitation and a parental consent form to all of the parents of sixth and seventh grade students (see Appendix A, B & C for copies of the introduction, invitation letter, and parental consent forms). The parents were encouraged to read the documents and return the forms through email with an electronic consent. They were also given the option of returning a hard copy of the form to the PI’s mailbox in the main office in a sealed envelope. To maximize participation, the PI also went to the four mainstream sixth and seventh grade math classes and gave a brief overview of the study using a prewritten recruitment script (see Appendix D for a copy of the recruitment script). The PI also handed out an invitation letter describing the nature of the study and a permission form to the students to take home to their parents that could be returned to the PI’s mailbox in a sealed envelope.

The PI visited the homerooms of each student whose parent returned the permission form to go through the assent form with them before starting the first session of the study (see Appendix E for the assent forms). This ensured that the students agreed to participate and that they understood the procedures of the study.

Measures for Quantitative Portion of the Study

Five types of outcomes were measured in this study, including (a) student math performance, (b) calibration, (c) self-regulated strategy use, (d) metacognition, and (e) prior math achievement. These outcomes were measured via math problems, single-item scales, questionnaires and interviewing techniques. Single-item scales have been used in some SRL-related studies (Cleary et. al, 2008, Zimmerman et al. 2011). Classic psychometrics viewed validity and reliability as properties of a test or instrument. These traditional conceptions have been revised because psychologists now understand that these properties can only be interpreted contextually. Since self-regulatory processes like monitoring or planning are constantly changing in relation to environmental stimuli, this study used contextually-bound single-item scales to measure confidence judgments. A series of contextually-bound questions was also administered during mathematical problem solving to collect information on the sources of information that participants use to form these judgments. This method is sensitive to small changes in mental and behavioral processes that can be missed by pre-post tests and other trait based measures (Cleary & Zimmerman, 2004). These methods have been shown to be valid and reliable (Zimmerman, 2008) and help answer the calls for a better understanding of the dynamic nature of regulatory behavior (Butler & Winne, 1995).

Math performance. Students' mathematics performance was assessed with (a) five math review questions for each of five training sessions; (b) shared items on their classroom probability unit tests (to gain a more ecologically valid measure of math performance). The five items for each session represented different content within the unit of probability and have been reviewed by a panel of experts to ensure the items were similar in difficulty level. This involved the PI compiling a set of 74 questions taken from a supplemental problems resource of the curriculum that the teachers did not use. The majority of the topics covered in both units were the same for the sixth and the seventh grade students since the curriculum used an upward spiral to deepen conceptual knowledge over the course of the two grades instead of teaching brand new material. These items covered subunits of the probability unit and were then given to four middle school math teachers, who were asked to rate their difficultly on a scale from 1 through 10 from the perspective of their average sixth/seventh grade math students, with 1 being 'not confident', 4 being 'somewhat confident', 7 being 'pretty confident' and 10 being 'very confident'. The mean of all four raters was calculated for each item. The items were then divided into five sessions, each containing five questions, and represented the progression of the topics within the broader probability unit, which included calculating experimental and theoretical probabilities, equally likely and non-equally likely outcomes, expected values and probability of two-stage outcomes. The PI selected the five items for each session by including two easy items (those with means below five), two medium items (those with means between five and seven) and one difficult item (those with means above seven) as rated by the panel of four teachers. In addition, each set of five questions had approximately the same level of mean item difficulty as rated by the panel of teachers, with the means of the five sessions ranging from 5.80 to 5.95.

For each math problem, students earned a score of 10 on an item if they solved the item correctly, a score of 5.5 if they got partial credit (awarded when only minor calculation errors led to an incorrect answer), and a score of 1 if they solved it incorrectly. This scoring system was used to align the performance scale to the predictive and postdictive confidence judgment scales described below. Students mean math performance was calculated for each session and used in the analysis of math performance.

All the teachers in the study agreed to use a shared item bank on the final unit test. This shared bank consisted of six questions that were present on both the sixth and seventh grade tests, as well as four items present only on the sixth grade test and seven items present only on the seventh grade test. These items were used to test the effectiveness of the intervention on math performance in a more natural context.

Calibration. Calibration represents the degree of difference between participants' confidence judgments compared to their actual performance. To measure calibration, researchers typically ask participants how confident they are that they can answer a question or that their answer is correct. These judgments can either be made at the local level, where estimates are made after each item, or at the global level, where a single judgment captures the expected performance on the entire task or test (e.g., asking participants how many questions they expect to get correct out of the total number of items) (Nietfeld et al., 2005). The current study used local judgments to obtain a more fine grained analysis of calibration accuracy.

Calculating calibration accuracy. Calibration accuracy can be calculated in a variety of ways, but always involves finding the difference between individuals' confidence judgments of their performance and their actual performance. The main distinction in the measurement literature is between absolute and relative accuracy. Absolute accuracy assesses the precision of a confidence judgment compared to performance on a criterion task. Relative accuracy assesses the relationship between correct and incorrect judgments, or a set of confidence judgments and performance outcomes. Said another way, absolute accuracy measures the precision of confidence judgments, whereas relative accuracy measures the consistency of these judgments (Schraw, 2009). Use of absolute accuracy indices is recommended for treatment research as these are typically more reliable and more likely to reveal individual differences than measures of relative accuracy (Hacker, Bol, & Keener, 2008).

Calibration accuracy for predictive (before performance) and postdictive (post performance) judgments was calculated by taking the absolute value of the difference between each item's predictive/postdictive calibration rating and its corresponding performance score. In this study, students earned a score of 10 on an item if they solved the item correctly, a score of 5.5 if they got partial credit (awarded when only minor calculation errors led to an incorrect answer), and a score of 0 if they solved it incorrectly. Accuracy scores represent the magnitude of calibration errors and ranged from 0 (perfect calibration) to 10 (complete lack of calibration) in this study. For example, if students report that they are totally confident (10) that they answered a question correctly and received a score of 10 on that item, the accuracy score for that item would be 0 (absolute value of 10 minus 10), indicating perfect accuracy. However, if students report that they pretty confident (7) and received a score of 10 on that item, the accuracy score for that item would be 3 (absolute value of 7 minus 10). Finally, if students reported that they were pretty confident (7) and received a score of zero on that item, the accuracy score for that item would be 7 (absolute value of 7 minus 0). For each session that students completed, these item-specific difference scores were summed and divided by the total number of items being tested, providing an overall predictive and postdictive accuracy score for each session.

Confidence judgments (prediction). To measure students' confidence judgments prior to solving each math review question, a 10-point Likert scale was used asking, "How sure are you that you will solve this problem correctly?" The scale ranges from 1 (*not sure*) to 10 (*very sure*) (see Appendix F for a sample scale). Previous research using single item scales to measure middle school students' math-specific self-efficacy (confidence judgments prior to solving problems) and self-evaluation judgments (confidence judgments after solving problems) shows high levels of internal consistency of these measures, with alphas ranging from .89-.96 (Chen & Zimmerman, 2007).

Confidence judgments (postdiction). The measure of post-performance confidence judgments read, "How sure are you that you solved this problem correctly?" The scale ranges from 1 (*not sure*) to 10 (*very sure*) (see Appendix F for a sample scale).

Self-regulated strategy use. Students' self-regulated strategy use while engaged in mathematics tasks was assessed using the 28-item Self-Regulation Strategy Inventory-Self-Report (SRSI-SR) (Cleary, 2006; Cleary & Chen, 2009) at three times throughout the intervention (see Appendix G for a copy of the SRSI-SR). This survey was developed to assess students' context-specific use of self-regulatory strategies during studying and homework completion. Although this inventory was initially developed for use with high school science students, a shortened version has been validated for use with middle school math students (Cleary & Chen, 2009). Students rate how often they used specific strategies on a 5-point Likert scale ranging from 1 (*almost never*) to 5 (*almost always*) with specific anchors for each scale unit.

Factor analytic evidence confirms the existence of three primary factors on the inventory: (a) environment and behavior management (12 items: $\alpha = .88$), (b) seeking/learning information (8 items: $\alpha = .84$), and (c) maladaptive regulatory behaviors (8 items: $\alpha = .72$) (Cleary, 2006). The environment and behavior management scale measures the frequency with which students use regulatory strategies during studying and homework completion, such as comprehension monitoring and time management. This scale includes items like "I tell myself exactly what I want to accomplish before studying" and "I quiz myself to see how much I am learning during studying". The seeking/learning information scale measures the frequency with which students seek help or use specific study tactic during studying and include items like "I ask my math teacher about the topics that will be on upcoming tests" and "I look over my homework assignments if I don't understand something". The maladaptive regulatory behavior scale measures the extent to which students engage in maladaptive regulatory behavior, such as forgetfulness and avoidance and includes items like "I avoid asking questions in class about things I don't understand" and "I wait to the last minute to start studying for upcoming tests".

Cleary and Chen (2009) combined the environment and behavior management and seeking/learning information factors into a composite measuring adaptive regulatory strategy use and found that this inventory has adequate levels of reliability with middle school students. The coefficient alpha values for the combined adaptive factor and the maladaptive regulatory scale were $\alpha = .89$ and $\alpha = .67$, respectively. These three factors have also been aggregated into a composite score representing overall self-regulatory strategy use with a high level of reliability ($\alpha = .92$) (Cleary, 2006).

Metacognitive strategy use. Students' metacognitive strategy use during math problem solving and studying was assessed using the 32 item Inventory of Metacognitive Self-Regulation (IMSR) at three times throughout the intervention (Howard, McGee, Shia, & Hong, 2000) (see Appendix H for a copy of the IMSR). The authors developed this self-report inventory to assess metacognitive awareness and regulatory skills for solving math problems. Students rate how certain sentences describe the way that they solve problems on a 5-point Likert scale ranging from 1 to 5. The specific anchors for each scale unit on the SRSI were used for this measure to create consistency and reduce confusion across surveys.

This measure was adapted from two existing public domain inventories that measure planning, monitoring, and evaluation, the Junior Metacognitive Awareness Inventory (Dennison, Murphy, Howard, & Hill, 1996) and the How I Solve Problems Questionnaire (Fortunato, Hecht, Tittle, & Alvarez, 1991). These two measures are highly correlated ($r = .68$) with one another (Sperling, Howard, Miller, & Murphy, 2002). To create this measure, Howard and colleagues (2000) first conducted an exploratory factor analysis on all of the items of both instruments using 339 students aged 10-19. After eliminating four items that focused on particular learning strategies (e.g., "I draw pictures or diagrams to help me understand while learning") the combined inventory produced a five factor structure and accounted for 42.7% of the sample variance. Second, Howard et al. tried to establish face validity of the items because they wanted the inventory to be accessible to classroom teachers. To do this, they conducted a content analysis of the remaining 32 items by writing definitions for each factor based upon the three or four items that loaded most heavily on their respective factors (.5 and above). They had a team of five raters come to consensus about which items they thought best represented the five factors. This resulted in the elimination of nine items. The remaining set of 23 items produced a five factor structure accounting for 56.3% of the sample variance.

To complete this measure, Howard et al. revised or rewrote the remaining 23 items to increase their reliability. In addition, they wrote new items to clearly demonstrate the existence of the five existing factors resulting from the process described above. They administered the revised 35-item inventory to a national sample of 829 students from grades 6-12. An exploratory factor analysis using a varimax rotation resulted in a five factor solution with eigenvalues over 1.12, which accounted for 51.6% of the variance. The instrument was found to be highly reliable. The overall alpha of the measure was .94 and the subscales ranged from .72-.87. Three items were eliminated because they weighed heavily on unexpected factors or weighed on multiple factors. Bulu and Pederson (2012) confirmed the reliability of this measure with middle school students after finding an alpha of .89.

The IMSR has been validated on middle school students and has been found to predict problem solving and content understanding (Howard, McGee, Shia, & Hong, 2001a, 2001b). Howard et al. (2001b) also reported that the measure discriminated between students with high metacognitive skills and low ability and those with high ability and low metacognitive skills and found that metacognitive self-regulation was a better predictor of problem solving success than standardized measures of ability. Bulu and Pedersen’s (2012) results further support the validity of the IMSR as an indicator of metacognitive skill. They found that students who scored highly on the IMSR did not benefit from scaffolding in a hypermedia environment like the students who had lower scores. Whereas students with lower metacognitive skills benefitted more from continuous scaffolding as opposed to faded scaffolding, students with higher metacognitive skill displayed similar problem solving skills across the conditions.

The five factors measured on the inventory are: (a) knowledge of cognition (6 items), (b) problem representation (5 items), (c) subtask monitoring (7 items), (d) evaluation (8 items), and (e) objectivity (6 items). The knowledge of cognition factor measures how much learners understand about their cognitive abilities and the ways they learn best. This factor includes items like “When it comes to learning, I know how I learn best” and “I use learning strategies without thinking.” The problem representation factor measures learners’ awareness of strategies they use to understand problems before solving them. This factor includes items like “I think to myself, do I understand what the problem is asking me?” and “I read the problem more than once.” The subtask monitoring factor measures how learners break problems down into subtasks and monitor the completion of each subtask. This factor includes items like “I use different learning strategies depending on the problem”, “I identify all the important parts of the problem”, and “I pick out the steps I need to do this problem”. The evaluation factor measures the degree to which learners are aware of checking their work throughout the entire problem solving process to evaluate if it is being done correctly. This factor includes items like “I look back at the problem to see if my answer makes sense” and “I stop and rethink a step I have already done.” The objectivity factor measures learners’ capacities to stand outside of themselves and think about their learning as it proceeds, which includes an awareness of one's learning goals and alternative choices in accomplishing a learning goal. This factor includes items like “I think of several ways to solve a problem and then choose the best one" and "I ask myself if there are certain goals I want to accomplish."

Prior math performance. The school also provided the PI with a measure of students' previous math achievement, their standard scores on the math section of the Iowa Test of Basic Skills (ITBS) for the previous academic year. This measure was used as a covariate to control for any differences in mathematical achievement prior to the treatment.

Measures of Qualitative Portion of the Study

Sources of confidence judgments. The sources that students use to form calibration judgments were measured using open-ended interview questions. After the intervention ended, the PI met with 29 participants individually to administer two additional math problems and used a set of six open-ended questions to gather information about the sources the participants used to make predictive and postdictive confidence judgments about these two questions, as well as a better understanding of their other monitoring processes. The PI selected the one easy item and one difficult item, as rated by the panel of four teachers, for this component of the study as problems of varying difficulty were likely to elicit different monitoring processes. Before participants actually solved each problem, they were prompted to make a predictive confidence judgment about their expected performance on the 10-point Likert scale described above. After the participants solved each problem, they were asked to make a postdictive confidence judgment about their performance on another 10-point Likert scale. These quantitative scales were repeated in this part of the study to explore how students' calibration accuracy impacted their answers to the interview questions. In addition they were asked the following four open-ended questions:

1. How did you make this prediction of __%?
2. Explain all of the steps you took to solve the problem. Try to be as exact as possible
3. Did you use any strategies to solve the problem?
    a. If so, which ones and why did you use these strategies?
4. How do you know whether you answered the question correctly?

Participants were then told if they got the question right and asked the remaining two open-ended questions:

5. What do you think is the main reason why you got this problem right/wrong?
6. Is there anything else you want to tell me about the problems or how you solved them?

These questions can also be found in Appendix I.

Procedure

This section begins with an overview of the procedures of the experiment and the intervention components. Next, each specific intervention component will be discussed in more detail. The PI began the intervention the week that the sixth and seventh grade mainstream math classes' began their units on probability. These units were taught in their math class by their math teacher for approximately three weeks. The treatment group received the intervention during five sessions, which the PI administered throughout these three weeks. The sessions were held during students' lunch periods, academic enrichment periods, and other non academic periods including, art, computer, and club time.

For each session they (a) completed five math questions reviewing recent material and accompanying predictive and postdictive confidence judgments focused on how well they think they will be able to/were able to complete each question (these judgments were used to measure calibration) as well as questions about specific strategies that they used to complete these problems; (b) were taught about Zimmerman's three stage self-regulated learning model and specific forethought, performance, and reflection strategies related to their math curriculum; (c) received feedback about their performance on the five math review questions; (d) received a running graph of their calibration accuracy as it progresses across the unit (the original proposal outlined that the students would self-construct their own graphs, but due to time constraints the PI constructed their calibration accuracy graphs for sessions two through five); and (e) completed a worksheet with reflective questions about the math unit.

The control group did not receive any of the active elements of the treatment during this time, which are outlined by elements (b) through (e) in the paragraph above. However, to ensure that they received equal amounts of additional instructional time, they also participated in five sessions during their probability unit. During these sessions control participants completed component (a) described in the paragraph above for data collection purposes. They then spent the remaining time receiving individualized math instruction using the computer program Math Whizz, an online math teaching program already used by the school. To ensure that all participants receive the benefits of the intervention, the control group received the intervention after the first five sessions were complete. They received five additional sessions of the full intervention incorporating elements (a) though (e) during their next math unit. In addition, both groups completed two surveys measuring self-regulatory strategy use and metacognition during math problem solving three times; before the first and third sessions and after the fifth session (see Appendix J for a visual overview of the intervention components). Finally, after the five sessions, the PI met with each participant individually to administer the open-ended interview questions to gather information about the sources the participants use to make predictive and postdictive confidence judgments and other monitoring processes.

Intervention components.

Math review questions. The PI gave students five math questions reviewing the material covered in class between sessions. While solving these questions they were asked to make predictive and postdictive confidence judgments focusing on how well they thought they would be able to/were able to complete each question. In addition, for each question they solve, they were asked an open ended question about what strategies they used to complete this question (see Appendix F for a sample of these review questions).

Instruction on Zimmerman's self-regulated learning model. The PI gave students an overview of the self-regulated learning (SRL) model in accessible language. Special emphasis was placed on the fact that learners gain information from each phase that they can use to adjust their approach in the future. Students were trained in various SRL strategies (e.g., planning, monitoring, reflection), and worked with these strategies throughout the sessions so that they could feel comfortable using them. These SRL strategies were taught in the context of the content and specific math strategies being taught in their math class to facilitate comprehension and generalizability. See Appendix K for more detail about the specific SRL training information that was covered during each session.

Feedback on review question performance. After solving the math problems, the PI told the students the correct answers to the questions they completed and informed them of how many items they got right or wrong.

Graph of calibration accuracy. The PI taught students a graphing procedure during the first session during which they constructed their own graphs of their confidence judgments along with their performance on the math questions. During the second, third, and fourth sessions, they were provided a completed calibration accuracy graph. This provided students with visual feedback about their calibration accuracy by showing the discrepancy between their predictive/postdictive confidence judgments and their actual performance with a concrete, continuous visual representation that highlights the trend in their calibration accuracy. After examining calibration accuracy graphs, the PI led a discussion about calibration, highlighting the differences between participants' confidence judgments and their actual performance.

Reflective worksheet. Students completed reflective worksheets aligned with their recent math instruction during each of the five training sessions of the intervention (see Appendix L for a copy of this worksheet). These worksheets guided students through rating their overall understanding of the material in the unit covered between sessions as well as identifying concepts or procedures they found difficult and what they could do to improve their understanding of these areas of weakness. Each worksheet asked the students (a) to consider the questions they just completed and explain what strategies or processes they did correctly; (b) to explain what strategies or processes went wrong on these questions; (c) to report how well they think they understood the material covered in their math class during the target unit; (d) to identify what concepts they found difficult to understand; and (e) to identify what they could do to improve their understanding of the concepts covered in this review.

Data Analysis: Quantitative

The major hypotheses of the study were tested using repeated measure analysis of variance (RMANOVA). A 2 (treatment: yes/no) x 2 (grade levels) multi-factorial repeated measure analysis of variance (RMANOVA) was used to test the effectiveness of the intervention on predictive and postdictive calibration accuracy. A 2 (treatment: yes/no) x 2 (grade levels) multi-factorial repeated measure analysis of variance (RMANOVA) was used to test the effectiveness of the intervention on math performance during the treatment sessions. A 2 (treatment: yes/no) x 2 (grade levels) multi-factorial ANOVA was used to test the effectiveness of the intervention on math performance on the unit test. A 2 (treatment: yes/no) x 2 (grade levels) repeated measure analysis of variance (RMANOVA) was used to test the effectiveness of the intervention on self-reported self-regulated and metacognitive strategy use during problem solving.

Power analysis.

The statistical software G-Power (version 3.1.5) was used to estimate the sample size needed to detect effects similar to previous self-regulated learning, calibration and math interventions summarized below. Prior studies that have examined the impact of self-regulatory skills training that incorporated calibration have produced significant improvements in students' calibration accuracy (mostly medium to large effect sizes, ranging from d = .39-1.34) (Brookhart, Andolina, Zuza, & Furman, 2004; Desoete, Roeyers, & De Clercq, 2003; Huff & Nietfeld, 2009; Nietfeld, Cao, & Osborn, 2006; Zimmerman et al., 2011). The current intervention shares components with a number of self-regulatory skills training programs that have increased students' mathematics achievement (small to large effect sizes, ranging from d = .31 -.75) (Desoete et al., 2003; Fuchs et al. 2003; Mevarech & Shimon, 2006; Mevarech & Kramarski, 1997; Mevarech & Kramarski, 2003; Perels, Dignath, & Schmitz, 2009; Zimmerman et al., 2011). Dignath, Buettner, and Langfeldt (2008) recently conducted a meta-analysis on SRL training programs with primary school students and found that, on average, students who participated in programs targeting mathematics performance improved achievement an entire standard deviation (mean effect size, d = 1). Dignath, Buettner and Langfeldt (2008) report that combining metacognitive self-regulatory skills and cognitive problem solving training produces a large increase in strategy use (mean effect size, *d* = .81). Comprehensive self-regulatory skills training programs similar to the current intervention have produced increases in student's use of self-regulatory strategy use and metacognition of various sizes, ranging from small to large effects (Cleary et al., 2008; Mevarech & Fridkin, 2006; Perels et al., 2009; Stoeger & Zielger, 2008).

Calculations were run with 2 groups and 5 repeated measurements under the assumption of an alpha level of both .05 and .1, power of 80%, correlations among repeated measures = .45, and a nonsphericity correction factor of .75 (in real data sets the correction factor is typically not below .75, which represents a moderate violation of sphericity assumption (Glass & Hopkins, 1996). The estimate of the correlation among repeated measures of calibration was derived from calibration research showing that calibration judgments within a domain are similar across time, with estimates ranging from .28-.55 within spatial aptitude (Jonsson & Allwood, 2003), .32- .65 within reading comprehension (Moore, Lin-Alger & Zabrucky, 2005) and .23-.62 in an educational psychology course (Hacker et al., 2000). Estimations of the correlations among repeated measures of math achievement were harder to find. Only one study was identified that reported the necessary correlational data. Norwich (1987) reported that the correlation between two math calculation tests was .55. However, it should be noted that each math calculation tasks only consisted of two moderately difficult items and that there was little academic intervention between the tests.

Sample size estimates for between and within factor ANOVA's as well as their interactions were calculated using G-Power and effect sizes (*f*-statistic) ranging from .25-.45 (medium to large effects) and are summarized in Table 1. Considering the estimates used in the calculations and the fact that using students' previous math achievement as a covariate would likely help reduce the error variance when analyzing the between-subject factors, the 51 students invited to participate in the study and 30 participants were deemed sufficient to detect the hypothesized effects of this intervention.

Table 1

Sample Size Estimates for Repeated Measures ANOVA's

	Sample Size		
Effect Size (f)	Repeated measures ANOVA, between factors	Repeated measures ANOVA, within factors	Repeated measures ANOVA, within-between interactions
.25	74 (58)	28	28
.3	52 (40)	20	20
.35	38 (30)	16	16
.4	30 (24)	12	12
.45	24 (20)	10	10

* Estimated samples outside of the parentheses are for alpha = .05, inside the parentheses are for alpha = .1

Data Analysis: Qualitative

Participants' answers to the qualitative interview questions were coded into categories and analyzed using descriptive statistics and chi-square analyses to determine whether those receiving the intervention or the control answered these questions differently and whether well and poorly calibrated participants answered these questions differently.

Participant's responses to the confidence judgment items (questions 1 and 4) were analyzed with a coding scheme adapted from Dinsmore and Parkinson (2012) based on Bandura's (1986) model of reciprocal determinism. Bandura's model proposes that an individual's regulatory behavior (i.e., rating one's confidence judgment) is shaped by both personal characteristics (e.g., metacognition or personality traits) and environmental sources (e.g., teacher feedback or type of task). As such, participants' responses were coded into seven a priori categories reflective of these personal and environmental influences. Categories of personal characteristics included: (a) prior knowledge (e.g., "I am doing well in class"; "I remember covering something like this in class", "I know how to solve problems like this"), (b) self-concept (e.g., "I am really good at math"), and (c) metacognition (e.g., "I checked my answer and knew I was right", "I thought back and realized I knew how to solve problems like these"). Categories of environmental sources included: (a) characteristics of the item (e.g., "the question was really difficult") and (b) social reasons (e.g., "my teacher says I am good at these types of questions"). Two other additional categories included: (a) guessing and (b) other considerations.

Participants responses to question 5 were analyzed with a coding scheme adapted from Weiner (1986) that consists of the following five a priori categories: (a) aptitude (fixed ability representing an internal, stable attribution); (b) skills/knowledge/strategy use (these can be learned over time representing an internal, unstable attribution); (c) effort (representing an internal, unstable characteristic); (d) task difficulty (representing an external, stable attribution); and (e) chance (representing an external, unstable attribution).

Participants' responses to question 2 were analyzed by counting the number of steps participants took to solve each problem. If the participant got the question right, the steps were checked for conceptual accuracy. If the participant got the question wrong, the steps were checked for patterns of errors and conceptual misunderstandings.

Participants' responses to questions 3 and 6 were coded after data were collected. Their answers to question number 3 were then grouped into the following thematic categories: (a) efficiency (e.g. the quickest or easiest way to get an answer); (b) accuracy (e.g. a way to get the correct answer); (c) prior knowledge (e.g. indicated that they have used the strategy for a similar problem before); (d) social (e.g. because they were taught to use the strategy); and (e) "I don't know" or no strategy.

Most students did not have any additional information that they wanted to report about the problems or how they solved them for question 6, so responses to this question were not analyzed.

These open-ended questions were analyzed by the PI and another qualified rater. The qualified rater was an educational psychology doctoral student who was trained on the coding schemes and provided examples of each of the categories of qualitative responses. The two raters worked together to code all of the answers according to the coding schemes outlined above. Any disagreements were discussed until consensus was reached. These data were used to determine whether those receiving intervention or the control answered these questions differently and whether well and poorly calibrated participants answered these questions differently.

CHAPTER IV

Results

Preliminary Analyses

To ensure that random assignment was effective in equalizing the treatment and control groups, students' prior mathematical achievement (scores on the math portion of the Iowa Test of Basic Skills (ITBS)), was compared using a univariate analysis of variance (ANOVA) for both sixth and seventh grade students. The grades were compared separately because the test is designed with different mean scores across grade levels. No significant differences on overall math achievement were found between the treatment group and the control group in either grade (sixth, $F(1,13) = .169$, $p = .689$, $\eta^2 = .015$; seventh, $F(1,16) = 1.176$, $p = .681$, $\eta^2 = .012$). As a further check of the effectiveness of random assignment, a 2 (condition: experimental vs. control) x 2 (grade levels) multi-factorial ANOVA was conducted examining participants math achievement on the first set of math problems they solved before their treatment began, as well as their first set of predictive and postdictive accuracy scores. These measures were taken before the treatment group received any intervention, so the measures should be similar across groups. Regardless of treatment assignment or grade, there were no significant differences among participants on their performance on the first set of math problems, their predictive accuracy, or their postdictive accuracy (performance, treatment, $F(1,29) = 1.808$, $p = .190$, $\eta^2 = .065$; grade, $F(1,29) = .486$, $p = .492$, $\eta^2 = .018$; predictive accuracy, treatment, $F(1,29) = .365$, $p = .551$, $\eta^2 = .014$; grade, $F(1,29) = .389$, $p = .538$, $\eta^2 = .015$; postdictive accuracy, treatment, $F(1,29) = .573$, $p = .456$, $\eta^2 = .022$; grade, $F(1,29) = .025$, $p = .876$, $\eta^2 = .001$. Because these tests found that there were no initial differences in math achievement, participants ITBS scores were not used in subsequent analyses. Descriptive statistics are presented for the major dependent variables in

Table 2. Participants were 14 sixth grade and 17 seventh grade students with approximately equal numbers of boys and girls in each grade. Further demographic information can be found in Table 3. Correlations between math performance and the self-regulated learning (SRSI-SR) and metacognition (IMSR) questionnaires are presented in Table 4. As expected, the correlations between self-regulated learning strategy use and metacognitive strategy use were large and positive. However, the medium size negative correlation between self-regulated strategy use and math performance found in the control group was unexpected and contrary to other research on this instrument (Cleary, 2006; Cleary & Chen, 2009). Correlations between the predictive and postdictive accuracy and performance are presented in Tables 5 and 6 respectively. The negative correlations in Tables 5 and 6 were expected because calibration accuracy is on a reverse scale where numbers closer to zero indicate higher levels of accuracy. As expected, there were large correlations between predictive and postdictive accuracy and performance. Overall, these correlations were largest when comparing calibration accuracy and performance within one session. However, there are other large and medium correlations among calibration accuracy and performance across sessions as well.

Table 2

Descriptive Statistics for the Major Dependent Variables

Dependent Variables	Control			Treatment		
	N	*M*	*SD*	*N*	*M*	*SD*
Iowa Test of Basic Skills	15	243.36	23.12	14	242.87	25.30
Calibration Session 1 – Predictive[a]	15	4.29	1.04	15	4.09	.93
Calibration Session 2 – Predictive[a]	15	4.42	1.76	15	4.13	1.29
Calibration Session 3 – Predictive[a]	15	4.19	1.35	15	3.03	1.16
Calibration Session 4 – Predictive[a]	15	4.68	1.35	15	4.41	.67
Calibration Session 5 – Predictive[a]	15	4.72	1.04	15	3.88	1.46
Calibration Session 1 – Postdictive[a]	15	4.08	1.27	15	3.87	.82
Calibration Session 2 – Postdictive[a]	15	3.53	1.80	15	3.48	1.61
Calibration Session 3 – Postdictive[a]	15	3.73	1.74	15	2.98	1.08
Calibration Session 4 – Postdictive[a]	15	3.72	1.74	15	2.97	.93
Calibration Session 5 – Postdictive[a]	15	4.48	1.97	15	2.76	1.50
Mean Math Performance Session 1[b]	15	3.52	1.53	15	4.36	1.91
Mean Math Performance Session 2 [b]	15	2.62	1.37	15	3.16	2.06
Mean Math Performance Session 3 [b]	15	4.54	2.16	15	5.50	1.56
Mean Math Performance Session 4 [b]	15	2.92	2.2	15	4.00	1.26
Mean Math Performance Session 5 [b]	15	2.80	1.44	15	4.60	1.99
Mean SRSI 1 Score[c]	13	3.33	.75	14	3.20	.71
Mean SRSI 2 Score[c]	13	3.32	.53	13	3.53	.63
Mean SRSI 3 Score[c]	15	3.30	.61	15	3.45	.66
Mean IMSR 1 Score[d]	15	3.38	.52	15	3.49	.48
Mean IMSR 1 Score[d]	15	3.29	.54	15	3.61	.64
Mean IMSR 1 Score[d]	15	3.29	.73	15	3.58	.66

[a] Predictive and postdictive accuracy scores ranged from 0-10, with 0 being the most accurate
[b] Math performance scores ranged from 0-5
[c] Self-Regulation Strategy Inventory-Self-Report (SRSI) scores ranged from 1-5
[d] Inventory of Metacognitive Self-Regulation (IMSR) scores ranged from 1-5

Table 3
Participant Demographics

Grade	Teacher	Treatment	*N*
Sixth	A	Experimental	4
		Control	3
	B	Experimental	3
		Control	4
Seventh	A	Experimental	4
		Control	3
	C	Experimental	4
		Control	5

Table 4
Correlations Among Math Performance and the Self-Regulated Learning and Metacognition Questionnaires Collapsed Across the Five Sessions

Dependent Variables	1.	2.	3.
1. Total Math Performance	—	-.311	.086
2. SRSI Total	.474	—	.778**
3. IMSR Total	.164	.920**	—

Correlations above the diagonal are from the treatment group; correlations below the diagonal are from the control group
* Correlation is significant at the 0.05 level (2-tailed)
** Correlation is significant at the 0.01 level (2-tailed)

Table 5

Correlations Among Predictive Accuracy and Math Performance Across the Five Sessions

Dependent Variables	1	2	3	4	5	6	7	8	9	10
1. Math Performance Session 1	—	.670**	.584*	.563*	.183	-.088	.182	-.210	-.027	.109
2. Math Performance Session 2	. 453	—	.612*	.490	.198	-.462	.232	-.421	.117	.199
3. Math Performance Session 3	.294	.816**	—	.413	.299	-.080	.192	-.445	-.025	.011
4. Math Performance Session 4	.179	.046	-.023	—	.139	.139	.109	-.213	-.173	.266
5. Math Performance Session 5	.603*	.234	.056	.474	—	.035	.176	.205	.329	-.275
6. Predictive Accuracy Session 1	-.336	.058	-.060	.051	-.098	—	.367	-.206	-.401	-.087
7. Predictive Accuracy Session 2	-.087	-.204	-.270	.434	.253	-.049	—	-.233	.092	-.025
8. Predictive Accuracy Session 3	-.016	-.666**	-.691**	.505	.306	-.303	.570*	—	.439	.034
9. Predictive Accuracy Session 4	.007	.476	.597*	-.380	-.228	-.022	-.278	-.748**	—	.251
10. Predictive Accuracy Session 5	-.385	-.038	.137	.324	-.351	.129	.371	-.019	.156	—

Correlations above the diagonal are from the treatment group; correlations below the diagonal are from the control group

* Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

Table 6

Correlations Among Postdictive Accuracy and Math Performance Across the Five Sessions

Dependent Variables	1	2	3	4	5	6	7	8	9	10
1. Math Performance Session 1	—	.670**	.584*	.563*	.183	.011	.107	-.096	-.098	.519*
2. Math Performance Session 2	. 453	—	.612*	.490	.198	-.318	-.118	-.436	.226	.209
3. Math Performance Session 3	.294	.816**	—	.413	.299	.080	.226	-.551*	.132	.496
4. Math Performance Session 4	.179	.046	-.023	—	.139	.166	.433	-.142	-.310	.205
5. Math Performance Session 5	.603*	.234	.056	.474	—	-.118	.194	-.150	.153	-.238
6. Postdictive Accuracy Session 1	-.377	.098	-.022	-.022	-.453	—	.361	.280	.021	.089
7. Postdictive Accuracy Session 2	.002	-.286	-.412	.767**	.372	-.142	—	-.081	.088	.068
8. Postdictive Accuracy Session 3	-.066	-.470	-.719**	.281	.146	-.035	.573*	—	.137	.127
9. Postdictive Accuracy Session 4	-.354	-.227	.070	-.090	-.505	-.179	-.050	-.056	—	.158
10. Postdictive Accuracy Session 5	-.048	.029	.358	.273	-.316	-.026	.180	-.223	.611*	—

Correlations above the diagonal are from the treatment group; correlations below the diagonal are from the control group

* Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

Primary Analyses

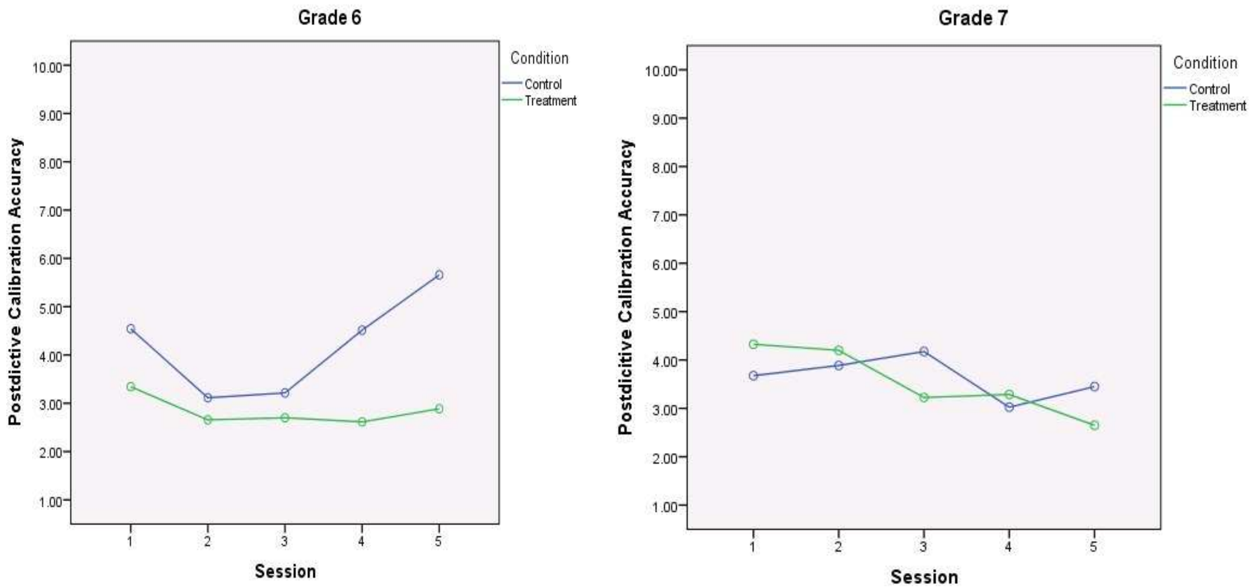
Predictive calibration accuracy. To test hypothesis one and to determine the effect of the intervention on predictive calibration accuracy, a 2 (condition: experimental vs. control) x 2 (grade levels) multi-factorial repeated measure analysis of variance (RMANOVA) was conducted on the measures of predictive accuracy collected during each of the five sessions throughout the study. Main effects of treatment were found for predictive accuracy across the five sessions, $F(1,26) = 8.314$, $p = .008$. According to Cohen (1988), these results can be categorized as a large effect size (small effect = .01; moderate effect = .06; large effect = .14).

Postdictive calibration accuracy. To test hypothesis one and to determine the effect of the intervention on postdictive calibration accuracy, a 2 (condition: experimental vs. control) x 2 (grade levels) multi-factorial repeated measure analysis of variance (RMANOVA) was conducted on the measures of postdictive accuracy collected during each of the five sessions throughout the study. Main effects of treatment were found for postdictive accuracy across the five sessions, $F(1,26) = 7.291$, $p = .012$, $\eta^2 = .219$. These results can be categorized as a large effect size (Cohen, 1988).

There was a significant interaction between treatment and grade for postdictive accuracy, $F(1,26) = 5.361$, $p = .029$, $\eta^2 = .171$. This finding shows that the difference in calibration accuracy between sixth graders in the treatment and wait list control groups was much larger than the difference in calibration accuracy between seventh graders in the treatment and wait list control groups. Further analysis shows that the sixth graders receiving treatment improved their calibration accuracy after the first session and maintained a similar degree of accuracy after that while the seventh graders receiving treatment continued to improve their calibration accuracy across all five sessions (see Figure 2).

Figure 2

Interaction Between Treatment and Grade for Postdictive Accuracy



Simple contrasts, or univariate F-tests, were conducted that compared participants calibration in the first session to their calibration for sessions two through five to gain more information about trends across time. These contrasts revealed that postdictive accuracy was significantly higher during session three and session four than during session one (session three compared to session one, $F(1,26) = 4.69$, $p = .04$, $\eta^2 = .153$; session four compared to session one, $F(1,26) = 4.537$, $p = .043$, $\eta^2 = .149$. In addition, contrasts also revealed a significant interaction between session and treatment for postdictive accuracy when comparing session five with session one. The treatment group showed significantly better calibration accuracy between the first and last session, whereas the control group did not (session five compared to session one, $F(1,26) = 4.244$, $p = .05$, $\eta^2 = .140$).

Mathematics performance. To test hypothesis two and to determine the effect of the intervention on mathematics performance a 2 (condition: experimental vs. control) x 2 (grade levels) multi-factorial repeated measure analysis of variance (RMANOVA) was conducted on the measures of math performance collected during each of the five sessions throughout the study. Main effects of treatment on math performance were found across the five sessions. This test revealed that the treatment group had significantly higher math performance on these problems as compared to the control group, $F(1,26) = 5.750$, $p = .024$, $\eta^2 = .181$. These results can be categorized as a large effect size (Cohen, 1988). It must be noted that the fifth session in this analysis violated Levene's test of equality of error variances, meaning that the variances of the subjects in the different conditions (treatment x grade) were not equal for this session. This suggests that the results of this analysis should be interpreted with caution.

A significant within subject effect of time was found across the five sessions, $F(4,104) = 8.472$, $p = .000$, $\eta^2 = .246$. Simple contrasts were conducted that compared participants math performance in the first session to their math performance for sessions two through five to gain more information about trends across time. As compared to participants performance during session one, these contrasts revealed that their performance was significantly lower during session two and significantly higher during session three (session two vs. session one, $F(1,26) = 12.452$, $p = .002$, $\eta^2 = .324$; session three vs. session one, $F(1,26) = 8.731$, $p = .007$, $\eta^2 = .251$).

To test the effectiveness of the intervention on math performance during a naturally occurring classroom assessment, a 2 (treatment: yes/no) x 2 (grade levels) multi-factorial ANOVA was conducted on the shared sixth and seventh grade items on the probability unit exam. This test revealed no significant effects of treatment or grade (treatment, $F(1,26) = .011$, $p = .918$, $\eta^2 = .000$; grade, $F(1,26) = 2.516$, $p = .125$, $\eta^2 = .088$). In addition, a univariate ANOVA was used to examine specific grade level effects of the intervention on math performance for both sixth and seventh grade participants using their performance on grade specific items on the probability unit exam. These tests revealed no significant effects of treatment for sixth grade, $F(1,12) = .255$, $p = .623$, $\eta^2 = .021$ or seventh grade, $F(1,14) = .164$, $p = .691$, $\eta^2 = .012$.

Self-regulated learning strategy use. To test hypothesis three and determine the effectiveness of the intervention on self-regulated learning strategy use during problem solving, a 2 (treatment: yes/no) x 2 (grade levels) multi-factorial repeated measures analysis of variance (RMANOVA) was conducted on the survey measurements collected three times throughout the study. This test was conducted on each of the three subscales on the SRSI-SR as well as the total score of the SRSI-SR.

Participants in seventh grade reported higher levels of environment and behavior management strategies, $F(1,22) = 6.024$, $p = .022$, $\eta^2 =.215$ and seeking/learning information strategies, $F(1,25) = 8.454$, $p = .008$, $\eta^2 = .253$ than participants in the sixth grade. In addition, a significant within subject interaction was found for time and treatment for the environment and behavior management scale, $F(2,44) = 4.210$, $p = .021$, $\eta^2 = .161$. Simple contrasts were conducted that compared participants self-reported self-regulatory strategy use before the first session (measurement 1) to their strategy use before session three (measurement 2) and after session five (measurement 3) to gain more information about trends across time. These contrasts revealed that participants receiving treatment reported significantly higher levels of environmental and behavior management strategies before session three (measurement 2) and after session five (measurement 3) as compared to before session one (measurement 1) (measurement two vs. measurement one, $F(1,22) = 5.565$, $p = .028$, $\eta^2 = .202$; measurement three vs. measurement one, $F(1,22) = 5.582$, $p = .027$, $\eta^2 = .202$).

Analysis of the overall scale produced similar results. Participants in seventh grade reported higher levels of self-regulated strategy use compared to participants in the sixth grade $F(1,20) = 6.11$, $p = .023$, $\eta^2 = .234$. A significant within subject interaction was found across session and treatment for the composite of the SRSI scale, $F(1.529,30.583) = 3.649$ $p = .049$, $\eta^2 = .154$. The greenhouse-geisser correction was used for this test as it violated the assumption of sphericity, $\chi^2\ (2) = 6.993$, $p = .03$. Contrasts revealed that participants reported significantly higher levels of overall self-regulated strategy use after session five (measurement 3) than before session one (measurement 1), $F(1,20) = 5.294$, $p = .032$, $\eta^2 = .209$.

Metacognitive strategy use. To test hypothesis four and determine the effectiveness of the intervention on metacognitive strategy use during problem solving, a 2 (treatment: yes/no) x 2 (grade levels) multi-factorial repeated measures analysis of variance (RMANOVA) was conducted on the survey measurements collected three times throughout the study. This test was conducted on each of the five subscales on the IMSR as well as the total score of the IMSR. This test revealed no significant effects of treatment or grade on the total scale (treatment, $F(1,26) = 1.656$, $p = .210$, $\eta^2 = .06$; grade, $F(1,26) = 3.859$, $p = .06$, $\eta^2 = .129$) or any of the five subscales.

Summary of Findings Related to Study Hypotheses

Table 7 summarizes this study's hypotheses and indicates which hypotheses were supported by the research findings. The results provided support for the hypotheses one and two, but not for hypotheses three and four.

Table 7

Summary of Research Hypotheses and Results

Hypotheses		Results
HO_1:	Students receiving the intervention will show improved calibration as compared to those on the control group who did not yet receive the intervention	Supported
HO_2:	Students receiving the intervention will show improved mathematics performance as compared to those on the control group who did not yet receive the intervention	Supported
HO_3:	Students receiving the intervention will show increased self-regulatory strategy use as compared to those on the control group who did not yet receive the intervention	Not supported
HO_4:	Students receiving the intervention will show increased metacognitive strategy use as compared to those on the control group who did not yet receive the intervention	Not supported
HO_5:	Students will make confidence judgments using information from both personal and task related factors. The lack of research in this area prevents any testable hypotheses from being made.	No testable hypotheses were made

Qualitative Analysis

Interviews were used to collect information about participants' monitoring processes during problem solving to explore (a) how they made their specific predictions and postdictions, (b) the steps they took to solve the problems, (c) what strategies they used to solve the problems, and (d) their perception of why they got the problems right/wrong. One participant did not complete an interview, so qualitative data were only collected for 29 participants. Their answers to these questions were coded into categories and then analyzed to determine what types of answers participants in the treatment and control groups provided.

Descriptive analysis. For each question, the percentage of participants that provided answers that fell into each category is tabulated below and interesting findings are highlighted.

The first question asked participants how they made their predictive judgments. Participants' answers to this question for both the easy and difficult problem in the interview are presented in Table 8 below. On the easy problem, the control group participants reported using more guessing or "gut" feelings to make their predictive judgments then the treatment group. On the difficult problem, the treatment group predominantly reported using metacognition and self-concept to form their predictive judgments, whereas the control group was more likely to report using their prior knowledge, the characteristics of the item, or guessing.

Table 8
Qualitative Responses of Participants in the Treatment and Control Groups – Question One – Sources of Predictive Judgments

Codes for predictive judgment	Treatment Group – Easy Problem	Control Group – Easy Problem	Treatment Group – Difficult Problem	Control Group – Difficult Problem
Prior knowledge	35.7%	33.3%	0%	13.3%
Self-concept	21.4%	13.3%	28.6%	0%
Metacognition	28.6%	26.7%	57.1%	33.3%
Item characteristics	7.1%	0%	14.3%	40%
Social	0%	0%	0%	0%
Guessing	0%	6.7%	0%	13.3%
Other – "gut"	7.1%	20%	0%	0%

The second question asked participants to explain all of the steps they took to solve the problems. The number of steps participants reported taking for both the easy and difficult problem in the interview are presented in Table 9 below. Regardless of treatment group, all participants used a similar number of steps to solve the easy problem and the difficult problem respectively. For the difficult problem the majority of participants did not provide any work or any answer. The majority of participants used two steps to solve the easy problem.

Table 9

Qualitative Responses of Participants in the Treatment and Control Groups – Question Two – Number of Steps Taken to Solve Each Problem

Number of steps	Treatment Group – Easy Problem	Control Group – Easy Problem	Treatment Group – Difficult Problem	Control Group – Difficult Problem
0	0%	6.7%	50%	53.3%
1	21.4%	20%	21.4%	20.0%
2	64.3%	66.7%	14.3%	26.7%
3	14.3%	6.7%	7.1%	0%
4	0%	0%	7.1%	0%

* If participants provided no work and no answer the number of steps participants took were counted as zero

The third question asked participants about the strategies they used to solve the problems and why they used them. Participants' answers about why they used the strategies they did for both the easy and difficult problem in the interview are presented in Table 8 below. On the easy problem, participants were mostly likely to report that they used strategies that helped them accurately or efficiently answer the problem. On the difficult question, the vast majority of participants did not use a strategy to solve the problem or were unable to report why they used the strategy they did.

Table 10

Qualitative Responses of Participants in the Treatment and Control Groups – Question Three – Why Participants Used the Strategies They Did to Solve Each Problem

Codes for strategies	Treatment Group – Easy Problem	Control Group – Easy Problem	Treatment Group – Difficult Problem	Control Group – Difficult Problem
Efficiency	50%	53.3%	14.3%	6.7%
Accuracy	28.6%	20%	7.1%	6.7%
Prior knowledge	7.1%	6.7%	0%	0%
Social	0%	13.3%	0%	0%
I don't know or no strategy	14.3%	6.7%	78.6%	86.7%

The fourth question asked participants how they made their postdictive judgments. Participants' answers to this question for both the easy and difficult problem in the interview are presented in Table 11 below. On the easy problem, the control group participants reported using more guessing or "gut" feelings to make their postdictive judgments, whereas the treatment group reported using more self-concept. On the difficult problem, the control group participants reported using more "gut" feelings to make their postdictive judgments, whereas the treatment group reported using more metacognition or prior knowledge.

Table 11
Qualitative Responses of Participants in the Treatment and Control Groups – Question Four – Sources of Postdictive Judgments

Codes for predictive judgment	Treatment Group – Easy Problem	Control Group – Easy Problem	Treatment Group – Difficult Problem	Control Group – Difficult Problem
Prior Knowledge	28.6%	26.7%	14.3%	0%
Self-concept	21.4%	0%	0%	6.7%
Metacognition	42.9%	53.3%	57.1%	33.3%
Item Characteristics	0%	0%	0%	0%
Social	0%	0%	0%	0%
Guessing	7.1%	13.3%	21.4%	20%
Other – "Gut"	0%	6.7%	7.1%	40%

After the participants were told whether they got the questions right or wrong, they were asked the fifth question about what they thought was the main reason they got the problem right or wrong. Participants' answers to this question for both the easy and difficult problem in the interview are presented in Table 12 below. For both the easy and difficult problems, the majority of participants reported that their strategy use or knowledge was the main reason why they got the problem right or wrong.

Table 12

Qualitative Responses of Participants in the Treatment and Control Groups – Question Five – Participants Perceptions of the Main Reason They Answered the Questions Correctly/Incorrectly

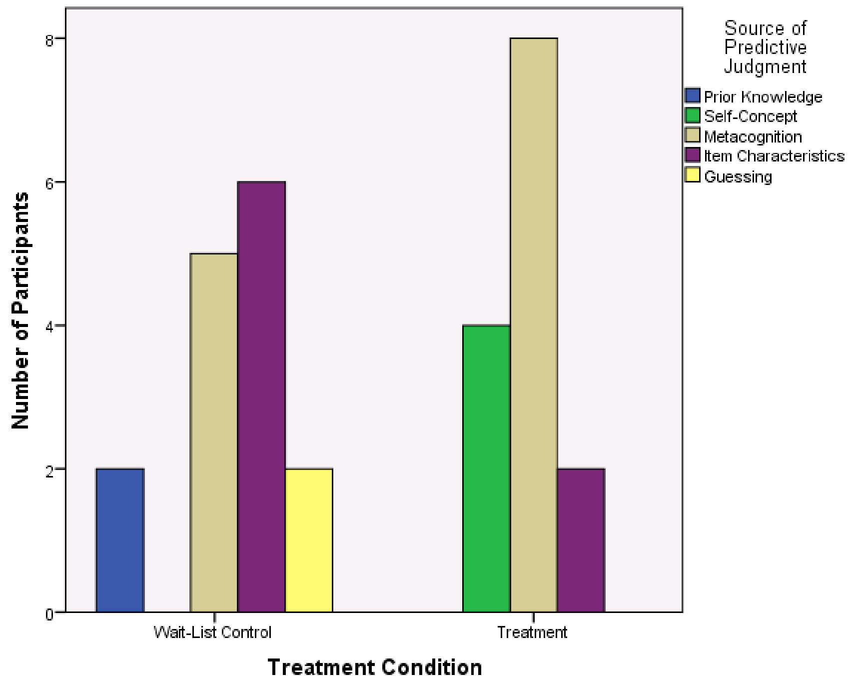
Codes for predictive judgment	Treatment Group – Easy Problem	Control Group – Easy Problem	Treatment Group – Difficult Problem	Control Group – Difficult Problem
Aptitude	7.1%	6.7%	0%	0%
Strategy use/knowledge	85.7%	73.3%	50%	73.3%
Effort	0%	6.7%	7.1%	0%
Task difficulty	7.1%	0%	7.1%	0%
Chance	0%	0%	21.4%	13.3%
I don't know	0%	13.3%	14.3%	13.3%

Only three participants provided an answer to the sixth question, which asked if there was anything else they wanted to report about the problems or how they solved them. Because so few students answered this question, the results were not analyzed.

Treatment versus control group responses. Chi-square analyses were used to determine whether those in the intervention or the control groups answered the qualitative questions outlined above differently. Chi-square analysis comparing the treatment group to the control group revealed significant differences between the sources participants in both groups used to make their predictive judgments on the more difficult interview problem, χ^2 (4, N = 29) = 10.671, p = .031. Specifically, participants in the treatment group were more likely to use metacognition and self-concept to make their predictive judgment whereas participants in the control group were more likely to use item characteristics, prior knowledge and guessing. See Figure 3 below for a graphic representation of these findings. Chi-square analyses comparing treatment and control group participants revealed no other significant differences in the sources participants used to form postdictive calibration judgments, the number of steps they took to solve the problems, what strategies they reported using and their perceptions of the main reasons why they got the question right/wrong.

Figure 3

Treatment Condition and Source of Predictive Calibration Judgments – Chi-Square Analysis

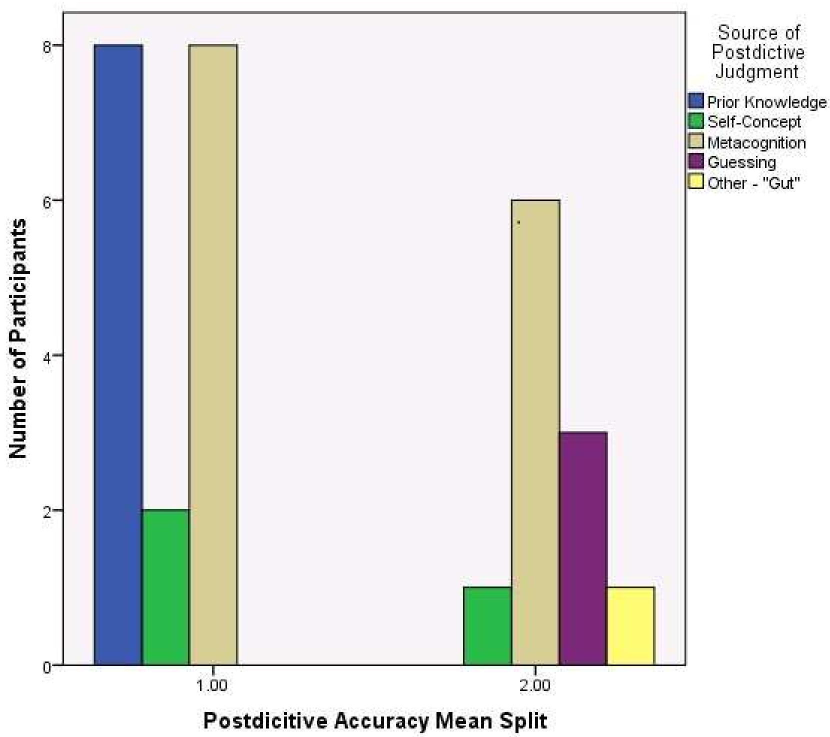


More accurate versus less accurate calibrators responses. Participants' predictive and postdictive calibration judgments from the open-ended interview questions were used to calculate calibration accuracy scores for both the easier and more difficult math problem. Median splits of these four scores were then used to categorize the participants as either more accurate or less accurate calibrators according to whether they fell in the top or bottom half of the median for each score (because of the small sample size, equal group proportions were approximated). These four median splits were then used to perform chi-square analyses to determine whether more accurate and less accurate calibrators answered the qualitative questions outlined above differently.

Chi-square analysis using the median split for postdictive accuracy for the easier question on the interview revealed significant differences between the more accurate and less accurate calibrators on the sources they used to make their postdictive judgments for that item, χ^2 (4, N = 29) = 11.606, p = .021. Specifically more accurate calibrators were more likely to use prior knowledge to make their postdictive judgment whereas less accurate calibrators were more likely to report that they guessed or used "gut" feelings to make their postdictive judgment. See Figure 4 below for a graphic representation of these findings. Chi-square analyses comparing more accurate and less accurate calibrators revealed no other significant differences in the sources participants used to form predictive calibration judgments, the number of steps they took to solve the problems, what strategies they reported using, and their perceptions of the main reasons why they got the question right/wrong.

Figure 4

Calibration Accuracy and Source of Postdictive Calibration Judgments – Chi-Square Analysis



CHAPTER V

Discussion

The primary goal of this study was to test whether a self-regulated learning intervention effectively improved participants' math achievement, as well as their abilities to monitor their performance and reflect on their learning, as compared to a control group. Multi-factorial repeated measures analyses of variances (RMANOVA's) were used to test the effects of the intervention on the following dependent variables: predictive and postdictive calibration accuracy, math performance, and self-regulatory and metacognitive strategy use. These analyses showed that participants who received the intervention had higher predictive and postdictive calibration accuracy and higher experimental math performance as compared to the control group, but did not report using more self-regulatory and metacognitive strategy use. Additionally, analysis of the qualitative data collected after the intervention suggests that the current treatment and calibration accuracy may impact the sources of information that students use to form calibration judgments. These findings will be discussed in more detail in the sections below.

Calibration Accuracy and Math Performance

The results of the study show that the intervention successfully increased the calibration accuracy of the participants and their math performance during the training sessions. These findings provide further experimental support for the efficacy of multi-component SRL interventions that explicitly train students to monitor and reflect on domain specific strategies simultaneously being taught (Graham & Harris, 2003; Huff & Nietfeld, 2009; Schmitz & Perels, 2011). It also provides additional support for incorporating distributed monitoring exercises like those used by Nietfeld et al. (2006) and Zimmerman et al. (2011), into interventions for middle school students as well as for college students. These structured worksheets were a main component of the current study and provided participants with opportunities to monitor and reflect on their regulatory behavior.

The intervention positively impacted both calibration accuracy and math performance, which is consistent with previous research findings that more accurate calibration leads to more effective regulation of learning (Dunlosky & Rawson, 2011; Stahl, Pieschl, & Bromme, 2006; Thiede, Anderson, & Therriault, 2003). However, it should be noted that these results were found with data collected during treatment sessions and not with natural classroom data. Therefore, the results lack some degree of ecological validity. This point is addressed further in the limitations section below.

These findings are notable considering the brevity of the intervention. Many of the effective multi-component SRL studies were very time and resource intensive, which makes their adaptation to everyday classroom teaching difficult (Brookhart et al., 2004; Fuchs et al.,2003; Schmitz & Perels, 2011). These findings provide support for the growing literature showing that brief and well-targeted instruction that provides metacognitively-focused strategy training within the context that these skills will be used can have large beneficial effects (Huff & Nietfeld, 2009; Perels et al., 2005; Perels et al., 2009). The literature as a whole continues to suggest that teaching students about the synergistic relationship between SRL processes and how these processes relate to information they are currently learning will likely be the most cost-effective way of improving students' self-regulatory capacities, and ultimately their achievement (Graham & Harris, 2003; Huff and Nietfeld, 2009; Nietfield et al., 2006). Based on the current results, it will likely be beneficial if more research incorporates multiple training sessions so that participants can fine-tune their monitoring and calibration skills over many self-regulatory learning cycles. This should allow for commensurate improvements in metacognitive monitoring to feed forward and inform more effective reflection and ultimately, regulation of learning (Schmitz & Perels, 2011; Zimmerman et al., 2011).

Results from the repeated measure analysis of variance conducted on postdictive calibration accuracy show that participants in sixth grade responded to the treatment more strongly than participants in seventh grade, as compared to the control group. This difference arose during the fourth and fifth sessions of the experiment, suggesting that this difference was due to the difficulty of the content. Although both grades were learning about probability, the curriculum used an upward spiral to deepen conceptual knowledge for the seventh grade participants. This may have explained why the sixth grade participants to be less familiar with the higher level concepts included in the study, thus decreasing the accuracy of their calibration judgments. However, the treatment may have been able to offset the detrimental effects of the more difficult material.

Metacognitive and Self-regulated Learning Strategies

Hypotheses three and four outlined above were not supported as the intervention did not improve participants' self-reported strategy use. However, seventh grade participants reported using more regulatory strategies during studying and homework, and seeking information or help from others to improve their studying more often as compared to the sixth grade participants in the study. This finding contrasts with a recent study by Cleary and Chen (2009) that found that seventh grade students used less self-regulatory strategies than their sixth grade peers. This discrepancy may be explained by that fact that the participants in their study attended a large suburban public school, whereas the participants in this study attended a small, resource-intensive private school that explicitly emphasizes independence. Furthermore, Cleary and Chen

(2009) sampled 880 students, whereas the current sample only included 30 participants. Participants in the current study have likely been held to high academic standards during their academic careers and received coaching about strategy use and owning their own learning. Being educated in contexts similar to the current study setting may make participants more aware of and skilled at regulating their own learning and metacognition.

It is unclear why the current intervention did not increase participants' self-reported strategy use. Cleary and Platten's (2013) case study analysis of their Self-Regulated Empowerment Program intervention may help explain the lack of intervention effects. In their student, Cleary and Platten found that even though participants did not report any increase in their regulatory strategy use on the SRSI-SR, other evidence collected throughout the intervention showed that three out of the four participants changed their regulatory behavior to some degree. It may be that participants in the current study needed more instruction and practice using these strategies before they were able to consciously report using them. Another potential explanation could be that the survey instruments were not theoretically or conceptually well aligned with the interventions strategy instruction and were not sensitive enough to pick up changes that did occur. The PI selected the survey instruments because they were readily available and had desirable psychometric proprieties. However, even though the IMSR measured some of the skills taught in the intervention, the items were derived from a different theoretical perspective than the one used to design the intervention. Even though the SRSI-SR is aligned with the major theoretical perspective underlying the study, it measures broad self-regulatory strategy constructs, as compared to the few specific self-regulatory skills taught to participants during the intervention. Use of different instruments to measure changes in self-regulatory and metacognitive strategy use that were more aligned with the dynamic nature of

SRL may have produced different results (Butler & Winne, 1995; Zimmerman, 2008). Finally, both treatment and control group participants were prompted to report their strategy use continuously throughout the study. This may have altered participants self-reports about strategy use if they came to believe that this was an important piece of the learning processes or what the PI was attempting to investigate.

Sources of Confidence Judgments

The current study expanded on the methodology of Dinsmore and Parkinson (2012) by exploring the sources that students use to form confidence judgments, as well as their self-reported strategy use and attributions during math problem solving using open-ended questions. Quantitative analysis revealed that participants exposed to the treatment had higher calibration accuracy than the control group. Qualitative analysis revealed that participants in the treatment group were more likely to use metacognition and self-concept to make their predictive judgment on the more difficult math problem in the interview, whereas participants in the control group were more likely to use item characteristics, prior knowledge, and guessing. These results are puzzling and cannot be easily explained by current theoretical understandings. However, participants who were more accurately calibrated were also found to use different sources to form calibration judgments than less accurately calibrated participants. These differences are largely consistent with current theoretical understandings. Research suggests that poor calibration accuracy may be the result of using cues or information to form calibration judgments that do not predict achievement well, such as one's self-concept or previous calibration judgments (Bol et al., 2005; de Bruin & Van Gog., 2012; Hacker et al., 2000; Redford et al., 2012). The current findings show that accurate calibrators were more likely to use their prior knowledge to form postdictive calibration judgments on the easier math problem in the interview, which is a source relevant to their ability to solve the problem at hand. Meanwhile, less accurate calibrators were more likely to use sources that should not be as predictive of success, including guessing and their "gut" feelings. The fact that more participants with lower calibration accuracy were unable to report how they formed their judgment further supports the appropriate cue logic above. It may be that participants with higher calibration accuracy have a larger knowledge base to draw from, which allows them to make more accurate calibration judgments.

Educational Implications

Calibrations well established link to regulatory behavior and academic achievement makes it a primary target for intervention and instruction. This study contributes to the growing self-regulated learning literature demonstrating that calibration is a skill that can be taught and suggests that students can become aware of and show improvement in their metacognitive monitoring skills in a relatively short period of time (Huff & Nietfeld, 2009; Nietfeld et al., 2006; Schmitz & Perels, 2011). More broadly, this study intended to help fill the need for educational interventions that improve students' self-regulatory and metacognitive skills, and ultimately their performance. Importantly, this intervention was designed to be flexible in order to facilitate adaptation to other classrooms or content areas. The procedures outlined above can be incorporated into preexisting curriculum or can be used as an adjunct to classroom instruction to facilitate more reflective, strategic approaches to learning. In addition, the study attempted to shed light onto the sources of information students use to form metacognitive monitoring judgments and further understanding of the factors that contribute to accurate monitoring. This will eventually help psychologists design more effective educational interventions targeting these skills.

Current findings and other successful calibration and self-regulated learning interventions can directly inform the practices of school psychologists, teachers and other educators looking to enhance their students' capacities to regulate their learning. This intervention can be used by educators, including school psychologists, to support teaching and learning in many diverse settings and contexts. This study provides valuable information to school psychology practitioners who can help teachers implement evidence-based practices in their classrooms through consultation and professional development. Interventions like this can serve as another tool for school psychologists to use to improve students' learning and academic achievement and fit well into the current legislative push for *response to intervention* (RTI), which is a data-based educational method defined by a three-tiered system of academic and behavioral support provided to students according to their response to instruction (Sailor, 2009). Under this model, all students receive Tier I instruction, which must be evidence-based. Progress is continuously monitored to determine if each student is responding to the instruction appropriately. Students who are not making appropriate progress are given increasing levels of support (Tiers II and III) until there is evidence that they are responding to the current interventions. This intervention can be used to supplement a general Tier I curriculum or adapted to assist struggling learners who need more individualized Tier II or III assistance.

Furthermore, school psychologists who familiarize themselves with the self-regulated learning intervention literature may improve their consultative skills. Understanding the dynamics of self-regulated learning should allow for greater insight into student learning and knowledge of effective interventions and may ultimately lead to more effective instructional consultation and a more fruitful collaborative problem-solving process between school psychologists and teachers.

A unique feature of this study was that participants formed and recorded both pre- and postdictive calibration judgments. This provided additional opportunities for participants to reflect on their accuracy and understand the dynamics between monitoring and performance, which was hypothesized to strengthen the effects of the intervention. Since the overall treatment was successful, future interventions may benefit from incorporating this method.

The study also begins to shed light on the mechanisms underlying the current treatments effects, which was found to improve participants' calibration accuracy as compared to a control group. Qualitative analysis suggests that the intervention did not broadly impact the sources of information participants use to form confidence judgments, which indicates that other mechanisms also contributed to the effects of the intervention. These findings can help inform research aimed at discovering the mechanisms whereby interventions improve calibration accuracy. The current findings that explored the different sources of calibration judgments used by more accurate and less accurate participants also strengthens the contention that teaching students to use appropriate cues to predict their performance will improve monitoring accuracy (de Bruin & Van Gog, 2012).

More research is also needed to determine which elements of self-regulatory interventions most effectively enhance monitoring and reflection processes. This study was not able to investigate this because the treatment incorporated features from many efficacious self-regulatory interventions, preventing analysis of specific intervention components in isolation.

Limitations

There are several limitations of the current study. First, the study was conducted on a small and select population, which may limit the external validity of the results. The research was conducted at a small, progressive, private middle school that enrolled students from a relatively high socio-economic status population. As such, the results may not generalize beyond this sample to the majority of middle school students (Pelham & Blanton, 2013). This sample was used because the PI served as a school psychology intern at the school and the administration encouraged research in their school. Further, the study was not powered to detect small to medium effect sizes. The small sample may have prevented detection of smaller intervention effects and may have prevented stronger, more meaningful results from surfacing (Glass & Hopkins, 1996). The small sample may have also reduced the reliability of the instruments used, leading to less psychometrically rigorous measurement

Second, current data do not allow for a thorough examination of the interventions effects on naturalistic math performance and calibration accuracy. The only analyzable measurement of natural math performance collected in this study was the shared item bank that all teachers agreed to use on the final unit test. Analyses of these few shared items on the final unit test did not reveal any performance differences for either treatment or grade. Therefore, the results of this study suffer from a lack of ecological validity (Kaufman & Kaufman, 2005). This suggests that it may be more difficult to create changes in the classroom than during experimental sessions. Thus, more research is needed to examine the longitudinal effects of monitoring and reflection interventions and how these interventions impact student achievement in the classroom.

Third, the PI had initially proposed to have participants graph their own confidence judgments and calibration accuracy after receiving training on this procedure during the first session. However, the participants took much longer than initially projected to construct their graphs during the first session, so the PI constructed the calibration accuracy graphs for each participant during the remaining sessions due to time constrains. The graphs were still used to provide visual feedback and serve as a platform for discussion about calibration, but participants were no longer graphing the feedback themselves, which may have made the feedback less salient.

Fourth, the PI should have applied more stringent procedures when coding the qualitative responses. All disagreements between the raters were discussed until consensus was reached. This prevented any information from being collected on the inter-rater reliability between the PI and the other rater. A better approach would have been to record inter-rater reliability and have a third rater solve any outstanding disagreements.

Conclusion

The purpose of this study was to investigate the effects of a self-regulatory strategy intervention designed to improve participants' calibration accuracy, self-regulatory skills, and math achievement. The study was designed to contribute to the growing literature base evaluating monitoring and calibration interventions as well as to begin to explore the mechanisms underlying their effects. The intervention incorporated elements of many efficacious monitoring and self-regulation interventions into one curriculum.

As hypothesized, those who received the intervention showed improvements in their predictive and postdictive calibration accuracy and higher math performance as compared to those in the control group. However, the intervention did not impact participants self-reported self-regulatory or metacognitive strategy use, refuting hypotheses three and four. Qualitative data suggest that participants use different sources of information for their calibration judgments depending on how accurate their calibration judgments were.

Research on interventions that improve students' abilities to monitor and regulate their learning, like the one used in the current study, is educationally valuable. The strong links between self-regulatory skills such as monitoring, reflection, and achievement and the fact that a large portion of students are found to be deficient in these skills make this is an important area for intervention (Bol et al., 2010; Dunlosky & Rawson, 2011; Hacker et al., 2008; Nietfeld et al., 2005). This research can be used to enhance school psychologist and teacher effectiveness, and can help fulfill the mandate to use “scientifically based” instruction in classrooms put forth by recent federal educational legislation (U.S. Department of Education, 2003).

Appendix A

Introductory Email from the **Head of School**

Dear Parent:

I am writing to inform you about a research study developed by Gregory DiGiacomo, one of the school psychology interns at Bay Ridge Prep, for his dissertation. The study investigates the effects of a brief five session program aimed at helping students become more strategic mathematical thinkers. This study will add an additional component to your child's math curriculum and allow us to look at its effectiveness. Attached is Greg's invitation, which outlines the course in more detail, and a permission form. If you would like to enroll your child into the study, please download the attached permission form, sign it electronically, and email the completed form to gdigiacomo@bayridgeprep.org. If you would prefer to submit a hard copy of the form, please print and sign the form, place it in a sealed envelope and return it to Greg's mailbox located in the main office.

Sincerely,

Charles Fasano

Appendix B

Invitation Letter from the PI

Dear Parent:

Hello, my name is Gregory DiGiacomo. I am a school psychology intern at Bay Ridge Prep. I have been working towards a doctoral degree in psychology with a focus on children and education. This experience has been both enlightening and challenging. Working with your children and the faculty at Bay Ridge Prep has reinforced my belief that I have chosen the right career path. After studying for six years, my final requirement to earn my doctorate is to conduct a research study. I would like to invite your child to participate in the study which is outlined below.

I have created a brief five session program to help students become more strategic mathematical thinkers. More specifically, your child will be taught to:

- Examine and discuss their current approach to solving math problems
- Reflect on if and why their approach is working
- Accurately judge their understanding of math concepts
- Learn to incorporate and execute these skills

These sessions will be built around your child's math curriculum and will reinforce the strategies they are currently being taught.

Participation provides an opportunity to practice these skills in a no-pressure environment. None of the work they complete during this training will impact their grades. Every student, regardless of ability, can benefit from this program because it will deepen conceptual understanding of mathematics, as well as foster a more reflective and analytical approach to problem solving.

If you have any questions, feel free to contact me at 631-793-9156 or gdigiacomo@gc.cuny.edu. Attached is a permission form. If you would like to enroll your child into the study, please download the form, sign it electronically, and email the completed form to the above email address. If you would prefer to submit a hard copy of the form, please print and sign the form, place it in a sealed envelope and return it to my mailbox located in the main office. Thank you for your consideration.

Sincerely,
Gregory DiGiacomo

Appendix C

Permission Form for Sixth Grade Parents

CITY UNIVERSITY OF NEW YORK

The Graduate Center
Department of Educational Psychology

PARENTAL/LEGAL GUARDIAN PERMISSION FORM AND AUTHORIZATION FOR CHILD'S PARTICIPATION IN RESEARCH

Project Title: Enhancing self-monitoring and self-reflection through a self-regulatory skills intervention embedded in a middle school mathematics curriculum.

Principal Investigator (PI): Gregory DiGiacomo, The Graduate Center of the City University of New York, 365 5th Avenue, 10016, 631-793-9156.

Faculty Advisor: Dr. Peggy Chen, Associate professor, Department of Educational Foundations & Counseling Programs, Hunter College, West 1123, 695 Park Avenue, NY, NY 10065, 212-772-4754

Your child is invited to participate in a research study. The study is conducted under the direction of Gregory DiGiacomo, a school psychology intern at Bay Ridge Preparatory and an Educational Psychology Ph.D. student at The Graduate Center of the City University of New York. The study will investigate the effects of a brief training program designed to improve students' abilities to monitor their performance and reflect on their progress during mathematical problem-solving.

Procedures: All sixth and seventh grade students enrolled in mainstream math classes at Bay Ridge Prep have been invited to participate. Approximately 25-30 students are expected to participate in this study (each grade will be seen separately). The training component of the study consists of five group sessions focused on strategy instruction (approximately 45 minutes each). The training will be conducted during one unit in their math class, which will take approximately three weeks. All sessions will take place at Bay Ridge Prep Middle School, 8101 Ridge Blvd., Brooklyn, NY 11209. Sessions will be held during your child's academic enrichment (two times) or lunch periods (three times) unless your child does not have an academic enrichment period. In this case, two sessions will be held afterschool (multiple dates will be available in case of scheduling conflicts). Any time your child gives up their lunch period to work with Greg, they will be provided with two slices of pizza or a peanut butter and jelly sandwich if they prefer.

If your child participates, they will be assigned to one of two groups, a treatment group and a delay-treatment group. The delay-treatment group will receive the training once the treatment group has completed the full five sessions. This is necessary so that the group that initially receives the training can be compared to a group that did not, but also ensures that all participating students will receive the benefits of the training. While the treatment group is receiving the training, the delay-treatment group will participate in five short data collection sessions which will occur during their homeroom periods (approximately 12 minutes each).

In addition, your child will fill out a brief survey three times over the course of the first five sessions (approximately 8 minutes each time). This survey is designed to measure their use of self-regulatory strategies and their thinking during math problem solving. After the initial training component ends the PI will meet individually with your child one time to ask them a few questions about the sources of information that they use to monitor their performance. This will occur during their homeroom period (approximately 5 minutes). Finally, records of your child's previous math achievement (scores on the Iowa Test of Basic Skills and current grade point average) will be collected and analyzed to account for their achievement levels prior to the study.

Benefits: Your child is likely to benefit from this program as they will be trained to approach math in a more reflective, analytic way. In turn, this may deepen their conceptual understanding and improve their math achievement. Many similar programs have improved student knowledge and use of strategies, as well as their achievement.

Voluntary Participation: Your child's participation in this study is voluntary. At anytime during the study, you may withdraw your child from participation without prejudice, penalty, or loss of benefits to which you are otherwise entitled. If you decide to remove your child from the study, please contact the PI, Gregory DiGiacomo, to inform him of your decision.

Possible Discomforts and Risks: There is no more risk involved in participating in this study than there is in a regular school classroom. Some students experience discomfort when they solve math problems. This study could potentially create some distress by adding to the amount of time in the week students spend doing math. To ensure that students experience as little discomfort as possible during the training, the principal investigator (PI) will emphasize that students are learning strategies that can help them overcome any difficulties they are having and that none of their work during theses sessions will affect their grades. If your child experiences any discomfort as a result of this study and would like to seek professional help, you should contact Dr. Jen Galbo, the school psychologist.

Confidentiality: During the study your child will fill out a number of surveys and worksheets. All information gathered will be kept confidential and will only be accessible to the PI and his faculty adviser. All personal identifying information will be removed from the data before data analysis. All electronic files will be password protected and kept on the PI's home computer. All paper documents collected during the study will be stored in a locked filing cabinet. The data will be stored for a minimum of three years, after which all data will be destroyed; all worksheets will be shredded and all electronic files will be permanently deleted. As long as the data exists, it will be kept secured. The information will be used to produce a paper for a graduate research project. Only aggregate data will be reported in any reports or publications derived from this research. If you would like a copy of the study, please provide me with your address and I will send you a copy when the study is completed.

Contact Questions/Persons: If you or your child have any questions about the research, you should contact the Principal Investigator, Gregory DiGiacomo at (631) 793-9156 or gdigiacomo@gc.cuny.edu. If you or your child has any questions concerning your child's rights as a participant in this study, you may contact the Hunter College Human Research Protection Program (HRPP) Office at (212) 650-3053 or hrpp@hunter.cuny.edu.

Statement of Consent: "I have read the above description of this research and I understand it. I have been informed of the risks and benefits involved, and all my questions have been answered to my satisfaction. Furthermore, I have been assured that any future questions that I may have will also be answered by the principal investigator of the research study. I voluntarily agree to allow my child to participate in this study.

By signing this form I have not waived any of my legal rights to which my child would otherwise be entitled.

I will be given a copy of this statement."

	☐	/ /
Printed Name of Subject's Legal Guardian	Electronic Signature of Subject's Legal Guardian	Date Signed

______________________	____________________________	________
Printed Name of Investigator	Signature of Investigator	Date Signed

Permission Form for Seventh Grade Parents

CITY UNIVERS1TY OF NEW YORK

The Graduate Center
Department of Educational Psychology

**PARENTAL/LEGAL GUARDIAN PERMISSION FORM
AND AUTHORIZATION FOR
CHILD'S PARTICIPATION IN RESEARCH**

Project Title: Enhancing self-monitoring and self-reflection through a self-regulatory skills intervention embedded in a middle school mathematics curriculum.

Principal Investigator (PI): Gregory DiGiacomo, The Graduate Center of the City University of New York, 365 5th Avenue, 10016, 631-793-9156.

Faculty Advisor: Dr. Peggy Chen, Associate professor, Department of Educational Foundations & Counseling Programs, Hunter College, West 1123, 695 Park Avenue, NY, NY 10065, 212-772-4754

Your child is invited to participate in a research study. The study is conducted under the direction of Gregory DiGiacomo, a school psychology intern at Bay Ridge Preparatory and an Educational Psychology Ph.D. student at The Graduate Center of the City University of New York. The study will investigate the effects of a brief training program designed to improve students' abilities to monitor their performance and reflect on their progress during mathematical problem-solving.

<u>Procedures:</u> All sixth and seventh grade students enrolled in mainstream math classes at Bay Ridge Prep have been invited to participate. Approximately 25-30 students are expected to participate in this study (each grade will be seen separately). The training component of the study consists of five group sessions focused on strategy instruction (approximately 45 minutes each). The training will be conducted during one unit in their math class, which will take approximately three weeks. All sessions will take place at Bay Ridge Prep Middle School, 8101 Ridge Blvd., Brooklyn, NY 11209. Sessions will be held during your child's academic enrichment (three times) and lunch periods (two times). Each time your child gives up their lunch period to work with Greg, they will be provided with two slices of pizza or a peanut butter and jelly sandwich if they prefer.

If your child participates, they will be assigned to one of two groups, a treatment group and a delay-treatment group. The delay-treatment group will receive the training once the treatment group has completed the full five sessions. This is necessary so that the group that initially receives the training can be compared to a group that did not, but also ensures that all participating students will receive the benefits of the training. While the treatment group is receiving the training, the delay-treatment group will participate in five short data collection sessions which will occur during their homeroom periods (approximately 12 minutes each).

In addition, your child will fill out a brief survey three times over the course of the first five sessions (approximately 8 minutes each time). This survey is designed to measure their use of self-regulatory strategies and their thinking during math problem solving. After the initial training component ends the PI will meet individually with your child one time to ask them a few questions about the sources of information that they use to monitor their performance. This will occur during their homeroom period (approximately 5 minutes). Finally, records of your child's previous math achievement (scores on the Iowa Test of Basic Skills and current grade point average) will be collected and analyzed to account for their achievement levels prior to the study.

Benefits: Your child is likely to benefit from this program as they will be trained to approach math in a more reflective, analytic way. In turn, this may deepen their conceptual understanding and improve their math achievement. Many similar programs have improved student knowledge and use of strategies, as well as their achievement.

Voluntary Participation: Your child's participation in this study is voluntary. At anytime during the study, you may withdraw your child from participation without prejudice, penalty, or loss of benefits to which you are otherwise entitled. If you decide to remove your child from the study, please contact the PI, Gregory DiGiacomo, to inform him of your decision.

Possible Discomforts and Risks: There is no more risk involved in participating in this study than there is in a regular school classroom. Some students experience discomfort when they solve math problems. This study could potentially create some distress by adding to the amount of time in the week students spend doing math. To ensure that students experience as little discomfort as possible during the training, the principal investigator (PI) will emphasize that students are learning strategies that can help them overcome any difficulties they are having and that none of their work during theses sessions will affect their grades. If your child experiences any discomfort as a result of this study and would like to seek professional help, you should contact Dr. Jen Galbo, the school psychologist.

Confidentiality: During the study your child will fill out a number of surveys and worksheets. All information gathered will be kept confidential and will only be accessible to the PI and his faculty adviser. All personal identifying information will be removed from the data before data analysis. All electronic files will be password protected and kept on the PI's home computer. All paper documents collected during the study will be stored in a locked filing cabinet. The data will be stored for a minimum of three years, after which all data will be destroyed; all worksheets will be shredded and all electronic files will be permanently deleted. As long as the data exists, it will be kept secured. The information will be used to produce a paper for a graduate research project. Only aggregate data will be reported in any reports or publications derived from this research. If you would like a copy of the study, please provide me with your address and I will send you a copy when the study is completed.

Contact Questions/Persons: If you or your child have any questions about the research, you should contact the Principal Investigator, Gregory DiGiacomo at (631) 793-9156 or gdigiacomo@gc.cuny.edu. If you or your child has any questions concerning your child's rights as a participant in this study, you may contact the Hunter College Human Research Protection Program (HRPP) Office at (212) 650-3053 or hrpp@hunter.cuny.edu.

Statement of Consent: "I have read the above description of this research and I understand it. I have been informed of the risks and benefits involved, and all my questions have been answered to my satisfaction. Furthermore, I have been assured that any future questions that I may have will also be answered by the principal investigator of the research study. I voluntarily agree to allow my child to participate in this study.

By signing this form I have not waived any of my legal rights to which my child would otherwise be entitled.

I will be given a copy of this statement."

	☐	/ /
Printed Name of Subject's Legal Guardian	Electronic Signature of Subject's Legal Guardian	Date Signed

Printed Name of Investigator	Signature of Investigator	Date Signed

Appendix D

Recruitment Script

(The teacher will be asked to leave the room during this announcement to ensure that students feel they can say they do not want to participate without feeling pressured by their teachers.)

Hi, for those of you who do not know me I am Greg DiGiacomo, one of the school psychology interns here at Bay Ridge Prep. I have really enjoyed working here this year so far and am excited to tell you all about a study I am conducting designed to improve student's mathematical thinking. I would like to invite you to participate in the study.

It involves five 45 minute lessons where we will review strategies to break down math problems and decide on the best approach to use to solve them. These lessons may enhance your understanding of math concepts and teach you strategies to use in the future. The lessons will be built around the next two units you will be covering in this math class and will take place during your academic enrichment periods, lunch periods and after (if you are in 6th grade and do not have an academic enrichment period). Each time you give up your lunch period, you will be provided with two slices of pizza or a peanut butter and jelly sandwich if you prefer.

Participating in this study will not hurt your grade in any way. Every one of you can benefit from the course, no matter how good you are at math. The lessons are designed to help you think about math in a more reflective way.

I am going to hand out some papers for you and your parents to look through that have more information about the study. Please bring these back to your house and give them to your parents. I am also going to email them to your parents. If you are interested in participating, your parents can electronically sign the permission form and send them back through email, or they can sign this form and drop it off in my mail box, which is located in the main office.

Does anyone have any questions?

Appendix E

Assent Form for Sixth Grade Students

CITY UNIVERSITY OF NEW YORK

The Graduate Center
Department of Educational Psychology

ASSENT TO PARTICIPATE IN A RESEARCH PROJECT

Child's Name: ________________

You are invited to participate in Gregory DiGiacomo's research study. The study will test the effects of a brief strategy training program. The program will teach you how to use learning strategies based on past successes and failures and help you judge how well you learn. This study may enhance your understanding of math concepts and teach you strategies to use in the future.

What will happen to me in this study?

You and 50 other students are being invited to participate in this program. Participants will take part in five group training sessions (approximately 45 minutes each). These training sessions will be completed over the course of one three week unit in your math class. All sessions will take place at Bay Ridge Prep Middle School, 8101 Ridge Blvd., Brooklyn, NY 11209. Sessions will be held during your academic enrichment (two times) or lunch periods (three times) unless you do not have an academic enrichment period. In this case, two sessions will be held afterschool (multiple dates will be available in case of scheduling conflicts). Each time you give up your lunch period, you will be provided with two slices of pizza or a peanut butter and jelly sandwich if you prefer.

Participating students will be split into two groups, a treatment group and a delay-treatment group. The delay-treatment group will receive the training once the treatment group has completed the full five sessions. While the treatment group is receiving the training, the delay-treatment group will participate in five short sessions during their homeroom periods to collect information for comparison (approximately 12 minutes each).

In addition, all participants will fill out a brief survey three times over the course of the first five sessions during their homeroom period (approximately 8 minutes each time). At the end of the first five sessions, I will meet individually with each participant to conduct a brief interview during their homeroom period (approximately 5 minutes).

Will I get hurt?

There is no more risk involved in participating in this study than there is in a regular school classroom. Some students experience discomfort when they solve math problems. This study could potentially create some distress by adding to the amount of time in the week you spend doing math. To minimize these risks, I will teach you strategies that can help you overcome any difficulties you are having. In addition, none of your work during theses sessions will affect your grades. If you experience any discomfort as a result of this study and would like to seek help, you should tell me, your parent/guardian, or someone else you know right away.

Will anything good happen to me?

You are likely to benefit from this training. Many similar programs have improved student achievement as well as knowledge and use of strategies. The targeted skills may improve your achievement in math and other subjects.

What if I do not want to do this?

You don’t have to be in this study. No one will be mad at you if you don’t want to do this. If you don’t want to be in this study, just tell us. If you want to be in this study, just tell us. Remember, it is ok to say yes now and change your mind later. Nothing will happen to you if you decide to stop.

Will anyone know I was involved?

Your name and the fact that you are in this study will be kept confidential.

Who can I talk to about this study?

You can ask questions any time. You can ask now. You can ask later. You can talk to me or someone else, like Dr. Jen Galbo.

Do you want to participate in this study? ☐ *Yes* ☐ *No*

Additional Information:

Project Title: Enhancing self-monitoring and self-reflection through a self-regulatory skills intervention embedded in a middle school mathematics curriculum.

Principal Investigator: Gregory DiGiacomo

Faculty Advisor: Dr. Peggy Chen

PERSON CONDUCTING ASSENT

I have explained the study to ______________________________ (*name of child*) in language he/she understands, and he/she has agreed to be in the study.

_______________________________	_______________________________	_____________
Name of Person Conducting Assent (*print*)	Signature of Person Conducting Assent	Date Signed
_______________________________	_______________________________	_____________
Name of Investigator (*print*)	Signature of Investigator	Date Signed

Assent Form for Seventh Grade Students

CITY UNIVERSITY OF NEW YORK

The Graduate Center
Department of Educational Psychology

ASSENT TO PARTICIPATE IN A RESEARCH PROJECT

My name is Gregory DiGiacomo. I am a student in the Educational Psychology Ph.D. Program at The Graduate Center of the City University of New York. You are invited to participate in my research study which will test the effects of a brief strategy training program. The program will teach you how to use learning strategies based on past successes and failures and help you judge how well you learn. This study may enhance your understanding of math concepts and teach you strategies to use in the future.

Procedures: You and 50 other students are being invited to participate in this program. Participants will take part in five group training sessions (approximately 45 minutes each). These training sessions will be completed over the course of one three week unit in your math class. All sessions will take place at Bay Ridge Prep Middle School, 8101 Ridge Blvd., Brooklyn, NY 11209 during your academic enrichment (three times) and lunch periods (two times). Each time you give up your lunch period, you will be provided with two slices of pizza or a peanut butter and jelly sandwich if you prefer.

Participating students will be split into two groups, a treatment group and a delay-treatment group. The delay-treatment group will receive the training once the treatment group has completed the full five sessions. While the treatment group is receiving the training, the delay-treatment group will participate in five short sessions during their homeroom periods to collect information for comparison (approximately 12 minutes each).

In addition, all participants will fill out a brief survey three times over the course of the first five sessions during their homeroom period (approximately 8 minutes each time). At the end of the first five sessions, I will meet individually with each participant to conduct a brief interview during their homeroom period (approximately 5 minutes).

Benefits: You are likely to benefit from this training. Many similar programs have improved student achievement as well as knowledge and use of strategies. The targeted skills may improve your achievement in math and other subjects.

Possible Discomforts and Risks: There is no more risk involved in participating in this study than there is in a regular school classroom. Some students experience discomfort when they solve math problems. This study could potentially create some distress by adding to the amount of time in the week you spend doing math. To minimize these risks, I will teach you strategies that can help you overcome any difficulties you are having. In addition, none of your work during theses sessions will affect your grades. If you experience any discomfort as a result of this study you should contact Dr. Jen Galbo, the school psychologist.

Voluntary Participation: Your participation in this study is voluntary, and you may decide not to participate without penalty. If you decide to leave the study, please contact me to let me know.

Confidentiality: During the study you will fill out a number of surveys and worksheets. Your name and the fact that you are in this study will be kept confidential.

Contact Questions/Persons: If you have any questions about the research now or in the future, you may ask the researcher now or contact the Principal Investigator, Gregory DiGiacomo, at (631) 793-9156 or gdigiacomo@gc.cuny.edu. If you have any questions concerning your rights as a participant in this study, you may contact Wankairys Decena at (212) 650-3053 or wdecena@hunter.cuny.edu.

Additional Information:

Project Title: Enhancing self-monitoring and self-reflection through a self-regulatory skills intervention embedded in a middle school mathematics curriculum.

Principal Investigator: Gregory DiGiacomo

Faculty Advisor: Dr. Peggy Chen

Statement of Consent:

"I have read the above description of this research and I understand it. All my questions have been answered to my satisfaction. I voluntarily agree to participate in this study.

I will be given a copy of this statement."

____________________	______________________________________	___________________
Printed Name of Subject	Signature of Subject	Date Signed
____________________	______________________________________	___________________
Printed Name of Person Explaining Assent Form	Signature of Person Explaining Assent Form	Date Signed
____________________	______________________________________	___________________
Printed Name of Investigator	Signature of Person Explaining Assent Form	Date Signed

Appendix F

Sample Review Questions with Corresponding Confidence Judgments

Directions: You will now examine a number of math problems from the Probability unit. Please read each math problem WITHOUT solving it. Then rate how confident you are that you can solve the problem correctly. Please circle ONLY ONE number to represent your confidence level.

1. A bag contains 3 red marbles, 4 white marbles, and 8 blue marbles. You pick a marble without looking. Find the probability of drawing a white marble.

1: How confident are you that you can solve the above question correctly?

1 2 3 4 5 6 7 8 9 10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

1: What strategy would you use to solve this problem?

____

2. A standard number cube with the numbers 1 through 6 is rolled. Find the probability of rolling a number greater than 2.

2: How confident are you that you can solve the above question correctly?

1 2 3 4 5 6 7 8 9 10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

2: What strategy would you use to solve this problem?

____

3. Juan tossed a coin 75 times. The coin landed heads up 50 times and tails up 25 times. Can you conclude that the coin is not a fair coin? Explain.

3: How confident are you that you can solve the above question correctly?

1	2	3	4	5	6	7	8	9	10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

3: What strategy would you use to solve this problem?

____

____

4. Betty empties her piggy bank, which contains 210 coins, out onto her desk. How many of the coins would you expect to be heads up?

4: How confident are you that you can solve the above question correctly?

1	2	3	4	5	6	7	8	9	10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

4: What strategy would you use to solve this problem?

____

____

5. Bobby tosses an action figure ten times. It lands on its base eighth times and on its side twice. Phillip tosses the same piece 100 times. It lands on its base 23 times and on its side 77 times. Based on their data, if you toss the piece one more time, is it more likely to land on its base or its side? Why?

5: How confident are you that you can solve the above question correctly?

1	2	3	4	5	6	7	8	9	10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

5: What strategy would you use to solve this problem?

____

____

Directions: Now you will get to solve the same five math problems that you just saw. Please show your work.

1. A bag contains 3 red marbles, 4 white marbles, and 8 blue marbles. You pick a marble without looking. Find the probability of drawing a white marble.

---

2. A standard number cube with the numbers 1 through 6 is rolled. Find the probability of rolling a number greater than 2.

---

3. Juan tossed a coin 75 times. The coin landed heads up 50 times and tails up 25 times. Can you conclude that the coin is not a fair coin? Explain.

4. Betty empties her piggy bank, which contains 210 coins, out onto her desk. How many of the coins would you expect to be heads up?

5. Bobby tosses an action figure ten times. It lands on its base eighth times and on its side twice. Phillip tosses the same piece 100 times. It lands on its base 23 times and on its side 77 times. Based on their data, if you toss the piece one more time, is it more likely to land on its base or its side? Why?

<u>Directions</u>: Now that you have solved these math problems please rate how confident you are that you solved each problem correctly. DO NOT solve the problems again, just circle THE ONE number that best represents your confidence level for each problem.

1. A bag contains 3 red marbles, 4 white marbles, and 8 blue marbles. You pick a marble without looking. Find the probability of drawing a white marble.

<u>1</u>: How confident are you that you solved the above math problem correctly?

1 2 3 4 5 6 7 8 9 10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

---

2. A standard number cube with the numbers 1 through 6 is rolled. Find the probability of rolling a number greater than 2.

<u>2</u>: How confident are you that you solved the above math problem correctly?

1 2 3 4 5 6 7 8 9 10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

---

3. Juan tossed a coin 75 times. The coin landed heads up 50 times and tails up 25 times. Can you conclude that the coin is not a fair coin? Explain.

<u>3</u>: How confident are you that you solved the above math problem correctly?

1 2 3 4 5 6 7 8 9 10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

---

4. Betty empties her piggy bank, which contains 210 coins, out onto her desk. How many of the coins would you expect to be heads up?

4: How confident are you that you solved the above math problem correctly?

1 2 3 4 5 6 7 8 9 10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

---

5. Bobby tosses an action figure ten times. It lands on its base eighth times and on its side twice. Phillip tosses the same piece 100 times. It lands on its base 23 times and on its side 77 times. Based on their data, if you toss the piece one more time, is it more likely to land on its base or its side?

5: How confident are you that you solved the above math problem correctly?

1 2 3 4 5 6 7 8 9 10

Not Confident *Somewhat Confident* *Pretty Confident* *Very Confident*

Appendix G

Self-Regulation Strategy Inventory-Self-Report (SRSI-SR) – (Cleary, 2006)

How Do You Study For Math Tests and Do Math Homework?

Directions: The purpose of this section is to see how you study for your **MATH tests** or do **MATH** homework. There are a total of 28 sentences. For each statement, please fill in **ONE** circle to indicate **HOW OFTEN** you do each of these things when studying for **MATH tests** or doing **MATH** homework or

To answer these questions, use the following 5-point scale:

1 Almost never	**2 Not very often**	**3 Somewhat often**	**4 Very often**	**5 Almost always**
○	○	○	○	○

Things I do when doing MATH homework or studying for MATH tests	**1 Almost never**	**2 Not very often**	**3 Somewhat often**	**4 Very often**	**5 Almost always**
1. I tell myself to keep trying hard when I get confused	○	○	○	○	○
2. I give up or quit when I do not understand something.	○	○	○	○	○
3. I try to study in a quiet place.	○	○	○	○	○
4. I ask my math teacher about the topics that will be on upcoming tests.	○	○	○	○	○
5. I use my class notes to study.	○	○	○	○	○
6. I study hard even when there are more fun things to do at home.	○	○	○	○	○
7. I quiz myself to see how much I am learning during studying.	○	○	○	○	○
8. I lose important dittos/worksheets that I need to study.	○	○	○	○	○
9. I make a schedule to help me organize my study time.	○	○	○	○	○

10. I use binders or folders to organize my study materials.	○	○	○	○	○
11. I think about the types of questions that might be on a test.	○	○	○	○	○
12. I try to see how my notes from math class relate to things I already know.	○	○	○	○	○
13. I try to identify the format of upcoming tests (e.g., multiple-choice or short-answer questions).	○	○	○	○	○
14. I try to study in a place that has no distractions (e.g., noise, people talking).	○	○	○	○	○
15. I forget to ask my teacher questions about things that confuse me.	○	○	○	○	○
16. I wait to the last minute to start studying for upcoming tests.	○	○	○	○	○
17. I try to forget about the topics that I have trouble learning.	○	○	○	○	○
18. I ask my teacher questions when I do not understand something.	○	○	○	○	○
19. I make pictures or diagrams to help me learn math concepts.	○	○	○	○	○
20. I make sure no one disturbs me when I study.	○	○	○	○	○
21. I tell myself exactly what I want to accomplish before studying.	○	○	○	○	○
22. I let my friends interrupt me when I am studying.	○	○	○	○	○
23. I look over my homework assignments if I don't understand something.	○	○	○	○	○
24. I carefully organize my study materials so I don't lose them.	○	○	○	○	○
25. I think about the best way to study for each math test.	○	○	○	○	○
26. I avoid asking questions in class about things I don't understand.	○	○	○	○	○
27. I finish all of my studying before I play video games or play with my friends.	○	○	○	○	○
28. I forget to bring home my study materials when I need to study for math tests.	○	○	○	○	○

Appendix H

Inventory of Metacognitive Self-Regulation (IMSR) – (Howard, McGee, Shia & Hong, 2000)

How Do You Solve Problems?

Directions: There are a total of 32 sentences. For each statement, please fill in **ONE** circle to indicate **HOW OFTEN** you do each of these things when you are trying to solve a **MATH** problem.

· Think about when you have to solve a hard problem. What do you do before you start?
· What do you do while you work the problem?
· What do you do after you finish working the problem?

There are no right answers--please describe yourself as you are, not how you want to be or think you ought to be.

To answer these questions, use the following 5-point scale:

1 Almost never	2 Not very often	3 Somewhat often	4 Very often	5 Almost always
○	○	○	○	○

Things I do when solving MATH problems	1 Almost never	2 Not very often	3 Somewhat often	4 Very often	5 Almost always
1. I try to understand what the problem is asking me.	○	○	○	○	○
2. I think of several ways to solve a problem and then choose the best one.	○	○	○	○	○
3. I look back at the problem to see if my answer makes sense.	○	○	○	○	○
4. I use different ways to memorize things.	○	○	○	○	○
5. I think to myself, do I understand what the problem is asking me?	○	○	○	○	○
6. I read the problem more than once.	○	○	○	○	○
7. I think about what information I need to solve this problem.	○	○	○	○	○
8. I use different learning strategies depending on the problem.	○	○	○	○	○
9. I look back to see if I did the correct procedures.	○	○	○	○	○
10. I think about how well I am learning when I work a difficult problem.	○	○	○	○	○

11. I use different ways of learning depending on the problem.	○	○	○	○	○
12. I go back and check my work.	○	○	○	○	○
13. I read the problem over and over until I understand it..	○	○	○	○	○
14. I stop and rethink a step I have already done.	○	○	○	○	○
15. I check to see if my calculations are correct.	○	○	○	○	○
16. When it comes to learning, I can make myself learn when I need to.	○	○	○	○	○
17. I ask myself how well I am doing while I am learning something new.	○	○	○	○	○
18. I check my work all the way through the problem.	○	○	○	○	○
19. I identify all the important parts of the problem.	○	○	○	○	○
20. I try to understand the problem so I know what to do.	○	○	○	○	○
21. I make sure I complete each step.	○	○	○	○	○
22. I can make myself memorize something.	○	○	○	○	○
23. When it comes to learning, I know my strengths and weaknesses.	○	○	○	○	○
24. I pick out the steps I need to do this problem.	○	○	○	○	○
25. When I am done with my schoolwork, I ask myself if I learned what I wanted to learn.	○	○	○	○	○
26. I double-check to make sure I did it right.	○	○	○	○	○
27. try to eliminate information in the problem that I don’t need.	○	○	○	○	○
28. I try to break down the problem to just the necessary information.	○	○	○	○	○
29. I use learning strategies without thinking.	○	○	○	○	○
30. When it comes to learning, I know how I learn best.	○	○	○	○	○
31. I ask myself if there are certain goals I want to accomplish.	○	○	○	○	○
32. I try more than one way to learn something.	○	○	○	○	○

Appendix I

Formation of Judgments – Qualitative Questions

Each of these questions will be asked twice while the participant solves two probability math problems.

After completing each problem, they will be asked:

- How did you make this prediction of __ %?

- Explain all of the steps you took to solve the problem. Try to be as exact as possible.
- Did you use any strategies to solve the problem? If so, which ones?
  - How/why did you use these?
- How do you know whether you answered the question correctly?

Then, participants will be told whether they got the question right or wrong and asked:

- What do you think is the main reason why you got this problem right/wrong?

- Is there anything else you want to tell me about the problems or how you solved them?

Appendix J

Visual Overview of Intervention Components and Data Collection

Visual Overview of Treatment Sessions				
Session I	Session II	Session III	Session IV	Session V
5 math questions with predictive and postdictive judgments (12 min)	**5 math questions with predictive and postdictive judgments (12 min)**	**5 math questions with predictive and postdictive judgments (12 min)**	**5 math questions with predictive and postdictive judgments (12 min)**	**5 math questions with predictive and postdictive judgments (12 min)**
Overview of three-stage SRL model and general strategies - assigning strategies to the different stages (20 min)	Training of specific SRL strategies and domain-specific strategies (18 min) - Focus on forethought	Training of specific SRL strategies and domain-specific strategies (18 min) -Focus on performance	Training of specific SRL strategies and domain-specific strategies (18 min) - Focus on reflection	Training of specific SRL strategies and domain-specific strategies (18 min)
Feedback on questions (5 min)	Feedback on questions (5 min)	Feedback on questions (5 min)	Feedback on questions (5 min)	Feedback on questions (5 min)
Overview of graphing procedure and practice graphing of calibration accuracy from current review questions (7 min)	Graphing of calibration accuracy (3 min)	Graphing of calibration accuracy (3 min)	Graphing of calibration accuracy (3 min)	Graphing of calibration accuracy (3 min)
Reflection worksheet with reflection of strategy use (7 min)	Reflection worksheet with reflection of strategy use (7 min)	Reflection worksheet with reflection of strategy use (7 min)	Reflection worksheet with reflection of strategy use (7 min)	Reflection worksheet with reflection of strategy (7 min)

* **Bold text indicates elements of intervention that also serve as data collection.**

Visual Overview of Control Group Sessions				
Session I	Session II	Session III	Session IV	Session V
5 math questions with predictive and postdictive judgments (12 min)	**5 math questions with predictive and postdictive judgments (12 min)**	**5 math questions with predictive and postdictive judgments (12 min)**	**5 math questions with predictive and postdictive judgments (12 min)**	**5 math questions with predictive and postdictive judgments (12 min)**
Individualized Math Instruction using Math Whizz (33 min)	Individualized Math Instruction using Math Whizz (33 min)	Individualized Math Instruction using Math Whizz (33 min)	Individualized Math Instruction using Math Whizz (33 min)	Individualized Math Instruction using Math Whizz (33 min)

* **Bold text indicates elements of intervention that also serve as data collection.**

Visual Overview of Data Collection
5 Math Review Questions all 5 sessions - **Given throughout the 3 week unit** **- With local predictive and postdictive judgments**
Classroom Assessment - **Given throughout the 3 week unit** **- 2 Quizzes** **- 1 Unit Test** **- With global predictive and postdictive judgments**
Metacognition Questionnaire & **SRL strategy survey** - **Given before session 1** - **Given before session 3** - **Given after session 5**
Interview Questions - **Given after session 5**

Appendix K

Details of Training Session

Session One

1. **During session one, I will explain the three-stage SRL model to the students. We will go through each stage and collectively assign various SRL strategies to the different stages so that students become familiar with the model.**

Begin with an introduction of the course and explain what the purpose is. Next hold a brief discussion of:

- What would you like to change about yourself academically?
- What has worked?
- What has not worked?

Today we are going to discuss a powerful way to look at your learning. Psychologists have developed a system to help people better understand how they learn. It helps people improve upon what they know. It consists of three phases.

a. Planning (Pre-action)

    - Breaking a problem down into smaller parts
    - Developing a strategy or a plan to help you solve a problem
        a. Thinking about similar problems
    - Forming a goal
        a. How well do you think you are going to complete this math problem?
    - Motivation
        a. Do I have what it takes?

b. Action

    - Giving effort
    - Using a strategy to help you solve a problem
    - Monitoring your performance
        a. Writing it down
    - Maintaining focus/attention

c. Reflection

    - Reflecting on your performance
    - Thinking about ways to improve your performance next time
        a. Would any particular strategies have helped?
    - Causes of outcomes – Attributions

- For the remainder of the 20 min training time, I will focus on the concept of monitoring and explain how important it is to the regulatory process.

- Monitoring – tracking your performance
- Allows you to judge your understanding of math problem/concept
    a. How much progress are you making?
        i. Why is this important
    b. You solved the problem correctly
        i. Great, figuring out why will help you continue succeeding
            1. Gives you feedback that you can use to improve your plan for the next math problem
    c. You cannot solve the problem
        i. Why not? What is preventing you from doing this?

- Gives you control of your learning process
- Examples:

1. Running a mile
    a. ¼ of the way you are breathing really heavy/very tired
        i. What does this feedback mean?
2. Solving a math problem
    a. You started trying to solve the problem and realize you cannot keep going
        i. What does this tell you?
        ii. What might you do next?

Session Two

2. During session two, we will explore the planning phase and its implications.

a. Planning (Pre-action)

- Breaking a problem down into smaller parts
- Developing a strategy or a plan to help you solve a problem
    a. Thinking about similar problems
- Forming a goal
    a. How well do you think you are going to complete this math problem?
- Motivation
    a. Self-efficacy – Do I have what it takes?

Once this model is introduced I will model how to break a problem down into smaller parts and how to plan out what strategies will help you solve a problem

For the remainder of the 18 min training time, students will practice dissecting problems and developing an approach to solving them.

Session Three

3. **During session three, we will discuss the action phase. Specifically, we will discuss monitoring in more depth and explore its link to self-reflection. These strategies will help students generate accurate internal feedback that will eventually be linked to adaptive regulatory actions.**

    a. Action

        - Giving effort
        - Using a strategy to help you solve a problem
        - Monitoring your performance
            a. Writing it down – recording
        - Maintaining focus/attention

    b. Monitoring – tracking your performance
        - Allows you to judge your understanding of math problem/concept
        - Control over the learning process
        - Building a bridge between last problem and the next one you solve
        1. Help isolate errors

    c. If you do not monitor and judge accurately, you will not know
        - Why you were successful/unsuccessful at solving a problem
        - How to change your strategy or plan to work better

    d. Overconfidence
        - Inaccurate monitoring of math knowledge
        1. Creates a lack of motivation to study for the next quiz
        2. Even though a strategy isn't working, you don't change it

    e. Under confidence
        - Even though you know the material, you don't think you do
        1. Staying up all night studying
        2. Anxiety can interfere with your thinking

    f. Monitoring leads to reflection

        - Thinking about ways to improve your performance next time
            a. Would any particular strategies have helped?
        - How do you feel about your performance?

For the remainder of the 18 min training time, the PI will work out examples and show how encountering problems should lead to reflection on how to change our approach using content-specific strategies currently being taught in the classroom,

    g. Example(s) of using a strategy successfully with a math problem

h. Example(s) of using a strategy unsuccessfully with a math problem

Session Four

4. During session four, we will discuss the reflection phase. Specifically, we will explore how reflection can help students modify their current approach or select more appropriate content-specific strategies in the future. Students will be encouraged to see errors as a learning opportunity and not as an indication of failure.

We will begin with an overview of monitoring and reflection and how they are related in the three stage model.

a. Examine the idea of reflection in more depth
   - Why is reflection important?
   - Helps you think about ways to improve your performance next time
     a. Focus on particular strategies, techniques, not ability
   - What led to the outcome you experienced?
     1. Focus on strategies as changeable – keys to success

b. Reflection
   - Reflecting on your performance
   - Thinking about ways to improve your performance next time
     a. Would any particular strategies have helped?

c. Errors as learning opportunities
   - Errors are inevitable, especially in math with problem-based learning
   - Great learners are not the ones that never make mistakes, but the ones that can learn from them
   - As long as you learn from these errors, they are actually a good thing

For the remainder of the 18 min training time, the PI will model math examples

a. You started trying to solve the problem and realize you cannot keep going
   i. What does this tell you?
   ii. What should you do next?
- Example(s) of using a strategy unsuccessfully with a math problem
- Example(s) of using a strategy successfully with a math problem

Session Five

5. During session five, we will review what we have learned about monitoring and reflection and practice these in the context of the current mathematical curriculum, emphasizing any content-specific strategies being taught in the classroom.

a. Monitoring – tracking your performance
   - Allows you to judge your understanding of math problem/concept
   - Control over the learning process

- Building a bring between last problem and the next one you solve
  1. Help build upon/reinforce successful strategy use
  2. Help isolate errors
- If you do not monitor and judge accurately, you will not know
  1. Why you were successful/unsuccessful at solving a problem
  2. How to change your strategy or plan to work better

b. How monitoring leads to reflection in three phase model

- Thinking about ways to improve your performance next time
    a. Would any particular strategies have helped?
- How do you feel about your performance?
    a. Why is this important?

c. For the remainder of the 18 min training time, we will work out examples and show how encountering problems should lead to monitoring of our solution processes and reflection on how to change our approach using content-specific strategies currently being taught in the classroom.

- Example(s) of using a strategy successfully with a math problem
- Example(s) of using a strategy unsuccessfully with a math problem

Appendix L

Reflection Worksheet

Think about the math questions you just completed and explain what strategies or processes you did correctly on these questions?

____

____

Think about the math questions you just completed and explain what strategies or processes went wrong on these questions?

____

____

What caused you to do well or to do poorly on the math problems you just completed? In other words, what is the main reason that you answered these problems correctly/incorrectly?

____

____

How well do you think you understand the material covered in the probability unit in your math class so far?

____

____

What concept(s) from the unit are you finding difficult to understand?

____

____

Specifically, what will you do to improve your understanding of the concept(s) you listed above? Describe an exact plan.

____

____

How satisfied are you with your performance on the math problems you completed during this session?

1	2	3	4	5	6	7	8	9	10
Not Satisfied	Somewhat Satisfied					Pretty Satisfied			Very Satisfied

References

Ainley, M., Corrigan, M., & Richardson, N. (2005). Students, tasks and emotions: Identifying the contribution of emotions to students' reading of popular culture and popular science texts. *Learning and Instruction*, *15*, 433-447.

Alexander, J: M., Carr, M., & Schwanenflugel, P. J. (1995). Development of metacognition in gifted children Directions for future research. *Developmental Review*, *15*, 1-37.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*, 586-598. doi:10.1037/0022-3514.41.3.586

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*, 213-238.

Barnett, J. E., & Hixon, J. E. (1997). Effects of grade level and subject on student test score predictions. *Journal of Educational Research, 90*, 170-174.

Bembenutty, H. (2009). Three essential components of college teaching: Achievement calibration, self-efficacy, and self-regulation. *College Student Journal*, *43*, 562-570.

Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests versus traditional review on performance and calibration. *Journal of Experimental Education*, *69*, 133-151.

Bol, L., Riggs, R., Hacker D. J., Dickerson, D., & Nunnery, J. A. (2010). The calibration accuracy of middle school students in math classes. *Journal of Research in Education, 21,* 81-96.

Bol, L., Hacker, D. J., O'Shea, P., & Allen, D. (2005). The influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *Journal of Experimental Education*, 73, 269-290.

Broden, M., Hall, R., & Mitts, B. (1971). The effect of self-recording on the classroom behavior of two eighth-grade students. *Journal of Applied Behavior Analysis, 4,* 191-199. doi:10.1901/jaba.1971.4-191

Brookhart, S. M., Andolina, M., ZuZa, M., & Furman, R. (2004). Minute math: an action research study of student self-assessment. *Educational Studies in Mathematics*, *57*, 213-227.

Buehl, M. M., & Alexander, P. A. (2002). Beliefs about schooled knowledge: Domain specific or domain general?. *Contemporary Educational Psychology*, *27*, 415-449. doi:10.1006/ceps.2001.1103

Bulu, S., & Pedersen, S. (2012). Supporting problem-solving performance in a hypermedia learning environment: The role of students' prior knowledge and metacognitive skills. *Computers in Human Behavior*, 28, 1162-1169. doi:10.1016/j.chb.2012.01.026

Butler, D.L., & Winne, P.H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*, 245-281.

Carr, M., Alexander, J., & Folds-Bennett, T. (1994). Metacognition and mathematics strategy use. *Applied Cognitive Psychology, 8*, 583-595. doi:10.1002/acp.2350080605

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ Lawrence Erlbaum Associates.

Confucius. (1955). *The sayings of Confucius* (J. R. Ware, Trans.). New York, NY: Signet

Chen, P. P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences, 14,* 77-90. doi:10.1016/j.lindif.2003.08.003

Chen, P. P., & Zimmerman, B. J. (2007). A cross-national comparison study on the accuracy of self-efficacy beliefs of middle-school mathematics students. *The Journal of Experimental Education, 75*, 221-244.

Cleary, T. J. (2006). The development and validation of the Self-Regulation Strategy Inventory—Self-Report. *Journal of School Psychology, 44,* 307-322.

Cleary, T. J., & Chen, P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, *47, 291-314.* DOI: 10.1016/j.jsp.2009.04.002

Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools, 41*, 537-550. doi:10.1002/pits.10177

Cleary, T. J., & Platten, P. (2013). Examining the correspondence between self-regulated learning and academic achievement: A case study analysis. *Education Research International,* vol. 2013, Article ID 272560, 18 pages, 2013. doi:10.1155/2013/272560

Cleary, T. J., Platten, P., & Nelson, A. (2008). Effectiveness of the self-regulation empowerment program with urban high school students. *Journal of Advanced Academics, 20*, 70-107.

Chi, M. T., Bassok, M., Lewis, M. W., & Reimann, P. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182. doi:10.1207/s15516709cog1302_1

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences*, *12*, 91-142. doi:10.1207/S15327809JLS1201_4

de Bruin, A. H., & van Gog, T. (2012). Improving self-monitoring and self-regulation: From cognitive psychology to the classroom. *Learning and Instruction*, *22*, 245-252. doi:10.1016/j.learninstruc.2012.01.003

Delclos, V. R., & Harrington, C. (1991). Effects of strategy monitoring and proactive instruction on children's problem-solving performance. *Journal of Educational Psychology*, *83*, 35-42. doi:10.1037/0022-0663.83.1.35

Dennison, R. S., Krawchuk, C. M., Howard, B. C., & Hill, L. (1996). *The development of a children's self-report measure of metacognition*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Desoete, A., Roeyers, H., & De Clercq, A. (2003). Can offline metacognition enhance mathematical problem solving?. *Journal of Educational Psychology*, *95*, 188-200. doi:10.1037/0022-0663.95.1.188

Dimmitt, C., & McCormick, C. B. (2012). Metacognition in education. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, J. Sweller (Eds.) , *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues* (pp. 157-187). Washington, DC US: American Psychological Association. doi:10.1037/13273-007

Dinsmore, D. L., & Parkinson, M. M. (2013). What are confidence judgments made of? Student' explanations for their confidence ratings and what that means for calibration. *Learning and Instruction*, *24*, 4-14, doi:10.1016/j.learninstruc.2012.06.001

Duijnhouwer, H., Prins, F. J., & Stokking, K. M. (2012). Feedback providing improvement strategies and reflection on feedback use: Effects on students' writing motivation, process, and performance. *Learning and Instruction*, *22*, 171-184. doi:10.1016/j.learninstruc.2011.10.003

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, *16*, 228-232. doi:10.1111/j.1467-8721.2007.00509.x\

Dunlosky, J. & Rawson, K. A. (2011). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction (2011)*,doi:10.1016/j.learninstruc.2011.08.003

Ewers, C. A., & Wood, N. L. (1993). Sex and ability differences in children's math self-efficacy and prediction accuracy. *Learning and Individual Differences*, *5*, 259-267. doi:10.1016/1041-6080(93)90006-E

Fortunato, I., Hecht, D., Tittle, C. K., & Alvarez, L. (1991). Metacognition and problem solving. *Arithmetic Teacher, 38*(4), 38-40.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1985). *Cognitive development*. Englewood Cliffs, NJ: Prentice-Hall.

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003). Enhancing third-grade student' mathematical problem solving with self-regulated learning strategies. *Journal of Educational Psychology, 95*, 306-315. doi:10.1037/0022-0663.95.2.306

Garavalia, L. S., & Gredler, M. E. (2002). An exploratory study of academic goal setting, achievement calibration and self-regulated learning. *Journal of Instructional Psychology, 29*, 221-230.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 702-718. doi:10.1037/0278-7393.11.1-4.702

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition, 15*, 84-93.

Glenberg, A. M., Wilkinson, A. C., and Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition, 10***,** 597–602.

Ghatala, E. S., Levin, J. R., Pressley, M., & Goodwin, D. (1986). A componential analysis of the effects of derived and supplied strategy-utility information on children's strategy selections. *Journal of Experimental Child Psychology*, *41*, 76-92. doi:10.1016/0022-0965(86)90052-4

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506-528. doi:10.1037/0033-295X.98.4.506

Gourgey, A.F. (2001). Metacognition in basic skills instruction. In H.J. Hartman (ed.) *Metacognition in learning and instruction* (pp.17–32). Dordrecht: Kluwer Academic.

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. Swanson, K. R. Harris, S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323-344). New York, NY US: Guilford Press.

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology, 85,* 541-553.

Hacker, D. J., Bol, L., Bahbahani, K. (2008). Explaining calibration in classroom contexts: The effects of incentives, reflection, and attributional style. *Metacognition and Learning, 3*, 101-121.

Hacker, D. J., Bol, L., Horgan, D. & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, *92,* 160-170.

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory*. (pp. 429-455). New York, NY US: Psychology Press.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (1998). *Metacognition in educational theory and practice*. Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Harackiewicz, J. M., Barron, K. E., & Elliot, A. J. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist*, *33*, 1–21.

Harackiewicz, J. M., & Linnenbrink, E. A. (2005). Multiple Achievement Goals and Multiple Pathways for Learning: The Agenda and Impact of Paul R. Pintrich. *Educational Psychologist, 40*, 75-84. doi:10.1207/s15326985ep4002_2

Harris, K. R., Santangelo, T., & Graham, S. (2008). Self-regulated strategy development in writing: Going beyond NLEs to a more balanced approach. *Instructional Science, 36*, 395-408.

Hattie, J. (2012). Calibration and confidence: Where to next? *Learning and Instruction, 24*, 62-66. doi:10.1016/j.learninstruc.2012.05.009

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66,* 99-136. doi:10.2307/1170605

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi:10.3102/003465430298487

Howard, B. C., McGee, S., Shia, R., & Hong, N. S. (2000). *Metacognitive self-regulation and problem-solving: Expanding the theory base through factor analysis*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Howard, B. C., McGee, S., Shia, R., & Hong, N. S. (2001a). *Computer-based science inquiry: How components of metacognitive self-regulation affect problem-solving*. Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Howard, B. C., McGee, S., Shia, R., & Hong, N. S. (2001b). *The influence of metacognitive self-regulation and ability levels on problem solving*. Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction, 20,* 18-29. doi:10.1016/j.learninstruc.2008.12.001

Huff, J.D., & Nietfeld, J.L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition Learning*, *4*, 161-176.

Jonsson, A., & Allwood, C. (2003). Stability and variability in the realism of confidence judgments over time, content domain, and gender. *Personality and Individual Differences, 34*, 559-574. doi:10.1016/S0191-8869(02)00028-4

Kaufman, A. S., & Kaufman, N. L. (2005). *Essentials of research design and methodology*. Hoboken, NJ US: John Wiley and Sons, Inc.

Keren, G. (1991). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica*, *77*, 217-273.

King, J. R., Biggs, S., & Lipsky, S. (1984). Students' self-questioning and summarizing as reading study strategies. *Journal of Reading Behavior, 16*, 205-218.

Kitsantas, A. (2002). Test preparation and performance: A self-regulatory analysis. *Journal of Experimental Education*, *70*, 101-113. doi:10.1080/00220970209599501

Kitsantas, A., Zimmerman, B.J. (2006). Enhancing self-regulation of practice: The influence of graphing and self-evaluative standards. *Metacognition and Learning*, *3*, 201-212.

Koku, P., & Qureshi, A. (2004). Overconfidence and the performance of business students on examinations. *Journal of Education for Business*, *79*, 217-225.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349-370. doi:10.1037/0096-3445.126.4.349

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 107-118. doi:10.1037/0278 7393.6.2.107

Kroll, M. D., & Ford, M. L. (1992). The illusion of knowing, error detection, and motivational orientations. *Contemporary Educational Psychology*, *17*, 371-378. doi:10.1016/0361-476X(92)90075-A

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121-1134. doi:10.1037/0022-3514.77.6.1121

Labuhn, A. S., Zimmerman, B. J., & Hasselhorn, M. (2010). Enhancing students’ self-regulation and mathematics performance: The influence of feedback and self-evaluative standards. *Metacognition and Learning, 5*, 173-194.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know?. *Organizational Behavior & Human Performance*, *20*, 159-183. doi:10.1016/0030-5073(77)90001-0

Lin, L., Moore, D., & Zabrucky, K. M. (2001). An assessment of students' calibration of comprehension and calibration of performance using multiple measures. *Reading Psychology, 22*, 111-128.

Lin, L., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology, 23*, 345-391. doi:10.1006/ceps.1998.0972

Linnenbrink, E. A., & Pintrich, P. R. (2000). Multiple pathways to learning and achievement: The role of goal orientation in fostering adaptive motivation, affect, and cognition. In C. Sansone, J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 195-227). San Diego, CA US: Academic Press. doi:10.1016/B978-012619070-0/50030-1

Locke, E. A. & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*, 705-717.

Lodewky, K. R., Winne, P. H., & Jamieson-Noel, D. L. (2009) Implications of task structure on self-regulated learning and achievement. *Educational Psychology*, *29*, 1-25.

Lundeberg, M. A., & Fox, P. W. (1991). Do laboratory findings on test expectancy generalize to classroom outcomes?. *Review of Educational Research*, *61*, 94-106. doi:10.2307/1170668

Maki, R. H. (1998). Test predictions over text material. In, D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice*. (pp. 117-144). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 609-616. doi:10.1037/0278-7393.16.4.609

Maki, R. H., Shields, M., Wheeler, A., & Zacchilli, T. (2005). Individual Differences in Absolute and Relative Metacomprehension Accuracy. *Journal of Educational Psychology, 97*, 723-731. doi:10.1037/0022-0663.97.4.723

Maki, R. H., & Serra, M. (1992). The basis of test predictions for text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 116-126. doi:10.1037/0278-7393.18.1.116

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*, 159-163. doi:10.1111/j.1467-8721.2009.01628.x

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 437-448. doi:10.1037/a0021962

Mevarech, Z. R. & Fridkin, S. (2006). The effects of IMPROVE on mathematical knowledge, mathematical reasoning and meta-cognition. *Metacognition Learning, 1,* 85-97.

Mevarech, Z. R., & Kramarski, B. (1997). IMPROVE: A multidimensional method for teaching mathematics in heterogeneous classrooms. *American Educational Research Journal*, *34*, 365-394. doi:10.2307/1163362

Mevarech, Z. R., & Kramarski, B. (2003). The effects of metacognitive training versus worked-out examples on students' mathematical reasoning. *British Journal of Educational Psychology*, *73*, 449-471. doi:10.1348/000709903322591181

Montague, M., & Bos, C. S. (1990). Cognitive and metacognitive characteristics of eighth grade students' mathematical problem solving. *Learning and Individual Differences*, *2*, 371-388. doi:10.1016/1041-6080(90)90012-6

Moore, D., Lin-Agler, L., & Zabrucky, K. M. (2005). A Source of Metacomprehension Inaccuracy. *Reading Psychology*, *26*, 251-265. doi:10.1080/02702710590962578

Naceur, A., & Schiefele, U. (2005). Motivation and learning - The role of interest in construction of representation of text and longterm retention: Inter- and intraindividual analyses. *European Journal of Psychology of Education*, *20*, 155-170. doi:10.1007/BF03173505

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The psychology of learning and motivation, 26*, 125-141.

Nietfeld, J.L., Cao, L., & Osborne, J.W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *The Journal of Experimental Education, 74*, 7 - 28.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning*, *1,* 159-179.

Norwich, B. (1987). Self-efficacy and mathematics achievement: A study of their relation. *Journal of Educational Psychology, 79,* 384-387. doi:10.1037/0022-0663.79.4.384

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*, 124-139. doi:10.1006/ceps.1998.0991

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*, 426-443. doi:10.1006/ceps.1995.1029

Pajares, F., & Miller, M. (1997). Mathematics self-efficacy and mathematical problem solving: Implications of using different forms of assessment. *Journal of Experimental Education, 65,* 213-228. doi:10.1080/00220973.1997.9943455

Pelham, B. W., & Blanton, H. B. (2013). *Conducting research in psychology: Measuring the weight of smoke*. Belmont, CA US: Wadsworth, Cengage Learning

Perels, F., Gürtler, T., & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. *Learning and Instruction*, *15*, 123-139. doi:10.1016/j.learninstruc.2005.04.010

Perels, F., Dignath, C., & Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education*, *24*, 17-31. doi:10.1007/BF03173472

Pintrich, P. R. (2000a). The role of goal orientation and self-regulation of learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp.451-502). San Diego: Academic Press.

Pintrich, P. R. (2000b). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92,* 544-555.

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82,* 33-40. doi:10.1037/0022-0663.82.1.33

Prenzel, M. (1992). The selective persistence of interest. In K. Renninger, S. Hidi, A. Krapp (Eds.), *The role of interest in learning and development* (pp. 71-98). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Pressley, M. (1977). Imagery and children's learning: Putting the picture in developmental perspective. *Review of Educational Research*, *47*, 586-622.

Pressley, M., Ghatala, E. S., Woloshyn, V. E., & Pirie, J. (1990). Sometimes adults miss the main ideas and do not realize it: Confidence in responses to short-answer and multiple-choice comprehension questions. *Reading Research Quarterly*, *25,* 232-249. doi:10.2307/748004

Pressley, M., & Harris, K. R. (2006). Cognitive strategies instruction: From basic research to classroom instruction. In P. A. Alexander, P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 265-286). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Pressley, M., & Levin, J. R. (1977). Task parameters affecting the efficacy of a visual imagery learning strategy in younger and older children. *Journal of Experimental Child Psychology*, 24, 53-59.

Ramdass, D. (2008). *Improving fifth grade students' mathematics self-efficacy calibration and performance through self-regulation training.* (Doctoral Dissertation). Retrieved from Digital Dissertations. (3365821).

Ramdass, D., & Zimmerman, B. J. (2008). Effects of self-correction strategy training on middle school students' self-efficacy, self-evaluation, and mathematics division learning. *Journal of Advanced Academics*, *20*, 18-41.

Rawson, K., & Dunlosky, J. (2007). Improving students' self-evaluation of learning for key concepts in textbook materials. *European Journal of Cognitive Psychology, 19*, 559-579. doi:10.1080/09541440701326022.

Redford, J. S., Thiede, K. W., Wiley, J., & Griffin, T. D. (2012). Concept mapping improves metacomprehension accuracy among 7th graders. *Learning and Instruction*, *22*, 262-270. doi:10.1016/j.learninstruc.2011.10.007

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin, 137*, 131-148. doi:10.1037/a0021705

Sailor, W. (2009). *Making RTI work: How smart schools are reforming education through schoolwide response-to-intervention.* San Francisco, CA: Jossey-Bass.

Schoenfeld, A. H. (1985). *Mathematical problem solving.* New York: Academic Press.

Schommer, M., & Surber, J. R. (1986). Comprehension-monitoring failure in skilled adult readers. *Journal of Educational Psychology, 78*, 353-357. doi:10.1037/0022-0663.78.5.353

Schmitz, B., & Perels, F. (2011). Self-monitoring of self-regulation during math homework behaviour using standardized diaries. *Metacognition and Learning, 6*, 255-273. doi:10.1007/s11409-011-9076-6

Schraw, G. (1994). The effect of metacognitive knowledge on local and global monitoring. *Contemporary Educational Psychology*, *19*, 143-154. doi:10.1006/ceps.1994.1013

Schraw, G. (2001). Promoting general metacognitive awareness. In H. J. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research and practice* (Vol. 19, pp. 3-16). Dordrecht, Netherlands: Kluwer Academic Publishers.

Schraw, G. (2009). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 415-429). New York, NY US: Routledge/Taylor & Francis Group.

Schraw, G. (1997). The effect of generalized metacognitive knowledge on test performance and confidence judgments. *Journal of Experimental Education*, *65*, 135-146. doi:10.1080/00220973.1997.994378

Schraw, G., & Dennison (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460-475.

Schraw, G., Dunkle, M. E., Bendixen, L. D., & Roedel, T. (1995). Does a general monitoring skill exist?. *Journal of Educational Psychology, 87*, 433-444. doi:10.1037/0022-0663.87.3.433

Schraw, G., Potenza, M. T., & Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology*, *18*, 455-463. doi:10.1006/ceps.1993.1034

Schunk, D. H. (1982). Verbal self-regulation as a facilitator of children's achievement and self-efficacy. *Human learning*, *1*, 265–277.

Schunk, D. H. (1983). Developing children's self-efficacy and skills: The roles of social comparative information and goal setting. *Contemporary Educational Psychology, 8*, 76-86. doi:10.1016/0361-476X(83)90036-X

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal*, *33*, 359-382. doi:10.2307/1163289

Schunk, D. H., & Ertmer, P. A. (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 US: Academic Press. doi:10.1016/B978-012109890-2/50048-2

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology*, *77*, 313-322. doi:10.1037/0022-0663.77.3.313

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). *Motivation in education: Theory, research and applications* (3rd ed.). Upper Saddle River, NJ: Merrill Prentice-Hall.

Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, *18*, 337-354. doi:10.1006/ceps.1993.1024

Schunk, D. H., & Zimmerman, B. J. (Eds.). (1998). *Self-regulated learning: From teaching to self-reflective practice*. New York: Guilford.

Schunk, D. H., & Zimmerman, B. J. (2006). Competence and control beliefs: Distinguishing the means and ends. In P. A. Alexander, P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 349-367). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly, 23,* 7-25. doi:10.1080/10573560600837578

Silver. E. A. (1979). Student perceptions of relatedness among mathematical verbal problems. *Journal for Research in Mathematics Education, 10,* 195-210.

Silver, E. A. (1981). Recall of mathematical problem information: Solving related problems. *Journal for Research in Mathematics Education, 12*, 54-64.

Spates, C. R., & Kanfer, F. H. (1977). Self-monitoring, self-evaluation, and self-reinforcement in children's learning: A test of a multistage self-regulation model. *Behavior Therapy, 8,* 9-16. doi:10.1016/S0005-7894(77)80115-9

Sperling, R. A., Howard, B. C., Miller, L., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*, 51-79. doi:10.1006/ceps.2001.1091

Stahl, E., Pieschl, S., & Bromme, R. (2006). Task complexity, epistemological beliefs, and metacognitive calibration: An exploratory study. *Journal of Educational Computing Research, 35*, 319-338.

Stoeger, H., & Ziegler, A. (2008). Evaluation of a classroom based training to improve self-regulation in time management tasks during homework activities with fourth graders. *Metacognition and Learning, 3,* 207–230.

Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12,* 437-475. doi:10.1023/A:1009084430926

Swanson, H. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology, 82,* 306-314. doi:10.1037/0022-0663.82.2.306

Thiede, K. W., & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28*, 129-160.

Thiede, K. W., Anderson, M. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*, 66-73. doi:10.1037/0022-0663.95.1.66

Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M. C. M. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes*, *47*, 331-362. doi:10.1080/01638530902959927.

Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky, A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 85-106). New York, NY US: Routledge/Taylor & Francis Group.

Thomas, A., & McDaniel, M. A. (2007). The negative cascade of incongruent generative study--test processing in memory and metacomprehension. *Memory & Cognition*, *35*, 668-678. doi:10.3758/BF03193305

U.S. Department of Education (2003) Indentifying and implementing educational practices supported by rigorous evidence: A user friendly guide. *Ed.gov. Retrieved* January, 19, 2014 from http://www2.ed.gov/rschstat/research/pubs/rigorousevid/rigorousevid.pdf

van Loon, M. H., de Bruin, A. H., van Gog, T., & van Merriënboer, J. G. (2013). Activation of inaccurate prior knowledge affects primary-school students' metacognitive judgments and calibration. *Learning and Instruction*, *24*, 15-25. doi:10.1016/j.learninstruc.2012.08.005

Vermeer, H. J., Boekaerts, M., & Seegers, G. (2000). Motivational and gender differences: Sixth-grade students' mathematical problem-solving behavior. *Journal of Educational Psychology,* 92, 308-315. doi:10.1037/0022-0663.92.2.308

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.

Whimbey, A.. & Lochhead, J. (1986). *Problem solving and comprehension.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman and D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 153-189). Mahwah, NJ: Lawrence Erlbaum Associates.

Winne, P. H. (2004). Students calibration of knowledge and learning processes: Implications for designing powerful software learning environments. *International Journal of Educational Research, 41*, 466-488.

Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice Hall.

Zabrucky, K. M., Agler, L. L., & Moore, D. (2009). Metacognition in Taiwan: Students' calibration of comprehension and performance. *International Journal of Psychology*, *44*, 305-312. doi:10.1080/00207590802315409

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*, 3-17.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner, *Handbook of self-regulation* (pp. 13-39). San Diego: Academic Press.

Zimmerman, B. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*, 166-183.

Zimmerman, B. J. (2008). Goal setting: A key proactive source of academic self-regulation. In D. H. Schunk, B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 267-295). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Zimmerman, B. J., Bonner, S., & Kovach, R. (1996). Goal 1: Understanding the principles of self-regulated learning. In, *Developing self-regulated learners: Beyond achievement to self-efficacy* (pp. 5-24). Washington, DC US: American Psychological Association. doi:10.1037/10213-001

Zimmerman, B. J., & Campillo, M. (2003). Phases and subprocesses of self-regulation. Motivating self-regulated problem solvers. In J. E. Davidson & R. J. Sternberg (Eds.), *The nature of problem solving* (pp. 239). New York, NY: Cambridge University Press.

Zimmerman, B. J., & Labuhn, A. (2012). Self-regulation of learning: Process approaches to personal development. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, J. Sweller (Eds.), *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues* (pp. 399-425). Washington, DC US: American Psychological Association. doi:10.1037/13273-014

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299-315). New York, NY US: Routledge/Taylor & Francis Group.

Zimmerman, B. J., Moylan, A., Hudesman, J., White, N., & Flugman, B. (2011). Enhancing self reflection and mathematics achievement of at-risk urban technical college students. *Psychological Test and Assessment Modeling*, *53*, 141-160.

Zohar, A., & Peled, B. (2008). The effects of explicit teaching of metastrategic knowledge on low- and high-achieving students. *Learning and Instruction*, *18*, 337-353. doi:10.1016/j.learninstruc.2007.07.001