

IMPROVING FIFTH GRADE STUDENTS' MATHEMATICS SELF-EFFICACY  
CALIBRATION AND PERFORMANCE THROUGH SELF-REGULATION  
TRAINING

by

Darshanand H Ramdass

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

2009

© 2009

DARSHANAND H RAMDASS

All Rights Reserved

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

3/27/2009

Date

Dr. Barry J. Zimmerman

---

Chair of the Examining Committee

3/27/2009

Date

Dr. Mary Kopala

---

Executive Officer

Dr. Peggy Chen

---

Dr. Bruce Homer

---

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

## Abstract

IMPROVING FIFTH GRADE STUDENTS' MATHEMATICS SELF-EFFICACY  
CALIBRATION AND PERFORMANCE THROUGH SELF-REGULATION  
TRAINING

by

Darshanand H Ramdass

Advisor: Professor Barry J. Zimmerman

This primary goal of this study was to investigate the effects of strategy training and self-reflection, two subprocesses of Zimmerman's cyclical model of self-regulation, on fifth grade students' mathematics performance, self-efficacy, self-evaluation, and calibration measures of self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, time, and strategy use. The participants were 88 fifth graders and the task involved subtraction fraction problems. Students were randomly assigned to one of four groups, strategy training and self-reflection training, strategy training only, self-reflection training only, and the control group.

A multivariate analysis of covariance showed significant main effects of strategy training with a large effect size. Follow-up univariate analyses of variance on each of the nine dependent measures revealed significant main effects for eight variables with the exception of self-efficacy accuracy. The effects sizes for these significant effects ranged from medium to large. Path analysis results also indicated that strategy training had direct and indirect effects on math performance. Self-efficacy bias was a mediating variable between strategy training and math performance. A second path analysis showed the effects of self-reflection training were weak, and it was not possible to confirm the

mediational role of math performance on self-evaluation accuracy. Correlation analyses indicated that all the variables correlated with math performance. Self-efficacy bias and self-evaluation bias correlated negatively with math performance.

However, the multivariate analysis of covariance did not reveal significant main effects for self-reflection training nor was there any interaction between strategy and self-reflection training. Nonetheless, upon examining the means scores of the self-reflection and the non self-reflection groups, a consistent pattern emerged. The mean scores for the self-reflection groups were higher than the non self-reflection groups. To explore whether this difference was statistically significant, a nonparametric chi-square analysis was conducted. The results of this test showed that the self-reflection training exerted a consistent, albeit a weak influence, indicating that it resulted in better math performance, self-evaluative judgments, and calibration judgments.

Overall, this research demonstrated that strategy training improved fifth grade students' math performance, self-evaluative judgments, and calibration measures of accuracy and less bias. The educational implications of the findings for educators were considered.

## Acknowledgements

I consider myself very fortunate in having Dr. Barry Zimmerman, an outstanding researcher and a remarkable educator, as my mentor and advisor for the past five years. This dissertation is a culmination of his immense help and guidance.

I am also indebted to my committee members, Dr. Peggy Chen and Dr. Bruce Homer, Dr. Shirley Feldmann, and Dr. John Hudesman for their time, effort, and extensive feedback in this project. I am also thankful to a number of professors at the Graduate Center who have helped me in manifold ways over the years notably, Dr Linnea Ehri, Dr. David Rindskopf, Dr. Carol Tittle, Dr. Alan Gross, and Dr. Mary Kopala.

This project would have never materialized without the help of Ms. Marie White, a professor at Nyack College. I am indeed grateful to her for suggestions, encouragement, and insightful feedback. Additionally, she helped in contacting various schools where I conducted this study.

I am also grateful to Ms. Marilyn Johansson, Mrs. Michelle Bergman, Mrs. Bharati Ramsamooj, Ms. Naomi Woodmansee and Ms. Carolyn Marko for allowing me to conduct this study at their schools. I extend heartfelt gratitude to the parents who allowed their children to participate voluntarily in this project. Without their help, this dissertation would have been impossible.

Finally, I would like to express special thanks and gratitude to my parents and siblings for this constant love and support.

THANK YOU ALL VERY MUCH

## TABLE OF CONTENTS

Copyright Page	ii
Approval Page	iii
Abstract	iv
Acknowledgements	vi
Table of Contents	vii
List of Tables	x
List of Figures	xi
Chapter	
1 INTRODUCTION	1
2 LITERATURE REVIEW	12
Self-Efficacy Theory	12
Definition of Self-Efficacy	12
Self-Efficacy Dimensions	13
Measuring Self-Efficacy	14
Self-Efficacy Scale	15
Self-Efficacy and Related Self Constructs	16
Self-Efficacy for Performance and Learning	17
Self-Efficacy and Self-Regulation	19
A Cyclical View of Self-Regulation	20
Early Self-Efficacy Studies	21
Self-Efficacy Studies in Educational Contexts	22
Self-Efficacy and Calibration	28
Measuring Calibration	29
Importance of Calibration	33
Research on Calibration	33
Synthesis	41
Rationale for Current Study	46
Hypotheses	49
3 METHODOLOGY	51
Participants	51
Task	52
Measures	52

	Self-Efficacy	52
	Mathematics Score	53
	Self-Efficacy Bias	53
	Self-Efficacy Accuracy	54
	Self-Evaluation	54
	Self-Evaluation Bias	54
	Self-Evaluation Accuracy	55
	Time on Task	55
	Strategy Use	55
	Research Design	55
	Procedures	55
	Data Analyses	58
4	RESULTS	60
	Descriptive Statistics	60
	Multivariate Analysis	60
	Univariate Analyses	62
	Correlational Analyses	63
	Supplementary Path Analysis	66
	Supplementary Chi-Square Analysis	70
5	DISCUSSION	71
	General Discussion	71
	Hypotheses	72
	Hypothesis 1	72
	Hypothesis 2	75
	Hypothesis 3	76
	Hypothesis 4	77
	Hypothesis 5	77
	Hypothesis 6	78
	Limitations	78
	Educational Implications	79
APPENDICES		
A	Pretest Items of Pilot Study	82
B	Booklet with Pretest, Training, Practice, & Posttest	84
C	Strategy Training Steps	97
D	Self-Reflection Training Methods	101

E Letter to School Principal	102
F Parental/Guardian Consent Form	103
G Student Assent Form	104
REFERENCES	105

## LIST OF TABLES

## Table

1	Means and standard deviations for self-efficacy, self-evaluation, math performance, self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, strategy use, and time on task	61
2	Univariate test results for each dependent measure	63
3	Intercorrelations between dependent measures	65
4	Decomposition of effects from the path analysis (strategy training)	68
5	Decomposition of effects from the path analysis (self-reflection training)	70

## LIST OF FIGURES

## Figure

1	Cyclical Phases and Subprocesses of Self-Regulation	7
2	Proposed Path Model from Strategy Training to Math Performance	66
3	Proposed Path Model from Self-Reflection to Math Performance	67
4	Final Reduced Path Model from Strategy Training	68
5	Final Reduced Path Model from Self-Reflection	69

## Chapter 1

### INTRODUCTION

This chapter begins with an overview of American students' mathematics performance based on national and international assessments. It is followed with a brief description of research evidence on self-efficacy, calibration, and a cyclical perspective of self-regulation indicating how these factors are critical to academic achievement and success. Finally, an intervention based on self-regulation training (strategy training, and self-reflection) is proposed, with the goal of improving fifth grade students' mathematical skills and self-efficacy calibration beliefs.

#### *American Students' Mathematics Achievement*

Mathematics is highly valued in the academic curriculum. Math skill is often used for entrance into special programs in college and is critical for students pursuing careers in technology and science (Pajares & Graham, 1999). Students who are unsure of their capability to perform in mathematics may refrain from choosing mathematics-related majors and careers. Moreover, if fewer American students enter science and technology fields due to poor performance in mathematics, it may continue to reduce America's technological edge over other countries.

The 2007 National Assessment of Educational Progress (NAEP) report showed a significant improvement in American students' mathematics scores at both the fourth and eighth grade levels compared to all past assessments. Over the past 17 years, the average scores for fourth-graders and eighth-graders increased by 27 and 19 points respectively (Lee, Grigg, & Dion, 2007). Although these results indicate progress among American students, key findings from two international assessments, the Trends in International

Mathematics and Science Study (TIMSS, 2007) and the Program for International Student Assessment (PISA, 2006) continue to raise doubts about American students' mathematical performance and the nation's ability to remain economically competitive. The TIMSS report of 2007 showed that American fourth-graders exceeded the international average in mathematics, but their performance was lower than fourth-graders of eight Asian and European countries. Likewise, eighth-graders performed above the international average, but their performance was lower than eighth-graders from five Asian countries (TIMSS, 2007). The latest PISA (2006) assessment reported that American 15-year old students performed below the international average in mathematical literacy. According to PISA, mathematical literacy involves the mastery of processes, understanding of concepts, and the capacity to apply knowledge and skills in various situations.

Taken together, these findings may have negative implications and possibly deter students from choosing science and technology-related majors and careers. Pajares and Miller (1994) observed that students generally begin to avoid math courses in high school, but this avoidance has its roots in elementary or middle school. According to the National Council of Teachers of Mathematics (2008) online report, approximately 75% of U.S. fourth-graders reported liking mathematics, stating it was practical and important. The report further noted that students would remain engaged in mathematics if they continue to perceive it as interesting and comprehensible. Brown and Lent (2006) stated that adolescents' ability and achievements in a specific domain such as mathematics are likely to create interests in that domain (mathematics-related occupations) and foster a sense of self-efficacy and positive expectations.

### *Self-Efficacy*

Academic success requires more than ability and subject matter skill; it also depends on students' use of advantageous self-regulated learning processes and on their motivational beliefs to use those skills (Bandura, 1986; Schunk & Pajares, 2004; Zimmerman & Kitsantas, 2005). In Bandura's (1986) social cognitive theory, a key source of motivation is *self-efficacy* judgments, which refer to an individual's perception of his/her capability to organize and execute courses of action necessary to achieve designated levels of performances. Self-efficacy is a predictive judgment that is assessed before performing a task. It has the greatest predictive value when it is assessed microanalytically, that is, self-efficacy beliefs should closely correspond with the criterial task (Bandura, 1986). In an analysis of these microanalytic studies, correlations between self-efficacy and academic performance have ranged from  $r = .49$  to  $r = .70$  and direct effects in path analysis studies have ranged from  $\beta = .349$  to  $\beta = .545$ . Among subject areas assessed, studies involving mathematics have shown higher correlations and stronger paths compared to reading and writing (Pajares, 1996a). Self-efficacy is one of the subprocesses of the forethought phase in Zimmerman's (1998; 2000) cyclical model of self-regulation.

### *Self-Efficacy and its Relation to Calibration*

Research on self-efficacy and its influence on academic achievement have generally shown positive outcomes (Pajares, 1996a; Zimmerman, 1995), however, research findings indicate that students are inaccurate, that is, they overestimate or underestimate their ability with a bias toward overconfidence (Hacker & Bol, 2004; Hackett & Betz, 1989; Klassen, 2002; Pajares & Miller, 1994). Bandura (1986; 1997)

hypothesizes that some overestimation of capability may allow students to increase effort, persistence, select challenging tasks, and complete them successfully. However, it is unclear to what degree high self-efficacy beliefs are beneficial when performance attainments are low. Bandura (1986) accurately notes that the most functional efficacy judgments are those that slightly exceed one's actual capability.

Moreover, an important goal of instruction is to align students' self-efficacy beliefs with their actual performance. The degree of alignment between self-efficacy and actual task performance is termed calibration (Brannick, Miles, & Kisamore, 2005). *Calibration* is a metacognitive measure that assesses any disparity between one's beliefs about performing a particular task and one's actual performance (Pajares & Kranzler, 1995). Well-calibrated people are accurate in judging their capability of performing a task and, in fact, can perform at their self-predicted level. On the other hand, poorly calibrated people who judge themselves as capable of doing a task do not perform at their anticipated level of proficiency (Schunk & Pajares, 2004).

#### *Possible Reasons for Poor Calibration*

A review of the calibration literature over the past forty years reveals that people's accuracy in judging the correctness of their answers is slightly above chance, with better accuracy relating to better performance. Studies over prolonged trials to improve students' metacognitive accuracy have remained unsuccessful, suggesting that these judgments are resistant to change (Bol & Hacker, 2001; Hacker & Bol, 2004). Moreover, researchers found that low-achieving students are less accurate and have a tendency toward overconfidence while high-achieving students tend to be underconfident, but more accurate (Bol & Hacker, 2001). Among possible reasons,

Hacker and Bol (2004) have attributed the low accuracy of students to the unreliability and insensitivity to change of existing measures of calibration. Additionally, some researchers use the attribution theory to account for differences in calibration accuracy between low- and high-achieving students. They speculate that low-achieving students have a self-serving attributional style, that is, they continuously overestimate their performance to protect their self-worth as relatively good students compared to others. On the other hand, high-achieving students attribute success to their personal skills, but they underestimate because they do not want appear immodest or overconfident (Bol, Hacker, O'Shea, & Allen, 2005; Hacker & Bol, 2004).

#### *Calibration and Mathematics Achievement*

Calibration is a very important issue in educational contexts. Inaccurate perceptions of one's capability may have negative consequences. Pajares and Miller (1994) have found that avoidance of math-related courses and careers were due to inaccurate perceptions of mathematics capability, and not because of inadequate preparation or lack of mathematics skill. Students who underestimate their confidence in certain skills they possess are unlikely to engage in tasks requiring those skills, will expend less effort, and give up easily in the face of difficulties (Bandura, 1986; Pajares & Miller, 1994). Likewise, students who grossly overestimate their capabilities may attempt tasks and fail, which may decrease their motivation (Bandura, 1986; Schunk & Pajares, 2004). Pajares (1996a) suggests that interventions should not lower students' self-efficacy, but improve the accuracy of their efficacy judgments by helping them to comprehend what they know and do not know. Researchers have found that self-regulation training improved performance (Butler & Winne, 1995; Pintrich & de Groot,

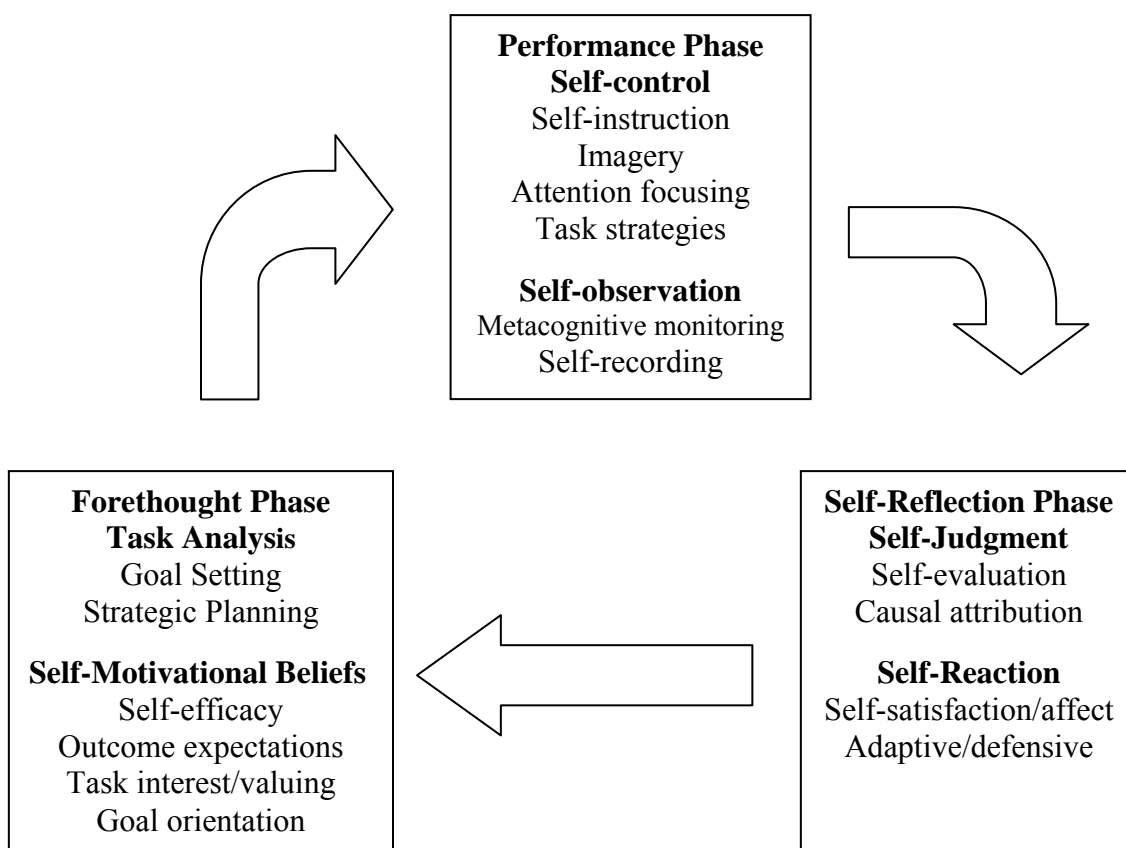
1990). Stone (2002) and Zimmerman (1990) hypothesized that self-regulated learners may be well calibrated. However, research evidence documenting this relationship is lacking.

#### *A Cyclical Model of Self-Regulation*

*Self-regulation* involves self-generated thoughts, feelings, and actions that are planned and cyclically adapted to accomplish learning goals. It is a cyclical process because feedback from previous performance is used to make improvements during current learning (Zimmerman, 2000). The cyclical nature of self-regulation is captured in Zimmerman's (1998; 2000) three phase model (see Figure 1). The *forethought phase* encompasses critical cognitive and affective processes and sets the stage for the performance phase. *The performance (volition control) phase* is comprised of processes that occur during learning and influences attention and action. The *self-reflection phase* involves processes after performance and affects one's reaction to that experience. Finally, self-reflections influence forethought and subsequent performance efforts, thereby completing the cyclical process (Zimmerman, 1998; 2000). Bandura (1986) remarks that self-regulation is not attained through will power, however, it operates through a set of subprocesses that must be developed for self-directed change.

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.

### *Why Study Mathematics*

A recent report by the National Mathematics Advisory Panel (2008) suggested that benchmarks are critical for the foundations in mathematics. These benchmarks were fluency with whole numbers, fluency with fractions, geometry, and measurement. Furthermore, the report stated that “difficulty with fractions is pervasive and it is an

obstacle to further progress in mathematics and other domains dependent on mathematics including algebra” (p. 28). The Panel’s report (2008) stated that there are few experimental studies in education. Experimental studies are necessary to provide answers of cause and effect. As a result, they recommended that small-scale experimental studies on the basic science of learning mathematics are needed to ensure the coherent growth of research in mathematics education.

An attractive feature of mathematics as a subject matter to study calibration is that it can be measured quickly and objectively. In a math task such as division or subtraction, students can be taught to assess their performance and determine whether they are making progress. Furthermore, in math, it is possible to briefly show students the actual problem, allow them to rate their efficacy to solve it, and then complete the problem (Klassen, 2002). As a result, the task analysis aspect of self-efficacy beliefs is much clearer for students doing a specific math task compared to a writing task.

Moreover, researchers have found that self-efficacy predicts mathematics performance better than math anxiety and previous math experience (Pajares & Miller, 1994). In a review article, Schunk and colleagues reported that intervention studies have improved mathematical skills and self-efficacy beliefs of struggling elementary students (Schunk & Ertmer, 2000). The National Mathematics Advisory Panel’s (2008) report remarked that children’s goals and beliefs about learning are critical in understanding how they gain proficiency in mathematics more than knowledge about how they learn the content matter. Therefore, a vital question is what leads to errors in solving mathematics problems and self-efficacy judgments? This proposed study will address this issue by

employing three subprocesses of Zimmerman's cyclical phase of self-regulation (see Figure 1).

### *Possible Sources of Errors and Self-Regulation Training*

According to this model, one potential source of errors stems from an inability to analyze key task elements accurately. If a learner does not understand what is necessary in solving a specific problem, he or she will make inaccurate self-efficacy judgments to perform the task, usually by overestimating success (Bandura, 1997). In Zimmerman's (1998; 2000) cyclical phases of self-regulation, self-efficacy occurs during the forethought phase, which is prior to the performance phase and self-reflection phase. Self-efficacy is assessing the task such as a math subtraction problem, and predicting how one will perform. Other components of the forethought phase are goal setting and strategic planning. *Goal setting* involves choosing explicit outcomes of learning such as, solving a set of fraction problems in mathematics (Locke & Latham, 1990). *Strategic planning* refers to the selection or application of learning strategies to attain specific goals (Zimmerman & Martinez-Pons, 1992). For successful performance, students have to employ strategies that are appropriate to the task. Schunk and Pajares (2004) suggest that interventions improving comprehension of task analysis and assessing one's capabilities may improve calibration. Faulty task analysis during the forethought phase can lead to inaccurate estimates of self-efficacy and poor academic functioning during the performance phase (Bandura & Schunk, 1981; Borkowski, 1992). That is, students who experience trouble in recognizing what the task entails may not employ the appropriate solution strategy.

As a result, inappropriate application of strategies is a second possible source of poor calibration. In Zimmerman's (2000) cyclical model, *strategy* is a key component of the performance phase. Schunk and Gunn (2001) define *strategy* as a systematic approach in problem solving, which involves procedural knowledge such as, how to apply algorithms in mathematics. Students who choose to implement incorrect strategies are likely to overestimate their self-efficacy judgments on future problems.

A third possible source of inaccurate estimates of self-efficacy involves poor self-reflection. *Self-reflection* occurs after doing a task and it involves comparing one's performance to a set standard (Zimmerman 1995; 2002). Self-reflection enables students to evaluate their performance, and alter their strategies if the desired outcome is not achieved (Bandura, 1986). Teaching students to self-reflect on their performance would enable them to locate the source of errors in their strategy use. Zimmerman and Kitsantas (1997) found that self-reflection training influenced students' intrinsic interest in a task and self-efficacy beliefs regarding future learning. Students who self-reflect and locate the source of errors are likely to have more accurate self-efficacy judgments compared to the students in a control condition regarding future tasks. In turn, this would influence the subsequent performance phase, thereby completing the cyclical phase of self-regulation.

Based on the above discussion, this study proposes to answer the following research questions.

1. Are inaccurate perceptions of self-efficacy and poor mathematics performance due to poor strategy choice or inadequate implementation of appropriate strategies?

2. Are inaccurate perceptions of self-efficacy and poor mathematics performance due to poor self-reflection of one's performance?
3. Would students who over-estimate their capability spend less time in solving math problems?
4. Would students who over-estimate their capability avoid using solution strategies?

## Chapter 2

### LITERATURE REVIEW

This chapter has four subdivisions. The first subdivision begins with a synopsis of self-efficacy theory, its definition, and dimensions. Next is a description of the optimal method of measuring self-efficacy, how it differs from other self-constructs, and the distinction between self-efficacy for performance and self-efficacy for learning. The second subsection discusses the relationship between self-efficacy and self-regulation focusing on Zimmerman's cyclical model of self-regulation. This is followed by a description of research studies exploring how self-efficacy influences various self-regulatory processes with a focus on academic achievement. In subdivision three, the relationship between self-efficacy and calibration is explored with a definition of calibration and how it is measured. Studies in calibration and its influence on academic achievement with a primary focus on mathematics are discussed. Finally, the fourth subdivision synthesizes the research in this chapter on self-efficacy, calibration, mathematics, and provides a rationale for this proposed study.

#### *A Synopsis of Self-Efficacy Theory*

This section defines self-efficacy and its dimensions. Next, self-efficacy measurement and scale are described. Finally, its relationship to other self-constructs are differentiated as well as the contrast between self-efficacy for performance and self-efficacy for learning.

#### *Definition of Self-Efficacy*

In his seminal article, Bandura (1977) describes *self-efficacy* as the belief or conviction that one can successfully organize and execute the behaviors or courses of

action necessary to attain specific outcomes. Although, optimal functioning requires skills and self-efficacy beliefs, self-efficacy functions independent of skills. It is related to the judgments of what an individual can do with those skills (Bandura, 1977; 1986). To discriminate between self-efficacy beliefs and ability, Collins (1982) selected children of high and low perceived self-efficacy for mathematics with different levels of math ability ranging from low to moderate to high. She found that ability was positively related to achievement, but children with high efficacy solved more math problems successfully and reworked more problems they had missed compared to students who were in the same ability group, but had low self-efficacy. Bandura (1986) postulated that perceived self-efficacy influences choice of activities, persistence, and effort. Moreover, people with stronger self-efficacy will expend more effort and persist longer at a task compared to those with weaker beliefs. Bandura (1977; 1986) mentioned that self-efficacy beliefs can be fostered through four sources, namely, (a) mastery experiences, (b) vicarious experiences, (c) verbal persuasion, and (d) physiological states, but notes that mastery experiences are the most influential. For example, as children gain mastery of a particular task, beliefs in their capabilities increase, with successes elevating efficacy beliefs and repeated failures lowering them especially if failures occur early and are not due to lack of effort or unfavorable external circumstances (Bandura, 1986; 1997).

#### *Self-Efficacy Dimensions*

Self-efficacy beliefs are multidimensional and differ in (1) level, (2) generality, and (3) strength (Bandura, 1977). *Level* refers to variations in the different level of tasks (Bandura, 1977; Zimmerman, 1995) such as, simple to moderate to difficult mathematics problems. *Generality* relates to how self-efficacy beliefs transfer across activities, such as

across different academic subjects. Schunk and Pajares (2004) stated that when different tasks require similar subskills, self-efficacy to perform or learn the subskills should predict those different outcomes. For example, writing and vocabulary skills taught in grade six are expected to generalize or transfer to grade seven. Bong (1997) also found that when underlying skills across different subject areas are related, self-efficacy beliefs generalize from one subject to the other for example, math and physics involve problem-solving skills.

Finally, the *strength* of efficacy beliefs is measured by the amount of certainty one has to perform particular tasks and it can range from low to moderate to high. Bandura (1977) advises that the best predictive judgment would emerge from a detailed assessment of the level, generality, and strength of self-efficacy beliefs that correspond to the task. It may not be possible to assess all three dimensions in a study that focuses on a particular skill for example, subtraction skills in mathematics. However, it would be possible to assess level and strength because it is within a particular subject area. The majority of self-efficacy studies have focused on these two dimensions.

### *Measuring Self-Efficacy*

Self-efficacy is a predictive construct and it is assessed before performing a task (Bandura, 1977; Zimmerman, 1995). There are three ways of assessing self-efficacy beliefs. The first method is a general evaluation of self-efficacy as an omnibus measure that assesses a global sense of efficacy and does not relate to a specific task (Pajares, 1996a). Unfortunately, students cannot make a judgment in the absence of a specific task. Decontextualized self-efficacy measures cannot predict performance; self-efficacy beliefs are not personality traits, rather they are context-based perceptions of capabilities that

individuals have of their capabilities to perform certain tasks (Bandura, 1986; Klassen, 2002).

The second method is a domain-specific assessment, which requires students to rate their capability to learn a particular subject or sport such as, science, math, or basketball. This method is more predictive than omnibus measures and preferred to general academic judgments. However, it is inferior compared to task-specific judgments.

Task-specific self-efficacy judgment is the third assessment method and it has more predictive power compared to global and domain-specific assessments (Bandura, 1997; Pajares, 1996a). Task-specific judgments matched to a specific outcome convey the greatest prediction and offer the best explanations of performance outcomes (Bandura, 1986). Using task-specific self-efficacy judgments, Schunk and his colleagues conducted a number of studies in mathematics with low-performing students (Bandura & Schunk, 1981; Schunk, 1981; 1982; 1983; Schunk & Hanson, 1985; Schunk, Hanson, & Cox, 1987). Schunk and colleagues used various types of arithmetic problems (division, subtraction) to assess self-efficacy and math achievement before and after the training sessions. At the end of the sessions, students' self-efficacy to solve mathematics problems as well as their math performance improved significantly.

#### *Self-Efficacy Scale*

In measuring self-efficacy, a “can do” statement is used rather than a “will do” statement (Bandura, 1997). For example, “how confident are you in your capability that you can write a descriptive essay?” To assess an individual's confidence, a self-efficacy scale is used to measure the strength of a person's belief with tasks of varying difficulty level. Bandura (1997; 2006) proposed to use a Likert scale with 10-unit intervals from 10

to 100 or 0 to 100 with 0 = cannot do at all, 50 = moderately certain can do, and 100 = highly certain can do. Another variation of this scale is 10 = not sure, 40 = maybe, 70 = pretty sure, and 100 = real sure (Bong, 2006). Self-efficacy scales do not have a negative number because a judgment of total incapability (0) has no lower appraisal (Bandura, 1986). Over the years, researchers have used various Likert scales to measure self-efficacy ranging from 1-5 (Bong & Hocevar, 2002) and 1-8 (Pajares & Graham, 1999). Bandura (2006) has cautioned against these types of scales stating that they are less reliable and less sensitive. Pajares, Hartley, and Valiante (2001) used two Likert scales ranging from 0 -100 and 1- 6 to assess students' writing self-efficacy. They found that the traditional 0 -100 self-efficacy scale had better predictive value for students' writing grade point average in the presence of other motivational variables such as task value in writing, apprehension, and self-concept. These results strengthen Bandura's (1997) suggestion that self-efficacy scales with too few steps would not capture the fine distinctions of a person's efficacy beliefs.

#### *Self-Efficacy and Related Self-Constructs*

Self-efficacy is different from other self constructs such as self-concept and self-esteem (Bandura, 1986). Self-concept is a general perception of one's self and one's reactions to that perception. Self-concept does not focus on accomplishing a particular task (McCombs, 1989). As a result, it can be domain specific, but not task specific. It is more global and less context specific (Pajares, 1996a). On the other hand, self-esteem is defined as how people evaluate themselves (Damon & Hart, 1982). For example, self-esteem statements are "I am happy with my math abilities" or "Everything I do goes wrong in math" (Bong, 2006). These self-constructs have a common element, which is a

belief in one's perceived capabilities; however, self-efficacy is an individual's perceived capability to execute courses of action to attain specific outcomes (Bandura, 1986; Pajares, 1996a; Zimmerman, 1995). Thus, self-efficacy has unique properties. First, self-efficacy is a judgment of one's capabilities to perform certain tasks or activities and not personal qualities or traits. Second, self-efficacy beliefs are multidimensional and not a single disposition. As a result, a student may have high self-efficacy for solving physics problems, but low self-efficacy for playing the violin. Third, self-efficacy is based on a mastery criterion of learning and not a normative one. It is a measure of what people can do with the skills they possess and not how they compare to others (Zimmerman, 1995). Pajares and Miller (1994) compared the roles of self-efficacy and domain-particular self-concepts in students' academic performance using path analysis procedures. They found that math self-efficacy was more predictive of problem solving than math self-concept, perceived usefulness of mathematics, and prior experience of mathematics. This study revealed that when self-efficacy and self-concept are distinguished, self-efficacy beliefs are the major predictor of math performance (Pajares & Miller, 1994).

#### *Self-Efficacy for Performance and Learning*

Social cognitive theorists have distinguished between self-efficacy for performance and self-efficacy for learning (Pajares 1996a; Schunk et al., 1987; Zimmerman & Kitsantas, 2005). *Self-efficacy for performance* involves judgments of one's capability to solve particular problems such as math subtraction problems. *Self-efficacy for learning* involves students' beliefs that they can learn the necessary skills and strategies to solve a particular problem (Zimmerman & Kitsantas, 2005). Research evidence indicates that self-efficacy for learning mathematical fraction problems is

predictive of posttest self-efficacy for math problem-solving (Schunk, et al., 1987), and perceived efficacy for self-regulated learning is predictive of perceived efficacy to perform (Zimmerman, Bandura, & Martinez-Pons, 1992).

Schunk and Hanson (1985) found that self-efficacy for learning fostered greater learning and higher achievement. In an intervention study, these two researchers randomly assigned students to one of six experimental conditions: male mastery model, male coping model, female mastery model, female coping model, teacher model, and no model. Students did a pretest on mathematics subtraction skills and judged their self-efficacy to learn to solve various types of subtraction problems. Next, they were trained to solve the subtraction problems and afterwards had time to practice on new problems. Posttest measures were self-efficacy for learning, posttest self-efficacy, and math skill.

The results showed that observing a peer model led to higher self-efficacy for learning, posttest self-efficacy, and math achievement than observing a teacher model or no model. Students in the teacher model condition scored higher on the above measures than the control group. Although Schunk and Hanson (1985) were interested in the effects of various types of modeling behaviors, one key finding was self-efficacy for learning. The results showed that self-efficacy for learning was positively related to posttests of self-efficacy and math skill.

In sum, this subsection defined self-efficacy, its dimensions, and the best method to measure it that would improve its predictability. Compared to other self-constructs such as self-concept and self-esteem, self-efficacy has unique properties that lends itself to research with predictable outcomes. This is possible only when self-efficacy is

assessed microanalytically (Bandura, 1986), that is, self-efficacy beliefs should be assessed specific to the task that it corresponds to.

### *Self-Efficacy and Self-Regulation*

This section explains the relation between self-efficacy and self-regulation, and cites some research evidence. Next, a cyclical view of self-regulation is discussed from a social cognitive perspective. Thereafter, a brief discussion ensues of self-efficacy with reference to early self-efficacy studies, and how the construct has evolved to study psychological variables in an educational context. Finally, there is a discussion of correlational and intervention studies that were done from the cyclical perspective of self-regulation.

According to social cognitive theory, self-efficacy is a major motivational construct of self-regulation (Bandura, 1986; Zimmerman, 2001). In Bandura's (1986) social cognitive model, people's efforts to self-regulate their behavior are guided through the interdependent interactions of personal/cognitive, behavioral, and environmental processes. These three processes constantly change during learning or performance and must be monitored (Bandura, 1986; 1997). Self-regulated learners are mentally engaged during learning compared to passive learners and are aware of what they must do to accomplish a task or goal (Zimmerman, Bandura, & Martinez-Pons, 1992). Researchers have found that students with a strong sense of self-efficacy engaged in higher-levels of self-regulatory behaviors (Pintrich & de Groot, 1990), implemented cognitive and self-regulatory strategies more effectively (Pintrich & de Groot, 1990; Zimmerman et al, 1992), and attained greater academic achievement (Wilhite, 1990) compared to those with low self-efficacy.

Multon, Brown, and Lent (1991) conducted a meta-analysis of 39 self-efficacy studies done between 1977 and 1988. They found that self-efficacy beliefs were related to performance ( $r = .38$ ), persistence ( $r = .34$ ), accounted for roughly 14% of the variance in academic achievement, and approximately 12% in academic persistence. More specifically, Multon et al. found that the strongest effect sizes were obtained by researchers who measured self-efficacy judgments specific to the cognitive skills, and administered the measures simultaneously. The strongest effect sizes were basic skills measures (.52), followed by classroom-based performance such as grades (.36), and standardized achievement tests (.13). Bandura (1986; 1997) stated that when task-specific judgments of self-efficacy are administered simultaneously with the task, they have the greatest predictive power. Multon et al. (1991) reported that self-efficacy judgments for the basic skills measures of performance were specific and they were administered simultaneously with the task, which supports Bandura's hypothesis.

Another meta-analysis study done by Stajkovic and Luthans (1998) explored the relationship between self-efficacy and work-related performance. The findings revealed that the average weighted correlation between self-efficacy and work-related performance was  $(G)r = .38$ , which was 28% gain in task performance. These meta-analysis studies demonstrate that self-efficacy beliefs are critical to successful academic achievement and performance.

#### *A Cyclical View of Self-Regulation*

The school environment is based on the premise that students will eventually learn to manage academic and personal self-regulation as they progress from childhood to adolescence (Zimmerman, 2002). Initially, educators provide ample support in

elementary and middle grades, but gradually decrease support as students progress to higher grades with the expectation that they will develop self-regulatory skills (Zimmerman, 2002). There is evidence that many students adopt self-regulatory skills such as effective learning strategies, but many do not (Zimmerman & Martinez-Pons, 1990). Students who do not develop these skills are at a disadvantage. Researchers have found that training students in various self-regulatory skills can improve self-efficacy and academic performance (Schunk & Ertmer, 2000; Zimmerman, 1995). In distinguishing skillful self-regulators from naïve self-regulators, Zimmerman (1998; 2000) emphasized that skillful self-regulators set specific goals for learning outcomes and possess a strong sense of efficacy to accomplish them. While working on various tasks, they monitor their performance and compare their achievements with target goals to determine progress. Positive self-evaluations of progress lead to continued use of strategies and motivation for achievement (Zimmerman, 1998; 2000). Schunk and Ertmer (2000) emphasized that high self-efficacy for learning in the forethought phase leads to self-efficacy for continuous progress in the performance phase, and self-efficacy for achievement in the self-reflection phase. In turn, self-efficacy would influence the forethought phase, thereby completing the cycle (Schunk & Ertmer, 2000).

#### *Early Self-Efficacy Studies*

Initially, self-efficacy studies were done in the context of therapeutic behavior change, primarily in fearful situations to explain coping behaviors (Bandura, 1977; Schunk & Pajares, 2004). Participants were adults who judged their self-efficacy accurately for performing the tasks, which resulted in high correspondence between self-efficacy and actual performance (Bandura, 1977). In one study, (Bandura, 1977), adult

snake phobics were assigned to three groups, participant modeling, modeling alone, or no treatment. In participant modeling, participants were assisted to engage in increasingly more threatening interactions with a boa constrictor. The modeling alone condition observed the therapist perform the same activities done in the participant modeling condition. Participants in the control condition took part in the assessment procedures without any intervention. The level, strength, and generality of the participants' self-efficacy were measured at various points in the experiment. These measures were obtained prior to treatment, following treatment, and after completing the posttest.

The results showed that the participant modeling condition produced higher, more generalized, and stronger efficacy expectations than the modeling alone condition, which in turn performed significantly better than the control condition.

Such studies are quite a contrast with those currently encountered in educational settings where participants are young children and adolescents who are unable to make accurate appraisal of their capabilities (Schunk and Pajares, 2004). Children may have difficulty understanding the range of skills needed to solve mathematics problem, write a descriptive essay, and answer comprehension questions. Furthermore, cognitive skills are difficult to judge and sometimes it is difficult to know when a skill is increasing such as becoming a better writer or a better problem-solver (Bandura, 1986; Schunk & Pajares, 2004). According to Klassen (2007), developing accurate self-efficacy beliefs is a metacognitive ability that demands a conscious awareness of the self and the task.

#### *Self-Efficacy Studies in Educational Contexts*

Over the years, researchers have modified the self-efficacy scale to measure cognitive skill learning in academic achievement (Schunk, 1981; Zimmerman, et al.,

1992). Schunk and Ertmer (2000) reported a number of correlation studies that supports the hypothesized relationship between self-efficacy and self-regulation. Zimmerman and Martinez-Pons (1990) found that gifted students displayed higher self-efficacy and strategy use compared to normally achieving students. Pintrich and de Groot (1990) found that self-efficacy, self-regulation, and cognitive strategy were positively intercorrelated and predicted academic achievement among seventh graders in English and Science. Other researchers demonstrated that self-efficacy correlated positively with performance (Bouffard-Bouchard, Parent, & Larivee, 1991), and affected achievement directly and indirectly by influencing goals (Zimmerman & Bandura, 1994). Moreover, students with high self-efficacy displayed more effort compared to those low in self-efficacy (Tuckman & Sexton, 1990), and self-efficacy enhanced academic achievement independent of ability and skills (Collins, 1982; Pajares and Kranzler, 1995).

The above correlational results reveal substantial support for the interrelationship of self-efficacy, self-regulation, and academic achievement. Over the years, researchers have conducted intervention studies designed to enhance self-efficacy and self-regulation. Schunk and Ertmer (2000) reviewed a number of such intervention studies from the perspective of Zimmerman's (1998; 2000) cyclical model of self-regulation, that is, they focused on studies aimed at processes in the forethought, performance, and self-reflection phases and how these studies influenced self-efficacy and self-regulation.

To investigate the influence of goal setting on self-efficacy and performance, Zimmerman and Kitsantas (1996; 1997) did two studies to evaluate the effects of goals on self-efficacy during dart throwing. In the 1996 study, ninth and tenth grade girls were randomly assigned to process-goal (focus on the steps in dart throwing) and product-goal

conditions (concentrate on the scores). Additionally, some students self-recorded their correct steps and some self-recorded their scores. The results showed that the process-goal group achieved higher self-efficacy and performance compared to product-goal condition. Self-recording also improved the students' self-efficacy and performance. These results were replicated in the 1997 study where Zimmerman and Kitsantas added another condition, a shifting-goal condition in which the girls shifted from process goals to product goals after they performed the steps automatically. The results showed that the shifting condition led to the highest self-efficacy ratings and performance.

In a subsequent study, Zimmerman and Kitsantas (1999) used a writing (cognitive) task to evaluate the effects of goal setting on self-efficacy and performance. Students were taught a revision strategy to rewrite sentences. Girls in the process goal condition focused on strategic steps to revise the writing task. Those in the outcome group focused on decreasing words in the revised passage. The optimal group shifted from process to outcome goals after attaining automaticity. In addition, some of the students in each group self-recorded their performance. The results showed that girls in the shifting goal condition surpassed those in the process or outcome goal condition and evidenced superior writing skills and enhanced self-efficacy beliefs. Girls in the outcome group displayed the least writing skill; however, self-recording improved the writing for all groups.

To investigate the influence of the subprocesses of the performance phase, Schunk (1983) examined the effects of self-monitoring on self-efficacy and self-regulated learning. Participants were struggling students and the task involved subtraction problems. The students were assigned to either a self-monitoring (review work at the end

of the session), external-monitoring (an adult reviewed and recorded the number of pages completed), or no monitoring conditions. He found that the self- and external-monitoring groups evidenced higher self-efficacy, persistence, and achievement compared to the control condition. This study revealed that monitoring of progress enhanced the students' perceptions of their learning and self-efficacy for continuous progress.

Another subprocess in the performance phase is task strategy. *Task strategy* involves employing various strategies to solve a task such as a set of steps to solve a math problem, or brainstorming ideas before writing an analytical essay (Zimmerman, 1998; 2000). Schunk and Cox (1986) did an intervention study to determine the influence of strategy training and attributional feedback on students with learning disabilities. The participants were 90 middle school students. They received subtraction training for six sessions. One group of students verbalized aloud while solving problems during all the sessions. Another group verbalized during the first three sessions and the control group did not verbalize.

The first goal of the study was to decide how verbalization during cognitive-skill learning influenced students' self-efficacy and math skills. It was hypothesized that students in the two verbalization conditions would develop higher self-efficacy and math skills than the non-verbalization condition. Additionally, the researchers postulated that the two verbalization conditions would not differ, but the overt verbalization during the first three sessions would help students learn to solve the problem in a strategic manner. The second goal was to investigate how the sequence of effort-attributional feedback (linking successful problem solving with effort) would affect students' self-efficacy and math skills.

The results of the study showed that overt verbalization of the steps to solve the problems facilitated task performance, self-efficacy, and math skills. Neither conditions of continued verbalization (entire six sessions) nor discontinued verbalization (first three sessions) enhanced achievement outcomes. Schunk and Cox (1986) suggested that the discontinued verbalization group abandoned the strategic approach when they were told not to verbalize any longer and might not have internalized the strategy. There were two conditions of effort feedback. Some students received feedback during the first half of the training and some received feedback in the second half of the training. Both conditions led to higher self-efficacy and subtraction skills. Students in the first-half-effort feedback condition showed significant improvement in effort attributions, but those in the second-half-effort also showed significant gains. Schunk and Cox (1986) concluded that these students were struggling in mathematics and feedback at any point in the study boosted their performance.

Studies investigating self-evaluation have shown benefits among adults (Bandura & Cervone, 1983; Schunk & Ertmer, 1999) as well as children (Schunk, 1996). In a two-part study, Schunk (1996) investigated how self-evaluations of capabilities during cognitive skill training affect children's achievement. Participants were 44 fourth-grade students. In study one, students were randomly assigned within gender, ethnic background, and classroom to one of four experimental conditions, (a) learning goal with self-evaluations, (b) learning goal without self-evaluations, (c) performance goal with self-evaluations, and (d) performance goals without self-evaluations. The dependent measures were goal orientation, self-efficacy, mathematics skill, and persistence. These dependent measures were given to each student in the pretest and posttest. Students

received six 45-minute instructional sessions over a week and the task involved fraction skills. The results showed that the learning goal with or without self-evaluation and the performance goal with self-evaluation led to higher self-efficacy, math skill, motivation, and task orientation compared to the performance goal without self-evaluation.

In study two, 40 fourth graders from two classes in one school were assigned to a learning or performance goal condition, but all self-evaluated their performance. Self-evaluation judgments were collected once at the end of the instructional program rather than after each session, as had been done in study one. Schunk (1996) hypothesized that learning goals would lead to higher self-evaluation scores and achievement than performance goals. The measures were goal orientation, self-efficacy, skill, and persistence and they were assessed at pretest and posttest. At the end of the sixth session, self-evaluation and self-satisfaction were also assessed. The results showed that combining a learning goal with self-evaluation improved mathematics achievement more than combining a performance goal with self-evaluation. Regarding self-evaluation, these two studies showed that when students assessed their capabilities or progress in learning, they became more competent, a perception that increased self-efficacy and academic achievement (Schunk, 1996).

In a later study, Schunk and Ertmer (1999) found that undergraduates in a process goal condition self-evaluated their work more compared to those in a product goal condition. Self-evaluation also improved self-efficacy, which is consistent with Zimmerman's (1998; 2000) cyclical model that self-evaluation of one's progress influences self-efficacy.

Clearly, a review of the research indicates that self-efficacy is a critical component of self-regulatory processes. Correlational and intervention research designs (Schunk and Ertmer, 2000) revealed that self-efficacy has a positive effect on self-regulatory behaviors in accordance with Zimmerman's (1998; 2000) cyclical model. Studies on goal setting (Zimmerman and Kitsantas, 1996; 1997; 1999), self-monitoring (Schunk, 1983), strategy training (Schunk and Cox, 1986), and self-evaluation (Schunk, 1996; Schunk and Ertmer, 1999) indicated that these processes enhance self-efficacy, and academic performance, which in turn influences students' use of self-regulatory processes.

### *Self-Efficacy and Calibration*

So far, the discussion has focused on various aspects of self-efficacy and various subprocess of self-regulation and their influence on academic achievement. This section focuses on self-efficacy and its relationship to calibration. Like self-efficacy, calibration is also a metacognitive measure. After discussing various methods how researchers have measured calibration, this section describes an optimal method of measuring calibration. This is followed by a discussion on the importance of calibration in educational contexts, a chronicle of calibration research, and finally, the issues that remain unanswered in calibration studies.

A key issue in research on self-efficacy is the accuracy of these judgments. Although, Bandura (1986) did not use the term *calibration* in his writings, he mentioned that "reasonably accurate appraisal of one's own capabilities is, therefore, of considerable value in successful functioning" (p. 393). Furthermore, misjudgments of personal efficacy in either direction have detrimental consequences. People who exceedingly

overestimate their capabilities would embark on activities beyond their control and experience failures. Likewise, those who underestimate their capabilities are likely to engage in self-limiting activities, thereby, decreasing their successes (Bandura, 1986). Self-efficacy, therefore, is related to calibration in the sense that one's judgment of performing a particular task has to be accurate. Social cognitive researchers have defined *calibration* as the degree of alignment between self-efficacy judgments and actual task performance (Brannick, Miles, & Kissamore, 2005) or how well self-efficacy judgments relate to actual task performance (Pajares & Schunk, 2004). Well-calibrated people are accurate in judging their capability of performing a task and, in fact, can perform at their self-predicted level. However, people with poor-calibration judge themselves efficacious, but cannot perform a task or make inefficacious judgements, but are able to perform it (Schunk & Pajares, 2004).

### *Measuring Calibration*

Social cognitive and metacognitive researchers have measured calibration in various ways over the years. Initially, researchers calculated congruence (an accuracy index) between self-efficacy and accuracy by comparing each posttest efficacy judgment with the subsequent accuracy score on a problem of comparable form and difficulty (Bandura and Schunk, 1981; Schunk, 1981). These researchers calculated two measures of incongruence; one was *overestimation*, which referred to children judging that they could solve a certain class of problems, but failed to solve them. The second was *underestimation*, defined as children judging that they could not solve a particular type of problem, but solved it correctly.

Another group of researchers used a correlational index to measure calibration (Glenberg, Sanocki, Epstein, & Morris, 1987). They described calibration as the correlation between ratings in comprehension and actual performance on an objective test with a correlation near one indicating good calibration and a correlation near zero as poor calibration.

Another method to measure calibration was developed by Schraw, Potenza, and Nebelsick-Gullet (1993). Schraw et al. (1993) used four different measures of calibration to assess the role of feedback and incentives. The first measure was *mean bias*, which was defined as the difference between estimated and true performance (the direction of judgment error). *Mean accuracy* was the second measure, and it was used to assess the magnitude of judgment error. *Mean accuracy* is the absolute value of the bias score. The third measure was a *coefficient alpha* to assess the internal consistency of judged performance scores and finally, they used the point-biserial correlation measurement by Glenberg et al. (1987) to determine the relationship between actual performance scores and judged performance scores. Schraw et al. (1993) tested four hypotheses in this study. First, easy test questions would lead to less overconfidence and better accuracy than difficult questions. Second, the incentive to improve calibration would decrease bias and increase accuracy on both tests. Third, coefficient alpha would be very high for estimates of performance accuracy and finally, correlations between true and estimated performances would be higher in the incentive than in the control conditions. Participants were 85 undergraduates and the experimental design was a 2 (feedback vs. no feedback) × 3 (incentive for improved performance, incentive for improved calibration, and control) × 2 (easy vs. difficult test) design.

The test involved 36 reading comprehension multiple-choice questions and a test of eight math questions. In a pilot study, each test was normed on 10 students, with a performance score of 76% for the easy test (reading comprehension) and 46% for the difficult test (math test).

The results showed main effects for incentives and test variables. Those who were rewarded for improved calibration showed less underconfidence than the control group and had a greater number of correct answers than the good performance group. These results indicated that although incentives improved monitoring of performance, accurate calibration was more critical in improving performance (Schraw et al., 1993). Moreover, feedback had no influence on performance, bias or accuracy. The researchers suggested that providing an incentive to improve calibration might have facilitated self-generated cognitive feedback, that is, learners shifted their attention from performance to monitoring their performance. Regarding test difficulty, bias and accuracy scores revealed that students were more accurate on easy items rather than difficult ones, suggesting that they were less effective in judging their performance when a task was difficult.

In a subsequent article, Schraw (1995) analyzed the various measures of assessing calibration and suggested that the bias and accuracy measures described by Yates (1990) and Keren (1991) are the most informative. Bias measures over- or underconfidence on a test or task. Bias is computed by taking the mean differences between predicted and actual performance scores and can range from -1 to 1. Scores larger than zero would be overconfidence, and scores less than zero would correspond to underconfidence. Scores close to zero would indicate no bias. Accuracy can be calculated by squaring the bias

score or taking its absolute value and can range from 0 to 1. Using the absolute value, the difference between the confidence judgment and performance for each test item can be calculated. Next, a mean value can be calculated by summing the absolute value of all the items and dividing them by the total number of items (Schraw, 1995).

Social cognitive researchers (Chen, 2003; Pajares & Graham, 1999; Pajares & Miller, 1997) and metacognitive researchers (Neitfeld, Cao, & Osborne, 2005) have used the above two measures of bias and accuracy over the past decade to measure self-efficacy calibration. In both studies by Pajares and colleagues, they found that calibration bias and accuracy correlated with self-efficacy; however, (Chen, 2003) did not find a correlation, which was due to a methodological difference of assessing calibration. Chen (2003) used a separate, but parallel test to measure calibration; whereas, Pajares and colleagues used self-efficacy and performance scores to calculate calibration scores. Chen (2003) reanalyzed the data and the results showed a significant correlation between both calibration measures and self-efficacy. Chen's (2003) study explored whether seventh-grade math students' calibration formed an important aspect of their self-efficacy beliefs. In her study, calibration was not an outcome measure, but it predicted self-efficacy strength and academic performance. Path analysis results revealed that calibration accuracy exercised a significant effect on students' math self-efficacy judgments. Regression results showed that calibration accuracy increased the predicted power of self-efficacy with students' math performance improving by 40%. Chen (2003) concluded that self-efficacy and calibration considerably enhanced students' math performance and were not only mediators of previous math performance.

### *Importance of Calibration*

In educational contexts, the issue of calibration is critical. Self-efficacy researchers have maintained that optimistic self-efficacy beliefs are critical to success in completing challenging tasks (Bandura, 1997; Pajares, 1996a). However, children who overestimate their capabilities attempt tasks and fail, which may decrease motivation. Similarly, those who underestimate their capabilities avoid tasks they are capable of doing, thereby, limiting their potential growth (Schunk & Pajares, 2004). Thus, poor calibration results in students applying effort in suboptimal ways. On the other hand, well-calibrated students, who judge their capabilities accurately, recognize their strengths and limitations and therefore, know where to employ their effort gainfully (Brannick, Miles, & Kissamore, 2005). Although, Bandura (1986; 1997) posited that optimistic self-efficacy beliefs can increase perseverance, and effort, gross misjudgments about one's efficacy can be misleading and potentially harmful (Klassen, 2002). Therefore, it is imperative that students construct accurate metacognitive understanding of their capabilities.

### *Research on Calibration*

Over the past three decades, researchers have found that normally achieving students at various academic levels frequently displayed poor calibration, mainly with a bias toward overconfidence (Hackett & Betz, 1989; Hacker & Bol, 2004; Pajares & Kranzler, 1995; Pajares & Miller, 1994). In a review study, Klassen (2002) found that learning disabled students miscalibrated with a tendency toward overestimation. On the other hand, there is some evidence that gifted or high-achieving students are better calibrated than regular-education students (Pajares, 1996b), but are somewhat

underconfident (Hacker & Bol, 2004).

The early clinical studies on calibration revealed that accuracy between self-efficacy and performance was very high, and it was normal to achieve 90% or higher calibration (Bandura & Adams, 1977). Possible reasons for these high results were that the participants were adults and the tasks were observable events in a controlled laboratory setting, that is, snake phobics in one group were assisted to engage in progressively more threatening interactions with a boa constrictor (Bandura, 1977). In contrast, Schunk and Pajares (2004) reported that educational studies have reported lower calibration percentages.

In an attempt to apply self-efficacy theory to a cognitive domain, Schunk (1981) trained 56 elementary children to solve arithmetic tasks and judge themselves on a self-efficacy scale. These children performed below average in arithmetic, persistence, and had low self-efficacy. His goal was to determine whether he could train them to improve their arithmetic skills, self-efficacy, and accuracy judgments. He assigned 12 students randomly to one of four treatment conditions, namely, (a) cognitive modeling-attribution, (b) cognitive modeling-no attribution, (c) didactic attribution, (d) didactic-no attribution, and 8 students to a control group.

Schunk (1981) tested three sets of hypotheses. The first set of hypotheses explored whether modeling, guided performance, and self-directed mastery would facilitate development of arithmetic skills and self-efficacy. The second set related to the effects of effort attribution on achievement during arithmetic training. Finally, the third set of hypotheses tested the relationship of self-efficacy to subsequent achievement. Schunk postulated that modeling was an ideal intervention to improve accuracy of self-

efficacy because it focused children's attention on problem-solving and corrective strategies.

Each student participated in a pretest, a training phase, and a posttest. The pretest comprised of 18 division problems ranging in various levels of difficulty. The children also did a self-efficacy pretest to familiarize themselves with the scale. Prior to rating themselves on the scale, they received training to judge their capability using the efficacy measure to jump progressively longer distances. Schunk developed the self-efficacy scale on the recommendations of Bandura (1977; 1986). Next, each child participated in three 55-minute training sessions. During the first 10 minutes, the children received instructions on division strategies. For the next 35 minutes, they practiced the strategies. In the final 10 minutes of the study, self-directed mastery, they solved problems independently.

During the instruction and practice phase, the trainer sat with the child, but left the child alone to work during the final 10 minutes. Before the training session, each student received an instructional package with an identical format. The first two pages explained the solution strategies with systematic examples. The next several pages consisted of several problems that the children solved during the practice phase. After solving each problem, they were informed of the correctness of their solutions and were asked to check their work if any errors were observed. The final two pages consisted of the self-directed mastery problems. Cognitive modeling instruction consisted of an adult who solved division problems while verbalizing the solution strategies to get the correct answer.

During the practice phase, the children received corrective modeling when they encountered difficulties. In the didactic condition, children studied the explanatory pages individually and worked on practice problems afterwards. If they faced difficulties, the trainer told them to review relevant sections of the explanatory pages. Schunk (1981) pilot tested these explanatory pages to enhance comprehension among this group of children. In the modeling- and didactic-attribution condition, the trainer attributed children's successes to high effort ("You worked really hard on that one") and difficulties to low effort ("You need to work harder") every 5-6 minutes during the practice phase of all three sessions. The children received attributions when it seemed most appropriate and they got each type of attribution approximately 20 times to make the effort attributions salient. Effort attribution and corrective feedback were verbalized separately to avoid confusion.

Posttest assessment was done within one week upon completion of training. The procedures were identical except that the self-efficacy measures were collected prior to the posttest. A new parallel set of division problems was used for the posttest. Both forms were administered to 13 fourth graders and their reliability coefficient were highly correlated ( $r = .92$ ). Additionally, students' scores were obtained from the school district on the math portion of the Metropolitan Achievement Tests (MAT) to find out whether math ability was related to children's response to the treatment.

The posttest results showed that all treatment groups evidenced significant improvement in self-efficacy, solved more problems, and persisted longer compared to the control group. The modeling condition was superior to the didactic treatment, but modeling-attribution condition had no effect on persistence. However, there was a

relationship between self-efficacy and persistence; students who judged they could solve more problems, persisted longer on solving them.

In this study, accuracy was calculated by two measures of incongruence. One was *overestimation*, which referred to children judging that they could solve a certain class of problem, but failed to solve it. The second was *underestimation*, defined as children judging that they could not solve a particular type of problem, but solved it correctly (Schunk, 1981). To analyze accuracy judgments, seven predictors were entered in a regression equation. Five of the seven predictors, self-efficacy, persistence, pretest accuracy, modeling-didactic variable, and the Metropolitan Achievement Test (MAT) score, individually accounted for a significant proportion in the explained variance of posttest accuracy. Schunk (1981) further explored the finding that posttest accuracy was the function of multiple influences in a causal model. A path analysis model showed direct causal links between treatment and self-efficacy ( $r = .12$ ), self-efficacy and persistence ( $r = .30$ ), self-efficacy and accuracy ( $r = .57$ ), and persistence and accuracy ( $r = .44$ ). It is interesting to note the strong, positive relationship between self-efficacy and accuracy, indicating that changes in accuracy depends mainly on changes in self-efficacy. ANOVA results showed that modeling children had significantly higher congruence than didactic children group. Modeling-attribution children showed higher congruence than those not receiving attribution, thereby supporting the hypothesis that modeling children would appraise their capabilities more accurately than didactic children who tended to overestimate. These findings regarding accuracy appraisal showed that it is possible to train children to improve self-efficacy, arithmetic skills, and accuracy judgments.

Bandura and Schunk (1981) conducted another study to test whether self-motivation through proximal subgoals would be most effective in developing mathematical skills, self-efficacy, and intrinsic interest in mathematics. Additionally, they hypothesized that self-efficacy would predict subsequent accuracy on math tasks and the level of intrinsic interest.

Participants were 40 elementary grade children. They were randomly assigned to one of four groups, (a) proximal goals, (b) distal goals, (c) no goals, and (d) no treatment (control group). The pretest measure consisted of 25 subtraction problems ranging from easy to difficult. The goal of this study was to examine how mathematics competencies, perceived efficacy, and interest might be developed when they are lacking; as a result, children who solved more than four problems in the pretest were excluded. The participating students exhibited poor mathematics skills, with one third unable to solve a single problem, while another third could only solve one correctly. Children received training in making a self-efficacy judgment similar to Schunk's (1981) study mentioned earlier.

Students were tested individually. Seven sets of instructional material were designed with various subtraction operations. Each child received an instruction set of seven pages and had seven 30-minute sessions to complete the problems. The instructions in the booklet were self-explanatory. If a student asked for further instructions, the experimenter reread the relevant sections without giving further hints. Students in the proximal goal condition had to set goals for themselves, such as completing at least six pages each session. This group got the same suggestion in the second session to indicate the importance of a continuing goal orientation. In the distal goal condition, the

experimenter told the students to consider completing the entire set of 42 pages by the end of the seventh session. The no goal group pursued the self-directed learning without any mention of goals. However, they had to complete as many pages as possible. Finally, the control group got the instructional material without any interventions.

The dependent measures were self-efficacy, intrinsic interest, mathematics performance, and accuracy of self-efficacy judgments. Self-efficacy was measured at the end of the treatment and after the arithmetic posttest. Self-efficacy scores at the end of the treatment were used to predict subsequent arithmetic performance and the self-efficacy scores after the posttest were used to measure intrinsic interest. Accuracy or congruent measures were calculated by comparing efficacy judgments at the end of the treatment with posttest performance of subtraction problems of comparable form and difficulty. Congruence occurred when children judged themselves capable of doing the task and in fact did it or when they judged themselves incapable and failed to do it. Mismatches between efficacy judgments and performance indicated incongruence (Bandura & Schunk, 1981).

Results of this study showed that children in the proximal subgoals condition who set attainable subgoals achieved mastery of arithmetic skills, improved their self-efficacy, and interest in the task that held little attraction for them. Moreover, the proximal group was highly accurate in their self-appraisals of efficacy (80%); whereas, children in the distal goal (54%), no goals (51%), and control condition (60%) displayed moderate congruence between self-efficacy beliefs and performance. In most instances, children made overestimations of their capabilities (Bandura & Schunk, 1981).

The above two studies by Schunk (1981) and Bandura and Schunk (1981) showed that modeling and setting proximal goals for one's self increased self-efficacy, arithmetic skills, persistence, and accurate self-appraisal. Taken together, these studies demonstrated that it is possible to train young children in elementary grades to improve their math skills, self-efficacy beliefs, intrinsic interest, and accuracy judgments.

However, a review of the literature on calibration studies over the past three decades has revealed that students are generally inaccurate in their judgments of what they know (Hacker & Bol, 2004). Numerous studies (Hacker, Bol, Horgan, & Rakow, 2000; Pajares, 1996b) have revealed that higher achieving students were more accurate in their predictions, but underconfident. However, lower achieving students were less accurate and overconfident (Bol, Hacker, O'Shea, Allen, 2005; Hacker et al., 2000).

In a quasi-experimental study, Bol and Hacker (2001) investigated whether practice tests would improve students' prediction and postdiction accuracy, and exam performance. Participants were 59 graduate students. Exam performance was measured by scores on the midterm and final exams. Practice tests were constructed to be as parallel as possible to the midterm and final exams without having identical items. Calibration was measured by asking students to predict how many of the 25 multiple-choice items they anticipated to answer correctly and how many points they expect to receive on the 50-point short-answer/essay questions. After taking the exams, they postdicted their performance on both types of items. Bol and Hacker (2001) found that high-achieving students were more accurate than low-achieving students. Low achieving students had significantly higher prediction and postdiction accuracy for essay than for multiple-choice items.

In summing up, this subsection explored the relationship between self-efficacy and calibration, defining calibration and its importance in academic achievement. Next, was a description of how various researchers assessed calibration, concluding that the bias and accuracy measures are the optimal measures. Current social cognitive researchers (Chen, 2003; Pajares & Miller, 1997) use the bias and accuracy measure proposed by Schraw et al. (1993), which is the method used in this study. Finally, early calibration studies and findings from Hacker and Bol (2004) indicate that the results on calibration research are mixed. Intervention studies showed it is possible to train elementary students to improve accuracy (Bandura & Schunk, 1981; Schunk, 1981). However, Hacker and Bol's (2004) review of test prediction studies suggest that students are generally inaccurate in their metacognitive judgments, and it is not possible to teach students to improve their prediction accuracy. Furthermore, it is possible that current metacognitive measures of accuracy are unreliable to detect change (Hacker & Bol, 2004). As a result, these findings warrant more research in the area of calibration in order to determine more optimal ways to improve the accuracy of students' self-efficacy beliefs.

### *Synthesis*

Up to this point, the previous sections discussed the research on self-efficacy, self-regulation, and calibration with studies focusing on mathematics. This section summarizes the findings of the previous sections and analyzes the work of various researchers or groups of researchers, which would shed more light on the research done in mathematics in relation to self-efficacy and calibration.

Schunk and colleagues (Schunk, 1998; Schunk & Ertmer, 2000) have done numerous studies with elementary and middle school students and demonstrated that it is possible to teach young children various self-regulatory skills to improve their self-efficacy judgments, congruence between self-efficacy judgments and performance (calibration), and mathematics skills. These were intervention studies with lengthy periods of training and practice. Self-efficacy was measured at the optimal level of specificity that corresponded with the performance outcome confirming Bandura's (1986; 1997) guidelines that these task-specific measures have the best predictive power of academic outcomes. Moreover, self-efficacy was measured in close proximity to the criterial task, another aspect that enhances measurement accuracy (Bandura, 1986). Modeling, goal orientation, self-verbalization, strategy training, self-monitoring, and self-evaluation were some of the self-regulatory behaviors that Schunk and colleagues employed to train struggling students with significant success (Schunk, 1998).

Likewise, Pajares and colleagues, Pajares and Miller (1994; 1997) and Pajares and Graham (1999) have done numerous correlational and descriptive studies on self-efficacy, calibration, and mathematics. Pajares and Miller (1994) explored the role of self-efficacy and self-concept beliefs in mathematics-related tasks. Participants were 350 undergraduates. The dependent measures were math self-efficacy, perceived usefulness of math, math anxiety, math self-concept, prior experience in math, and math performance. Path analysis results showed that students' math self-efficacy was more predictive of their ability than the other variables. Pajares and Miller (1994) also calculated an accuracy score. They used a Likert scale from 1 to 5 for self-efficacy judgments and overestimation was judging an item 4 or 5 and getting it incorrect.

Underestimation was judging an item 1 or 2 and getting it correct. The results showed that 57% of the students overestimated their performance and 20% underestimated their performance. The researchers concluded that the undergraduates' math self-efficacy beliefs would be critical in making math-related decisions such as pursuing advance math courses, math-related majors, and careers. Inaccurate perceptions of one's capabilities would lead to avoiding math-related courses and careers. Furthermore, they suggested that it would be advantageous to have self-efficacy assessments early in students' academic career in order to identify inaccurate perceptions and employ appropriate interventions to improve them.

Another study by Pajares and Graham (1999) explored the role of self-efficacy, motivation constructs, and math performance of middle school students. Participants were 273 regular and gifted sixth grade students. The variables in the study were math self-efficacy, math self-concept, engagement, self-efficacy for self-regulated learning, math scores, and calibration measures of bias and accuracy. The results showed that students' mathematics self-efficacy beliefs were the only motivational variable that predicted math performance both at the beginning and end of the year, which supported Bandura's (1986; 1997) claim that self-efficacy beliefs predict academic achievement. Pajares and colleagues used the task-specific method of assessing of self-efficacy that corresponded with the criterial task in their research and have consistently found that self-efficacy predicts performance when measured in this manner. Pajares and Graham (1999) also found that students were overconfident, which has been a consistent finding by other researchers as well. However, gifted students were less biased toward

overconfidence and had more accurate self-efficacy beliefs (better calibrated) than regular education students.

Chen (2003) investigated the accuracy and predictability of 107 seventh grade students' self-efficacy beliefs on mathematics tasks using path analysis procedures. Among the potential causes of inaccurate self-efficacy beliefs that Chen investigated were task difficulty, gender, and students' prior achievement. Task difficulty was investigated based on findings of Schraw, Potenza, and Nebelsick-Gullet (1993) who found that students were better calibrated in their self-evaluative judgments on an easy task or test rather than on a difficult one. Chen tested four hypotheses. First, students' calibration judgments would play a mediational role between the effects of prior achievement and gender, and students' math self-efficacy beliefs and test performance. Second, self-efficacy, math performance, and calibration measures of bias and accuracy would predict postperformance measures of effort attributions and self-evaluative judgments. Third, regarding the task difficulty, linear trends would occur between difficulty and self-efficacy, and difficulty and calibration. Finally, calibration indexes combined with self-efficacy measures would increase the predictability of math performance. The dependent measures were math performance, mathematics self-efficacy measured on a Likert scale 1 to 8, mathematics effort judgment, self-evaluation, and previous math achievement.

Path analysis results showed that calibration accuracy had a significant effect in enhancing students' math self-efficacy judgments and a direct effect on postperformance self-evaluation and effort judgment. Math self-efficacy was also critical in predicting students' math performance, postperformance self-evaluation, and effort judgments.

Gender was not a significant cause of calibration and self-efficacy beliefs, but it correlated with self-evaluation; boys evaluated their math performance more positively than girls. Prior math achievement had a greater indirect effect than a direct effect on math performance, with indirect effects mediated via self-efficacy and calibration accuracy equally. Regarding task difficulty, students reported better calibration, higher self-efficacy beliefs, and favorable self-evaluations on easier math problems compared to difficult ones. Students' math performance had no direct or indirect effect on their effort judgments (Chen, 2003).

In a review article on calibration, Klassen (2002) noted that compared to writing and reading, students were more accurate in assessing mathematics tasks. One reason was the task-specific measurement of self-efficacy. In math, it is possible to briefly show students the actual task, allow them to rate their efficacy to solve it, and give them time to complete it. The task-analysis aspect of self-efficacy beliefs is much clearer for students doing a specific math task in comparison to a writing task (Klassen, 2002).

According to The National Council of Teachers of Mathematics (NCTM) content standards for grades 3-5, students' interest in math is likely to diminish if learning becomes a process of memorizing and imitating. As a result, the standards focus on multiplicative thinking, equivalence, and computational fluency. Multiplicative thinking allows students in these three grades to build their understanding of fractions as a part of a whole and as division. In mathematics, students can use solution strategies to solve problems and employ various strategies to self-correct any errors. Thus, math tasks that can be self-corrected for accuracy would enable students to make judgments of their performance rather than depending solely on external feedback.

In conclusion, the intervention and correlational research on mathematics, self-efficacy, and calibration have shown that self-efficacy predicts mathematics achievement when measured optimally, that is, the measurement should be task-specific and correspond to the performance criteria. Furthermore, it is possible to train students to enhance their self-efficacy beliefs, math skills, and accurate self-efficacy judgments (calibration).

#### *Rationale for Current Study*

The conclusions of the literature review suggest that American students' math achievement should be enhanced in order to compete globally. The subject domain in this study is mathematics and the tasks would involve subtraction fraction problems. An attractive feature of mathematics as a subject matter to study calibration is that it can be measured quickly and objectively. Moreover, students across America participate in National Assessment of Educational Progress tests, which are given in different subject areas in fourth, eighth, and twelfth grades. Mathematics is one subject area where fourth and eighth graders have been improving over the past two decades (Lee, Grigg, & Dion, 2007). It is important for this trend to continue if American children are to compete with Asian and European children. The latest TIMMS (2007) report indicated that while American fourth and eighth graders scored above the international average in math, they performed poorly compared to their counterparts in Asian and some European countries.

This proposed study incorporated some aspects of The NCTM Standards for elementary mathematics education (NCTM, 2008). First, it incorporated a variety of self-regulatory strategies to solve problems. Second, it included problem-solving strategies, which would build students' mathematical skill and self-efficacy. Third, many

operations (subtraction, multiplication, and addition) are included in a domain (fractions). A recent report stated that many students and adults have difficulties with fractions, a foundation skill essential for success in algebra (National Mathematics Advisory Panel, 2008). Finally, the problems are in word-format, which is also consistent with the NAEP and international tests.

Second, students' low mathematics scores are linked to poor calibration, which is a key metacognitive judgment. The research on calibration is inconclusive, indicating that there is a tendency toward overestimation (Hacker & Bol, 2004; Klassen, 2002; Pajares, 1996a). However, some researchers note that students have better calibration on easier tasks compared to difficult ones (Chen, 2003). In addition, gifted and high-achieving students have better calibration compared to low-achieving students (Hacker & Bol, 2004; Pajares & Graham, 1999). Most of these studies were correlational or quasi-experimental; as a result, intervention studies in calibration are lacking. It is hypothesized that training students to distinguish the strategies necessary to solve tasks of varying difficulty level would lead to improved self-efficacy calibration judgments. Dempster (1995) suggests devising a plan or set of operations to solve problems. These plans or steps move the problem solver from the initial state to the solution state (Dempster, 1995). Bandura and Cervone (1983) remark that when knowledge of the task and one's capability to do it is clear, self-efficacy influences learning and performance outcomes. Therefore, it is important to train students in cognitive and metacognitive strategies in order to improve their math performance and accurate judgments. Calibration measures of accuracy and bias will be used in this proposed study as described by Keren (1991)

and Yates (1990), which was later modified by Schraw (1995) and (Pajares & Graham, 1999).

Third, many studies have reported that students are generally overconfident and it has been difficult to train or teach students to improve calibration (Hacker & Bol, 2004). It is important to understand why students are overconfident on difficult tasks mentioned above. It is possible that students have difficulty determining an easy task from a moderately difficult to a difficult one and that training them to use solution strategies may improve calibration judgments and math performance. Deficient self-knowledge (overestimates or underestimates of self-efficacy) may result in faulty task understanding, not using solution strategies, and difficulty in self-reflecting on one's performance (Klassen, 2002). Furthermore, Bandura (1986) hypothesized that optimistic overestimates of capability may allow students to select challenging tasks, persist at them, and complete them successfully. As a result, the time spent on each problem, and the use of solution strategies in solving them were recorded in this proposed study.

Fourth, from a social cognitive perspective, math achievement and calibration would be improved by a three-phase intervention: forethought, performance, and self-reflection. Consequently, this study differs from previous studies in that it seeks to test whether an intervention based on Zimmerman's (1998; 2000) cyclical-phase model of self-regulation will improve self-efficacy, math skill, self-evaluation, self-efficacy calibration, self-evaluation calibration, time on task, and use of strategy. The current study included one subprocess from each phase. In the forethought phase, students rated their self-efficacy judgment to solve each problem. In the performance phase, intervention students participated in strategy training, and in the self-reflection phase,

students reflected on possible causes of errors if their responses were incorrect. All students received feedback regarding correctness in the practice phase where they worked on a new set of problems. It was hypothesized that this cyclical self-regulatory intervention will improve students' self-efficacy, math skills, and calibration.

Although, Schunk and associates have conducted self-efficacy and mathematics research among elementary students, much of the research on self-efficacy has focused on students from the intermediate grades to college (Pajares, 1996a). One reason for lack of research at earlier ages is the belief that young children have limited cognitive skills (Gaskill & Murphy, 2004). However, numerous studies (Schunk & Ertmer, 2000; Montague & Bos, 1986) have shown that it is possible to train low-performing elementary children to estimate their self-efficacy judgments and improve math skill. Nevertheless, their intervention studies focused primarily on improving math achievement and various motivational constructs, such as self-efficacy, persistence, self-verbalization, and task engagement. Only two studies, mentioned earlier (Schunk, 1981; Bandura & Schunk, 1981), examined calibration as an outcome measure.

### *Hypotheses*

Hypothesis 1: Students in the strategy training groups will perform significantly better than the non-strategy training groups on the dependent measures of mathematics performance, self-efficacy, self-evaluation, self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, time on task, and strategy use.

Hypothesis 2: Students in self-reflection training groups will perform significantly better than the non self-reflection groups on the dependent measures mentioned above.

Hypothesis 3: Students' over-estimates of self-efficacy will correlate negatively with their time spent on the difficult math problem.

Hypothesis 4: Students' overestimates of self-efficacy will correlate negatively with an advantageous strategy choice.

Hypothesis 5: Students' self-efficacy, self-efficacy bias, and self-efficacy accuracy judgments are hypothesized to play a mediational role between the effects of strategy training and mathematics achievement.

Hypothesis 6: Students' mathematics performance is hypothesized to play a mediational role between the effects of self-reflection training and self-evaluation, self-evaluation bias, and self-evaluation accuracy judgments.

## Chapter 3

### METHODOLOGY

This chapter described the methodology of this study that examined whether an intervention based on a cyclical model of self-regulation would improve fifth graders' mathematics skill, self-efficacy, self-evaluation, self-efficacy bias, self-efficacy accuracy, self-evaluation bias, and self-evaluation accuracy. The chapter begins with a selection of participants. Next is a description of the measures and the procedures for conducting the study. Finally, the chapter concludes with a description of data analysis.

#### *Participants*

The participants were recruited from two parochial and one private school in New York City. The students were normally achieving, English-speaking fifth graders. A letter describing the study (see Appendix E), parental/guardian consent forms (see Appendix F), and student assent forms (see Appendix G) were sent to the schools. A total of 205 fifth graders' parents got consent forms, asking them to allow their children to participate in the study. Ninety-four returned the consent and assent forms. Of these, six students' data were part of the pilot study and were not included in the final analysis. The final sample included in the study was 88 students comprising of 39 boys and 49 girls. Each student was randomly assigned to one of four groups. According to Cohen (1988), a power analysis indicated that a sample of 64 students is necessary to detect a medium effect size at  $\alpha = .05$  level with 80% statistical power and a sample of 60 to detect large effect sizes at the  $p < .05$  level of significance (Cohen, 1988). As a result, the sample of 88 would yield medium to large effect sizes at  $\alpha = .05$  level with 92% and 99% statistical power respectively.

### *Task*

The mathematics fraction problems were developed based on guidelines from two math texts for elementary grades (Everyday Mathematics, 2004; Schwartz, 2008) and an internet based software, Jamit Fractions, Worksheet 6. “Jamit Fractions is a unique math educational software for learning fractions online, suitable for students from 8 years old to adults. The program features 10 topics. Each topic is presented in a short lesson with complete references, examples, and an individual practice game for the current topic to ensure that the topic is thoroughly learned” (<http://jamit.com.au/fractions.htm>).

The worksheets on this webpage were classified as easy, medium, hard, and very hard. In addition, the difficulty level of the items was determined by a pilot study and in consultation with the schools’ principals. A pilot study was conducted with six students initially. The results showed ceiling effects in the pretest phase, indicating that the problems were too simple (Appendix A). As a result, the number of items and difficulty level was increased. The new set of items (Appendix B) was piloted on eight students to determine whether the mathematics problems were of the appropriate difficulty level for the fifth grade students. The pretest test results did not indicate ceiling effects and this final set of math problems ranged from simple and difficult.

### *Measures*

*Self-efficacy.* The following task-specific scale assessed students’ judgment of their capability to do each mathematics problem. Students made a judgment on each problem before solving it. The math self-efficacy measure is indicated below and the range was from 0 to 100.

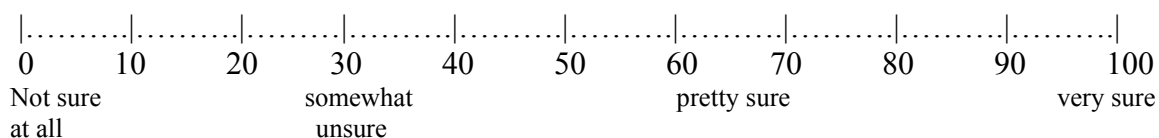


(0), the bias score would be  $(100-0) = 100$ , signifying overconfidence. Therefore, self-efficacy calibration bias can range from -100 to + 100 with scores greater than zero indicating overconfidence and scores below zero signifying under confidence. A mean bias score was calculated for the six bias scores.

*Self-efficacy accuracy.* A second calibration measure was calculated by subtracting the absolute value of each bias score from 100. This score conveyed the magnitude of the judgment error and it ranged from 0, which is complete inaccuracy to 100, complete accuracy (Pajares & Miller, 1997). The mean score for these six accuracy scores was computed for each student.

*Self-evaluation.* This scale assessed students' confidence about their answers. Students rated their performance on this scale after solving each math problem. This scale was adapted from Chen's study (2003), which ranged from 1 to 8. The scale used in this study ranged from 0 to 100. The following prompt was given to all participants.

"After solving the problem, how sure are you that you have solved it correctly? Circle whichever number you believe is appropriate for this specific problem."



The internal consistency reliability for self-efficacy and for self-evaluation scales were both  $\alpha = .92$ , indicating high levels of reliability for each scale.

*Self-evaluation bias.* Using the self-evaluation ratings and the math scores on the posttest, the researcher calculated self-evaluative bias scores similar to the self-efficacy bias score above.

*Self-evaluation accuracy.* Using the self-evaluation ratings and the math scores on the posttest, the researcher calculated self-evaluative accuracy similar to the self-efficacy accuracy measures above.

*Time on math task.* The time each student spent on each problem was recorded in seconds for the practice phase and for the posttest as well.

*Strategy use.* Students' response on each problem indicated whether they used the solution strategy or not. A step-by-step solution strategy that resulted in a correct answer was scored as 3 points. An incomplete strategy solution with some correct steps was scored as 1 point, and 0 points for no strategy use.

### *Research Design*

The research design is a  $2 \times 2$  factorial with strategy versus no strategy, and self-reflection versus no self-reflection as independent variables and math performance, self-efficacy, self-evaluation, self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, time on problem, and strategy use serving as dependent measures.

### *Procedures*

The researcher conducted the study with students individually in a room in each school. Students were randomly assigned to one of four respective groups, (1) strategy training and self-reflection (2) strategy training and no self-reflection (3) no strategy training with self-reflection and (4) no strategy training and no self-reflection (control). Each session lasted approximately an hour and consisted of 20 fraction subtraction problems with a pretest, a training phase, a practice phase, and a posttest (Appendix B). Students in the strategy training groups were taught a step-by-step strategy to solve

various math problem (Appendix C). Students in the self-reflection training groups were taught how to reflect on their errors (Appendix D).

*Pretest phase.*

Each student completed a set of six fraction problems.

Instructions: I am here to find out how children your age solve subtraction fraction problems. I may ask you some questions periodically to understand what you are doing. Show your work and solve as many problems as possible. If you have difficulty with any problem, move on to the next one.

After the pretest, students assigned to strategy training groups were shown a set of four mathematics subtraction problems and were taught how to solve each problem using a step-by-step strategy. However, students in the self-reflection and control groups were advanced directly to the practice phase. At the end of the study, the self-reflection only group students were taught the strategy methods for each problem, and the control group students were trained in both strategy and self-reflection methods for ethical reasons. Thus, the training was additive and cyclical, that is, self-efficacy judgment preceded strategy training, which in turn preceded self-reflection training.

*Training phase.*

*Strategy Training.* Instructions to groups one and two: I will teach you some strategies to use when solving various fraction problems. There is a way to judge the difficulty of a fraction problem depending on the strategy needed to solve it.

*Instructions to self-reflection group only:* It is important to read each problem very carefully before trying to solve it. If you miss an important part of the problem, it will affect your ability to solve it. I want you to read these problems carefully. After

solving each problem in the practice phase, if there are any errors, I will circle them. I told them, “This is the part you got wrong. What can you do to correct it?” After this, the student was told to continue and solve the next problem and the process was repeated depending on the student’s response. If there were errors in the response, the above short self-reflective questions were repeated. However, if the response was correct, the student was told to proceed and solve the next problem.

*Instructions to control group:* It is really important to read each problem very carefully before trying to solve it. If you miss an important part of the problem, it will affect your ability to solve it. I want you to read these problems carefully.

*Practice phase*

Instructions to each student: Here is a new set of four problems to practice. You may choose any order you wish to do them. I will give you feedback after each problem about the correctness of your answers.

Students in each group got independent time to practice on a set of four new math items. The researcher recorded the time students spent on each problem and noted whether they used the solution strategy. Each student received feedback regarding the correctness or incorrectness of his/her answers after each problem. Moreover, responses with errors were circled for students in the self-reflection groups. The researcher told them, “This is the part you got wrong. What can you do to correct it?”

*Posttest Instructions*

Okay, you have practiced and I want to see how much you have learned. I cannot give you feedback on these problems, but do your best. I may ask you some questions to gain a better understanding of how you are solving the problems.

First, all students completed the self-efficacy ratings for each problem before attempting to solve it. Each problem was shown for a few seconds and students made a judgment without solving the problem mentally. After the self-efficacy judgments, students solved the six math problems in any order they chose. Upon completion of each problem, students rated their performance on the self-evaluation scale. The researcher recorded the time spent on each problem and noted whether they used the solution strategy.

#### *Data Analyses*

Descriptive statistics were conducted for each dependent measure: self-efficacy, self-evaluation, mathematics performance, self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, time on each problem, and use of strategy.

A 2 X 2 multivariate analysis of covariance (MANCOVA) was conducted to test for main effects and interactions. The students pretest score was used as a covariate in the statistical model. The MANCOVA analysis revealed significant main effects for strategy training, and it was followed by analysis of covariance on each dependent measure to determine whether students in the experimental group performed significantly better than students in the control group. Significant interactions will be analyzed using posthoc tests. These analyses will determine whether hypotheses 1 and 2 are accepted.

Next, correlations were done to determine whether the dependent measures of time spent on task, and strategy use would correlate negatively with calibration measures of bias. These results would answer hypotheses 3 and 4. These analyses were followed by supplementary analyses. Path analysis was done to determine whether the model indicated causality. These analyses provided answers for hypotheses 5 and 6. Finally,

nonparametric chi-square tests were done on the mean scores of the self-reflection and control groups to determine whether the difference between the two groups were statistically significant.

## Chapter 4

## RESULTS

*Descriptive Statistics*

Table 1 presents the means and standard deviations for the following dependent measures: mathematics performance, self-efficacy, self-evaluation, self-efficacy accuracy, self-efficacy bias, self-evaluation accuracy, self-evaluation bias, strategy use, and time on task.

*Multivariate Analysis*

Although, the hypotheses did not incorporate gender as an independent variable, a supplementary multivariate analysis of covariance was conducted to assess the effect of gender across the dependent measure array, but it did not reveal a significant effect. Thus, gender was dropped from subsequent analyses. To examine the effects of self-regulatory training on the nine dependent measures, an overall multivariate analysis of covariance was performed. The results showed a significant main effect for strategy training, Wilks' Lambda = .43,  $F(7, 77) = 14.68$ ,  $p < .05$ , partial  $\eta^2 = .57$ , indicating that strategy training significantly improved the strategy groups' performance compared to the non-strategy groups. However, the multivariate main effect for self-reflection was nonsignificant, suggesting that the self-reflection groups did not differ from the non self-reflection groups. There was also a nonsignificant result for the interaction effects between strategy training  $\times$  self-reflection training.

Table 1

Means and standard deviations for self-efficacy, self-evaluation, math performance, self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, strategy use, and time on task

Measures	Strategy groups (n = 44)		No strategy groups (n = 44)	
	M	SD	M	SD
Math performance	70.64	19.48	33.33	26.84
Self-efficacy	61.02	22.01	50.08	23.48
Self-evaluation	64.36	22.67	47.20	25.93
Self-efficacy bias	-9.62	18.23	16.74	28.71
Self-efficacy accuracy	65.68	15.77	60.83	17.60
Self-evaluation bias	-6.29	13.50	13.86	27.23
Self-evaluation accuracy	80.23	12.83	68.94	17.90
Strategy use	2.07	0.58	0.95	0.79
Time (seconds)	117.54	33.13	87.37	32.01
Measures	Self-reflection groups (n = 44)		No self-reflection groups (n = 44)	
	M	SD	M	SD
Math performance	53.41	31.40	50.57	28.67
Self-efficacy	56.74	23.83	54.36	22.95
Self-evaluation	58.94	24.14	52.61	27.08
Self-efficacy bias	3.33	26.63	3.79	28.35
Self-efficacy accuracy	65.68	17.84	60.83	15.49
Self-evaluation bias	5.53	23.29	2.05	24.14
Self-evaluation accuracy	75.08	16.82	74.10	16.34

Table 1 continued

Strategy use	1.55	0.92	1.48	0.86
Time (seconds)	103.33	37.68	101.57	34.15

---

Note. N = 88

### *Univariate Analyses*

To determine the effects of strategy training on each separate dependent measure, a univariate analysis of covariance was conducted with strategy as the independent variable, pretest as a covariate, and the nine measures above as dependent variables. The results showed significant main effects for strategy training on the following eight dependent measures: mathematic performance, self-efficacy, self-evaluation, self-efficacy bias, self-evaluation bias, self-evaluation accuracy, time, and strategy use. The only measure that did not attain statistical significance was self-efficacy accuracy (see Table 2). These findings indicate that strategy training significantly enhanced the performance of the strategy training groups compared to the non-strategy training groups.

Table 2

Univariate test results for each dependent measure

Measures	df	F	partial $\eta^2$	Sig.
Math performance	1	96.33*	.53	.01
Self-efficacy	1	3.82*	.04	.05
Self-evaluation	1	10.25*	.11	.01
Self-efficacy bias	1	24.63*	.23	.01
Self-efficacy accuracy	1	1.21	.01	.27
Self-evaluation bias	1	17.78*	.17	.01
Self-evaluation accuracy	1	10.15*	.11	.01
Strategy use	1	21.70*	.53	.01
Time	1	17.44*	.17	.01

\*  $p < .05$ .

### *Correlation Analyses*

Pearson correlation coefficients for the dependent measures are reported in Table 3. All the variables correlate significantly with mathematics, except self-efficacy bias and self-evaluation bias, which correlated negatively. The negative correlation of self-efficacy bias ( $r = -.67$ ) and self-evaluation bias ( $r = -.56$ ) indicate that as students became overconfident, their mathematics performance decreased. As a result, it was underconfidence that led to improved math performance. On the other hand, both self-efficacy ( $r = .22$ ) and self-evaluation accuracy ( $r = .42$ ) correlated positively with math, indicating that as students judged their capability accurately, it improved their math

performance, with self-evaluation accuracy showing a stronger correlation than self-efficacy accuracy. Both self-efficacy bias ( $r = -.36$ ) and self-evaluation bias ( $r = -.32$ ) scores correlated negatively with time, indicating that underconfidence resulted in less time spent on the math task. However, both accuracy measures, self-efficacy and self-evaluation, did not correlate with time.

To test for the third and fourth hypotheses, the math problems were categorized as simple and difficult problems. Using the criterion score of above 65 as simple items, the mean score on math items 1, 2, and 4 were above 65 and they were classified as simple. Likewise, using the criterion score of below 46 as difficult items, the mean score on math items 3, 5, and 6 were below 46 and were classified as difficult. The correlation between time and difficult math items correlated positively ( $r = .37$ ), indicating students spent more time on difficult math items. However, this was not different from the correlation of time and simple math items ( $r = .42$ ). The correlation between difficult math items and self-efficacy bias was ( $r = -.53$ ) and self-evaluation bias ( $r = -.40$ ) respectively, signifying a negative correlation, which imply that as the items became difficult, students underconfidence increased. Therefore, students underestimated their performance instead of overestimating it, a finding that does not support the third hypothesis that students' over-estimates of self-efficacy will correlate negatively with their time spent on the difficult math problems.

Students in the training groups were taught a strategy to solve the problems. Strategy use correlated positively with math performance ( $r = .99$ ), suggesting that strategy use influenced math performance. Moreover, strategy use correlated negatively with both bias measures, self-efficacy bias ( $r = -.69$ ) and self-evaluation bias ( $r = -.58$ ),

indicating that as students became more underconfident, strategy use increased. This result does not support the fourth hypothesis, which stated that students' overestimates of self-efficacy would correlate negatively with strategy use. On the contrary, underestimation correlated negatively with strategy use. Self-efficacy accuracy did not correlate with strategy use; however, self-evaluation accuracy correlated positively with strategy use, suggesting that students who evaluated their math performance accurately performed better on subsequent problems.

Table 3

Intercorrelations between dependent measures (N = 88)

Measures	1	2	3	4	5	6	7	8	9
1. Math performance	–	.50**	.65**	-.67**	.22*	-.56**	.42**	.99**	.46**
2. Self-efficacy		–	.85**	.31**	-.01	.30**	-.06	.47**	.16
3. Self-evaluation			–	.02	.01	.27*	-.02	.62**	.24*
4. Self-efficacy bias				–	-.25*	.87**	-.52**	-.69**	.36**
5. Self-efficacy accuracy					–	-.27*	.73**	.20	.03
6. Self-evaluation bias						–	-.56**	-.58**	-.32**
7. Self-evaluation accuracy							–	.42**	.14
8. Strategy use								–	.44**
9. Time									–

\* Correlation is significant at the .05 level (two-tailed)

\*\* Correlation is significant at the .01 level (two-tailed).

### Supplementary Path Analysis

An important follow-up question is which of the six self-regulatory judgments mediated the influence of strategy training and self-reflection on math performance. The software Lisrel 8.80 Student Edition (Jöreskog & Sorbom, 2006) was used to conduct a path analysis to test two hypotheses. (a) The role of strategy training on math performance through the mediating variables of self-efficacy, self-efficacy bias, and self-efficacy accuracy (Figure 2); and (b) the role of self-reflection on self-evaluation, self-evaluation bias, and self-evaluation accuracy through math performance as the mediating variable (Figure 3). This methodology uses manifest variables to test the effect of strategy and self-reflection training on students' mathematics performance through various mediating variables. In model one, the exogenous variable was strategy training, which was hypothesized to influence mathematics performance. The next sequential variables were endogenous variables, self-efficacy, self-efficacy bias, self-efficacy accuracy. In model two, the exogenous variable was self-reflection and the next sequential endogenous variable was math performance, the mediating variable on self-evaluation, self-evaluation bias, and self-evaluation accuracy.

*Figure 2*

Proposed path model from strategy training to math performance

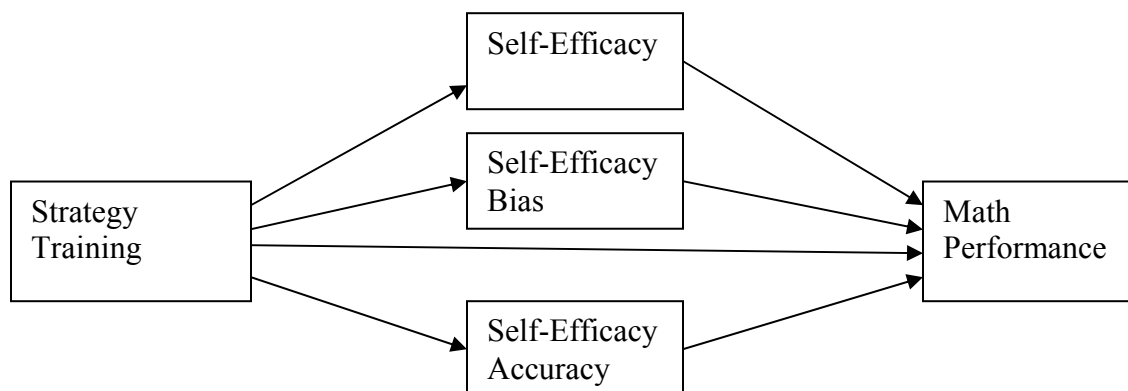
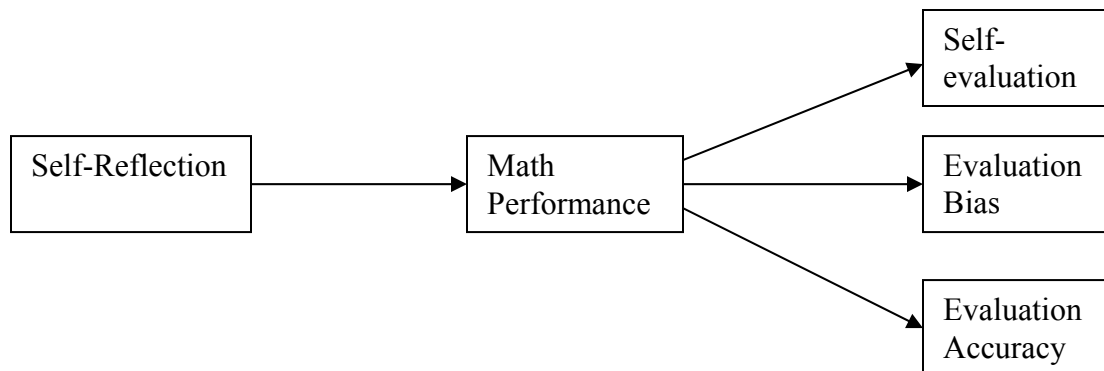


Figure 3

Proposed path model from self-reflection to math performance



Estimations of the proposed model (Fig. 2) using LISREL 8.80 Student Edition (Jöreskog & Sorbom, 2006) showed a nonsignificant  $\chi^2$  value,  $\chi^2 = (1, N = 88) = 0.01, P = 1.00$ ; a goodness-of-fit index (GFI) of 1.00; and a comparative fit index (CFI) of 1.00. The initial paths (Fig. 2) were nonsignificant with a poor fit for the model. The following paths were removed sequentially: (1) strategy to self-efficacy accuracy (2) self-efficacy accuracy to math performance (3) strategy to self-efficacy (4) self-efficacy to math performance. Initially, the direct path from strategy to math performance was removed due to nonsignificance. However, the resulting fit from strategy to efficacy bias and from efficacy bias to math performance was also poor. The direct path from strategy to math performance was added, and the resulting model indicated a good fit in Figure 4.

The final model suggests that strategy training influenced math performance both directly and indirectly through self-efficacy bias. Bias reflects the direction of judgment errors with negative scores indicating underconfidence and positive scores indicating overconfidence. The path suggests that strategy training decreased students' bias, (underconfidence), which improved their math performance. The squared multiple

correlations show that strategy training had a larger effect on self-efficacy bias than on math performance directly, indicating that the calibration measure mediated the influences of strategy training on math performance. The predictor variables (strategy training and self-efficacy bias) accounted for 57% of the variance in mathematics performance. Table 4 depicts the decomposition of the path analysis into direct and indirect effects. The path model shows that the direct effects of strategy training on math skills are much larger than the indirect effects. Although, the indirect effect is smaller, it is significant, implying it played a pivotal role in math performance.

Figure 4. Final reduced path model from strategy training

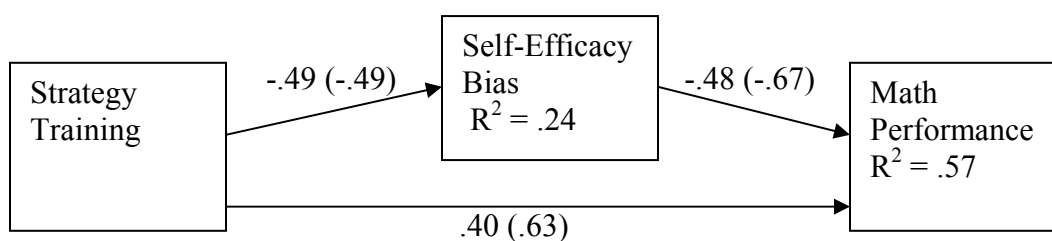


Table 4

Decomposition of effects from the path analysis (strategy training)

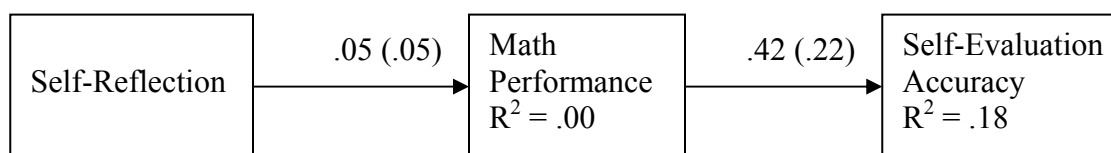
<u>Effect</u>	<u>Direct effect</u>	<u>Indirect effect</u>	<u>Total effect</u>
Of strategy			
On self-efficacy bias	-.49*	0	-.49*
On math performance	.40*	.23*	.63*
Of self-efficacy bias			
On math performance	-.48	0	-.48*

Note. \*  $p > .05$

Next, estimations of the proposed model (Fig. 3) using LISREL 8.80 Student Edition (Jöreskog & Sorbom, 2006) showed a nonsignificant  $\chi^2$  value,  $\chi^2 = (1, N = 88) = 0.01, P = 0.93$ ; a goodness-of-fit index (GFI) of 1.00; and a comparative fit index (CFI) of 1.00. The initial paths revealed a poor fit, and the following paths were removed sequentially from the proposed model (a) self-reflection to math performance (b) math performance to self-evaluation (c) self-reflection to math performance (d) math performance to evaluation bias. The final path is in Figure 5.

Figure 5

Final reduced path model from self-reflection



Although, the final model indicated a good fit, only the path from math performance to self-evaluation was significant. The path from self-reflection to math performance was nonsignificant. The direct and indirect effects are presented in Table 5.

Table 5

Decomposition of effects from the path analysis (self-reflection training)

Effect	Direct effect	Indirect effect	Total effect
Of self-reflection			
On math performance	.05	0	.05
On self-evaluation accuracy	0	.02	.02
Of math performance			
On self-evaluation accuracy	.42*	0	.42*

---

Note. \*  $p > .05$

*Supplementary Chi-Square analysis*

Although the self-reflection training did not have a significant influence on the self-regulatory measures or math skill according to the multivariate analyses, an examination of the mean score of each dependent measure revealed that they were generally higher for the self-reflection group. To test whether these numerical patterns were statistically significant, a nonparametric chi-square analysis was conducted on the mean scores of the nine dependent measures (see Table 1). Between the self-reflection and the non self-reflection group, the group that had the higher mean for each dependent measure was scored 1 and the group with the lower mean was scored 0. The results showed that the difference between the two groups was statistically significant,  $\chi^2(1, N = 9) = 5.44 p < .05$ . This result suggests that self-reflection training did exert a weak, but statistically reliable effect on self-efficacy, self-evaluation, self-efficacy accuracy, self-efficacy bias, self-evaluation accuracy, math performance, strategy use, and time.

## Chapter 5

### DISCUSSION

This primary goal of this study was to investigate the effects of strategy training and self-reflection training, two subprocesses of Zimmerman's cyclical model of self-regulation, on fifth grade students' mathematics performance, self-efficacy, self-evaluation, calibration measures of self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, time on task, and strategy use. To determine the effects of the self-regulatory training, a multivariate analysis of covariance was conducted. Findings from this analysis revealed a significant main effect for strategy training, signifying that strategy training enhanced the performance of the intervention groups compared to the non-strategy groups. The size of the main effect was .57, which is classified as statistically large (Cohen, 1988), indicating a large difference between the two groups. Follow-up analysis of covariance indicated that except self-efficacy accuracy, there was a main effect of strategy training on the remaining eight dependent measures. The size of the main effect on each dependent measure ranged from medium to large with the exception of self-efficacy, which was marginally significant and had a small effect (see Table 2). Two supplementary path analyses were conducted, and the results showed that strategy training had direct and indirect effects on math performance. Moreover, the dependent measure, self-efficacy bias played a mediational role in predicting math performance. The second path analysis showed the effects of self-reflection training on math performance and self-evaluation accuracy were nonsignificant, but the effects of math performance on self-evaluation accuracy were

significant. Because the effects of self-reflection training were weak, it was not possible to confirm the mediational role of math performance on self-evaluation accuracy.

Finally, although there was a non-significant effect of self-reflection on the dependent measures, a supplementary nonparametric chi-square analysis revealed that with the exception of self-efficacy bias, self-reflection exerted a weak, but statistically reliable influence on the remaining eight dependent measures, suggesting that it influenced students' overall math performance. On the whole, the results indicated that the self-regulatory processes of strategy training and self-reflection improved students' math performance, calibration measured as less bias and more accuracy, self-efficacy, and self-evaluative judgments, albeit the independent measure of self-reflection being much weaker compared to strategy training intervention.

## Hypotheses

### *Hypothesis 1*

The first hypothesis posited that students in the strategy training groups will perform significantly better than the non-strategy training groups on the dependent measures of mathematics, self-efficacy, self-evaluation, self-efficacy bias, self-efficacy accuracy, self-evaluation bias, self-evaluation accuracy, time on task, and strategy use. The findings from the multivariate analysis indicated a statistically significant main effect of strategy training on the dependent measures. This result implies that self-regulatory strategy training greatly enhanced self-efficacy, self-evaluative judgments, calibration judgments (less bias and more accuracy), and math performance of the strategy training groups compared to the non-strategy groups. Follow-up analysis of covariance showed statistically significant main effect for strategy training, except for self-efficacy accuracy,

indicating that students in the non-strategy training groups overestimated their capability compared to the strategy groups. This remains an issue as Pajares (1996a) suggested that interventions should not lower students' self-efficacy, but the goal should be how to improve students' calibration accuracy. This study suggests that strategy training would enable students to decipher the task and determine the skills necessary to complete it, thereby leading to judgments that are more accurate.

Schunk and Cox (1986) did a study with learning disabled middle students using math problems and found that strategy training and attributional feedback improved the students' self-efficacy and skill performance compared to the control group. Likewise, this current study demonstrated that strategy training can indeed improve students' math skill, self-judgments, and calibration measures. Moreover, the results suggest that it is possible to train students to improve their calibration judgments, contrary to the findings of Hacker and Bol (2004) that students' metacognitive judgments are resistant to change. Two earlier studies, involving mathematics problems and accuracy as outcome measures, (Bandura & Schunk, 1981) found that students in a modeling condition appraised their capability more accurately compared to those in a didactic condition. Moreover, students in the model condition with effort attribution assessed their capability more accurately. Likewise, Schunk (1981) found that students in a proximal subgoal condition were highly accurate of their appraisals compared to those in distal and no-goal conditions.

Klassen (2006) has analyzed a number of studies of students with learning disabilities and their self-efficacy beliefs and academic performance using writing, math, and reading tasks. In these studies, the authors calculated calibration indices of underestimation, overestimation, and accuracy. Klassen found that in a few studies

involving mathematics (Alvarez & Adelman, 1986; Schunk, 1985; Schunk & Cox, 1986), the methodology used to assess students' accuracy judgments was clear. Chen's (2003) study also involved mathematics problems and she found that calibration accuracy had a significant effect on students' mathematics performance and math self-efficacy judgments. Likewise, Bandura and Schunk (1981) and Schunk (1981) also used mathematics items in their study. According to Klassen, mathematics tasks can be measured more objectively compared to writing or reading tasks. In math, the researcher can show the problem to the students briefly for a self-efficacy assessment and then evaluate the same problem as a performance measure. In writing or reading, the researcher has to measure calibration using a qualitative approach. This current study adds to the limited body of calibration literature suggesting that tasks involving mathematics can be measured more objectively, thereby allowing students to make better judgments of their capability. The path analysis results also support Pajares' (1996a) assertion that mathematics-related studies have shown higher correlations and stronger paths compared to reading and writing.

On the other hand, most of the tasks used by other researchers involved exams, quizzes, (Bol & Hacker, 2001; Bol, Hacker, O'Shea, & Allen, 2005; Hacker, Bol, Horgan, & Rakow, 2000) and self-regulatory scales, predicting performance, prior work, and exam scores (Garavalia, 2005; Garavalia & Gredler, 2002). Thus, tasks involving mathematics have shown to improve students' calibration judgments. This poses another challenging problem because it may not be possible to improve students' calibration judgments in non-mathematics task. Clearly, more research is needed in this area, and the potential holds for interesting findings.

Another critical feature is how self-efficacy is measured. Bandura (1986; 1997) suggests a microanalysis approach, which involves measuring the self-judgment and performance in close proximity in order to obtain the maximum accuracy. The studies mentioned above followed this task-specific guideline in measuring self-efficacy, which is one possible reason for improved accuracy in students' performance. Omnibus measures of self-efficacy do not relate to a specific task, but evaluate global performance on an exam or a quiz, a method used by other researchers, Hacker, Bol and colleagues. The predictive power of such measures is much weaker compared to a task-specific method (Bandura, 1997; Pajares, 1996a). Although, there was no hypothesis delineating the role of gender (a between-subject variable) in this study, it was tested as an independent variable in the multivariate analysis with non-significant results. Because of non-significant results, it was dropped from subsequent analyses. This finding, of non-significant gender main effects among fifth graders, adds to those of previous researchers who also found non-significant differences among seventh graders (Chen, 2003), middle school (Pajares & Graham, 1999), and high school students (Pajares & Kranzler, 1995).

### *Hypothesis 2*

The second hypothesis stated that students in the self-reflection training groups would perform significantly better than the non self-reflection groups on the dependent measures. The multivariate analysis of covariance indicated a non-significant main effect of self-reflection training, suggesting that it did not contribute to students' self-evaluative judgments, math performance, and self-evaluative calibration measures. As a result, a follow-up univariate analysis was not done. However, upon examining the mean scores of the self-reflection and the non self-reflection groups, a consistent trend emerged with

the exception of one variable, self-efficacy bias. This variable was lower for the self-reflection group. The mean scores of the remaining eight variables were higher for the self-reflection groups. To test the significance of these results, a nonparametric chi-square test was conducted. The results indicated a statistically significant difference between the two groups, suggesting that self-reflection training exerted a consistent and reliable effect, albeit a weak one. Schunk and Ertmer (2000) noted that there is a need for more research on self-reflection, a critical component of self-regulated learning. The significant nonparametric findings suggest that self-reflective training has the potential to improve students' learning, but it is not clear whether in the absence of academic skills, students would be able to reflect on their performance.

### *Hypothesis 3*

The third hypothesis stated that students' over-estimates of self-efficacy would correlate negatively with the time spent on the difficult math problem. The bias measure indicated the direction in the errors of judgment and it ranged from -100 to 100 with negative numbers indicating underconfidence or underestimation and positive numbers overconfidence or overestimation. Both self-efficacy bias ( $r = -.36$ ) and self-evaluation bias ( $r = -.32$ ) correlated moderately, but significantly with time on math task. The negative direction indicate that as students' bias decreased (underestimation), they spent more time on the math problems, which led to better math performance ( $r = .46$ ). The accuracy measures of self-efficacy and self-evaluation did not correlate with time spent on the task (see Table 3).

Correlation analyses between difficult math items and self-efficacy bias was ( $r = -.53$ ) and self-evaluation bias ( $r = -.40$ ), suggesting that as math items increased in

difficulty, students became underconfident. As a result, this finding does not support the third hypothesis that students' overestimates of self-efficacy will correlate negatively with time spent on difficult math items.

#### *Hypothesis 4*

The fourth hypothesis proposed that students' overestimates of self-efficacy would correlate negatively with an advantageous strategy choice. Both self-efficacy bias ( $r = -.69$ ) and self-evaluation bias ( $r = -.58$ ) had significant and moderately strong correlations with strategy use. These correlations indicate that as students' bias decreased (underestimation), use of the strategy increased, which led to better math performance ( $r = .99$ ). Furthermore, in the final path analysis model (see Figure 4), self-efficacy bias played a significant mediational role between strategy training and math performance, implying that it was key mediating variable influencing math performance.

#### *Hypothesis 5*

The fifth hypothesis stated that students' self-efficacy, self-efficacy bias, and self-efficacy accuracy judgments are hypothesized to play a mediational role between the effects of strategy training and mathematics performance. The path analysis results indicated that only self-efficacy bias played a mediational role between strategy and math performance. As a result, the findings do not support the hypothesis. Self-efficacy bias is a calibration index of underconfidence and overconfidence. Students' underconfidence resulted in better math performance, implying that metacognitive judgments played a key role in math performance. Additionally, strategy training had direct effects on math performance, indicating that math skills were critical for success in the math tasks.

### *Hypothesis 6*

The sixth hypothesis proposed that mathematics performance would play a mediational role between the effects of self-reflection training and self-evaluation, self-evaluation bias, and self-evaluation accuracy. That is, self-reflection would influence self-evaluative judgments and self-evaluative calibration judgments after students perform the math task. The final path indicated that self-reflection influenced self-evaluation accuracy judgments through math performance. The path from self-reflection to math performance was not significant; however, the path from math performance to self-evaluation accuracy was significant. Thus, the effects of self-reflection training on math performance were so weak that it was not possible to confirm the mediational role of math performance on self-evaluation accuracy. Clearly, stronger self-reflection training is needed in future efforts to test this hypothesis.

### *Limitations*

The participants in this study were fifth grade students from private, parochial schools. Therefore, these results may not generalize beyond these populations. One major reason for approaching these schools was that they were more cooperative and willing to allow the students to participate voluntarily in the research. The public school authorities were not supportive in allowing their students to participate. Therefore, it was not feasible to conduct this study with students from public schools within a reasonable time.

It is not clear why the self-reflection intervention was not significant like the strategy training intervention. One possible reason is that the sample size was too small and with a larger sample size, it might be feasible to obtain statistically significant results.

Another reason might be that self-reflection training was not as robust enough to have a more substantial influence on performance outcomes.

Although, self-efficacy was significant, it exerted only a weak effect on math performance. One reason was that students in the non-strategy groups overestimated their performance, which inhibited the effects of self-efficacy judgments. Underconfidence measured by the calibration index of self-efficacy bias was significant, signifying that less bias improved students' math performance. Prior studies by Schunk and colleagues showed that self-efficacy was a strong predictor of academic achievement.

### *Educational Implications*

The need for new methods to improve the math performance of American students is widely recognized. The present study examined whether deficiencies in a number of key self-regulatory processes can lead students to overestimate or underestimate the accuracy of their academic performance. This study demonstrated that these key sources of inaccurate estimations could be improved through self-regulatory training, specifically strategy training. It was not possible in one hour to teach students the importance of strategy training. Therefore, classroom practices should incorporate strategy training in the context of a content area such as math.

Second, this study adds to the literature that children's self-perceptions of their capability influence their subsequent performance. When such perceptions are accurate, they lead to better performance. Self-efficacy and self-efficacy calibration measures are important metacognitive capabilities that predict performance. Teachers can be taught to incorporate these metacognitive procedures into math curricular experiences in order to help those students whose underconfidence and overconfidence may be interfering with

the acquisition of appropriate math skills and strategies. For example, a simple task might involve having students estimate how well they might do on a set of problems. After completing the problems, the teacher can show them how well they judged their capability. These practices will enable students to better assess their capability on a wide array of tasks.

Third, self-evaluations and self-evaluation calibration are critical metacognitive skills students must develop in order to succeed. Students who self-evaluate their performance accurately will do better on future tasks. Moreover, accurate self-evaluations were related to subsequent self-efficacy beliefs and math performance. It is simple to include self-evaluations in the classroom and frequent self-evaluations are beneficial. This can be done in various ways for example, checking one's work against a standard set of measures in the absence of external feedback.

Finally, this study suggests that self-regulatory training occurs in a cyclical interrelated manner and not as isolated components. Less bias and accurate estimations of one's capability led to better use of math strategy, which in turn improved math performance. Self-evaluation and self-reflection enabled students to evaluate their performance, which influenced self-efficacy judgments, thereby completing the cycle. Although, self-reflection training did not exert a strong influence as strategy training on the outcome measures, the results shows that it has the potential to improve students' learning as well. Schunk and Ertmer (2000) suggest that self-reflection may be more important where external evaluation or feedback is infrequent or when students face difficulties in learning. The challenge is to find ways to motivate students to engage in

self-reflection by themselves. Researchers suggest to teach students that self-reflection is similar to any other academic task that must be planned (Schunk & Ertmer, 2000).



3. From  $\frac{5}{8}$  yard length of cloth,  $\frac{3}{8}$  yard was cut. How much yard remained?
4. From a  $4\frac{1}{2}$  gallon container of milk, Sara used  $2\frac{1}{3}$  gallons to make a cake. How much milk remained?
5. A bag of flour weighed  $\frac{7}{8}$  pound. A chef used  $\frac{3}{4}$  pound to make a loaf of bread. How much flour remained?

## Appendix B

Name: \_\_\_\_\_

Date: \_\_\_\_\_

Age: \_\_\_\_\_

Gender Male Female

<b>Pretest</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>Math score</b>						

<b>Practice</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
Time (Seconds)				
Strategy use				
Math score				

<b>Posttest</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
Time (Seconds)						
Strategy use						
Math score						

**Pretest: Read instructions carefully.**

*Solve each problem in the space below. You may solve the problems in any order you desire.*

1. A bag of candy weighed  $3\frac{4}{5}$  ounce, Cindy ate  $1\frac{2}{3}$  ounce. How much remained?

**Answer** \_\_\_\_\_

2. Jim had  $\frac{13}{16}$  pound of coffee. After making a pot of coffee, he had  $\frac{6}{15}$  pound remaining. How much coffee did he use?

**Answer** \_\_\_\_\_

3. From  $\frac{12}{13}$  yard length of cloth,  $\frac{5}{8}$  yard was cut. How much yard remained?

**Answer** \_\_\_\_\_

4. From a  $10\frac{15}{28}$  gallon container of milk, Sara used  $7\frac{9}{14}$  gallons to make a cake. How much milk remained?

**Answer** \_\_\_\_\_

5. In a race, three top runners had times of  $12\frac{2}{3}$  seconds,  $12\frac{3}{8}$  seconds, and  $12\frac{2}{5}$  seconds. How much time (in seconds) separated the first and third runner?

**Answer** \_\_\_\_\_

6. A container had  $15\frac{4}{48}$  gallons of water. During the day, people drank  $7\frac{54}{102}$  gallons of water. How many gallons of water remained?

**Answer** \_\_\_\_\_



**Practice: You may do the problems in any order you desire.**

1. Tim had a bag of peanuts weighing  $3\frac{7}{9}$  pounds. He and his friends ate  $2\frac{4}{5}$  pounds of the peanuts. How many pound of peanuts remained?

**Answer** \_\_\_\_\_

2. From a bag of potato chips weighing  $\frac{11}{13}$  ounce, Molly ate  $\frac{15}{26}$  ounce. How many ounce of chips remained?

**Answer** \_\_\_\_\_

3. Three students finished a race in  $6\frac{4}{39}$  minutes,  $6\frac{2}{13}$ , and  $6\frac{10}{78}$ . How many minutes separated the first and third runners?

**Answer** \_\_\_\_\_

4. A box had  $\frac{27}{39}$  ounce of tortilla chips. At a picnic, children ate  $\frac{32}{56}$  ounce of chips. How much remained?

**Answer** \_\_\_\_\_













## Appendix C

1. From a  $9\frac{4}{48}$  gallon container of milk, Sara used  $6\frac{11}{12}$  gallons to make a cake. How much milk remained?

The following is a step-by-step strategy training method to solve the problem.

- Set up the problem  $9\frac{4}{48} - 6\frac{11}{12} =$
- Find the common denominator; divide both denominator and numerator  $\frac{4}{48}$  by 4 to get  $\frac{1}{12}$  or multiply both by 4 and get  $\frac{44}{48}$ .  $9\frac{1}{12} - 6\frac{11}{12}$
- If the second numerator is smaller, subtract the second whole number from the first.  
Next subtract the second numerator from the first.
- If the second numerator is bigger and it is in the above example, borrow one from the first whole number and add to the smaller numerator.  $9 - 1 = 8$ , which is  $\frac{12}{12}$  or  $\frac{48}{48}$ .  
Add 12 to 1 =  $\frac{13}{12}$  or 48 to 4 =  $\frac{52}{48}$ .
- Do step 3, first subtract the whole number  $8\frac{13}{12} - 6\frac{11}{12} = 2$ , then subtract the second numerator from the first one  $\frac{2}{12}$ .
- Answer is  $2\frac{2}{12}$  gallons remained.

2. From  $\frac{8}{17}$  yard length of cloth,  $\frac{37}{85}$  yard was cut. How much yard remained?

The following is a step-by-step strategy training method to solve the problem.

1. Set up the problem  $\frac{8}{17} - \frac{37}{85} =$
2. Find the common denominator. One way of doing it in this problem is to add 17 to itself a number of times and see if it eventually adds up to 85.  $17 + 17 = 34$ .  $34 + 34 = 68$ .  $68 + 17 = 85$ . Therefore,  $17 \times 5 = 85$ .
3. Convert to  $\frac{8}{17}$  a common denominator by multiplying  $17 \times 5$ . You have to also convert the numerator by multiplying by 5 =  $\frac{40}{85}$ .
4. Subtract  $\frac{40}{85} - \frac{37}{85} = \frac{3}{85}$  yard remained.
5. Answer is  $\frac{3}{85}$ .

**3. In a math test, the three top students finished in  $9\frac{3}{4}$  minutes,  $9\frac{5}{6}$  minutes, and  $9\frac{7}{9}$  minutes. How much minutes separated the first and third student?**

The following is a step-by-step strategy training method to solve the problem.

1. Read the problem carefully and notice that the whole numbers are alike. The difference will be in fraction minutes and not in whole minutes.
2. Set up the fractions  $\frac{3}{4}$      $\frac{5}{6}$      $\frac{7}{9}$
3. Find the common denominator. Here, you can multiply the largest denominator by the smallest ( $9 \times 4$ ) and get 36, which is the common denominator of all three numbers.
4. Convert the fractions with the new denominator. Divide 36 by each denominator and multiply each numerator by that number to get  $\frac{27}{36}$     $\frac{30}{36}$     $\frac{28}{36}$
5. Figure who finished first and who finished third. The first person finished in the least fraction time =  $\frac{27}{36}$ . The third person finished in the highest fraction time =  $\frac{30}{36}$ .
6. Subtract the smallest fraction from the largest.  $\frac{30}{36} - \frac{27}{36} = \frac{3}{36}$ .
7. Answer is  $\frac{3}{36}$ .

- 4. A box had  $\frac{72}{81}$  pound of cookies. For breakfast, some children ate  $\frac{60}{72}$  pound of cookies. How much pound of cookies remained?**

The following is a step-by-step strategy training method to solve the problem.

1. Set up the problem.  $\frac{72}{81} - \frac{60}{72} =$
  2. Simplify the fractions. Think of a number that would divide both numerator and denominator without a remainder. If you cannot think of a number then start with the smallest number.
  3. The first fraction has even and odd numbers so try 3.  $72 \div 3 = 24$  and  $24 \div 3 = 8$ ;  $81 \div 3 = 27$  and  $27 \div 3 = 9$ . Therefore,  $\frac{72}{81} = \frac{8}{9}$ .
  4. Reduce the second fraction.  $\frac{60}{72}$ , here both numbers are even. Find a number that would divide both without a remainder. If you cannot, start with the smallest 2.
  5.  $60 \div 2 = 30$  and  $30 \div 2 = 15$ . Likewise,  $72 \div 2 = 36$  and  $36 \div 2 = 18$ . Further simplify  $\frac{15}{18}$  by dividing both numbers by 3 to get  $\frac{5}{6}$ .
  6. The simplified fractions are  $\frac{8}{9} - \frac{5}{6}$ . Find the common denominator by multiplying both denominators to get 54.
  7. Convert both numerators by cross multiplying with the other denominator to get  $\frac{48}{54} -$
- $$\frac{45}{54} = \frac{3}{54}$$
8. Answer is  $\frac{3}{54}$

## Appendix D

### Self-Reflection Training

Self-reflection training occurred during the practice phase.

1. In the practice phase, students in each group solved four problems. After students solved each problem, every student got feedback regarding the correctness or incorrectness of the response.
2. However, for in the self-reflection groups, the researcher circled any error or errors in the response.
3. After circling the errors, the researcher told the student in the self-reflection condition “This is where you went wrong. What can you do to correct it?”
4. Next, the student then did the following problem, and if there were errors, the researcher repeated the same procedures in steps 2 and 3.
5. If there were no errors, the researcher told the student that the response was free from error and he/she may proceed to do the next problem.

## Appendix E

Date, 2008

Dear Ms. or Mr. (Principal/Assistant Principal)

I am a doctoral candidate in Educational Psychology at the Graduate Center, which is part of City University of New York (CUNY). Under the supervision of Dr. Barry Zimmerman, I am currently conducting research on the topic “Improving Elementary School Students’ Mathematics Self-Efficacy Calibration and Performance through Self-Regulation Training” for my dissertation.

A pilot study conducted in spring of 2007 focused on my dissertation topic which was: Investigating the Effects of Strategy Training on the Self-Evaluative Judgments of Middle School Students”. My interest in improving students’ accuracy judgments and performance on math division tasks (fractions) is what guides my work in this area of learning.

I propose to do the study individually with students and teach them various methods to solve math division problems more effectively. Students will be randomly assigned to groups and depending on their group, they will be taught the strategies at the beginning or at the end of the session. This entire session will take about sixty minutes. I am gathering a sample of approximately 84 fifth grade students for this study.

The Institutional Review Board of Graduate School and University Center, CUNY has approved this study. I have applied to the Proposal Review Committee, Department of Education, New York City for permission to conduct this study in the public school and I am currently awaiting a reply.

Thank you very much for considering my study. I know that the experience can only add to the wonderful work that you are doing to provide your students with quality learning experiences. My email address is: [dramdass@gc.cuny.edu](mailto:dramdass@gc.cuny.edu) and my cell phone # is: 347-239-9213.

Sincerely,

Darshanand Ramdass

## Appendix F

**PARENTAL/GUARDIAN CONSENT FORM**

My name is Darshanand Ramdass, and I am a Ph.D. student in the Educational Psychology Program at The Graduate Center of the City University of New York (CUNY). I am conducting research designed to improve the accuracy of students' judgments on math tasks and math solving skills. I would like your permission for your child to participate in this study. Separately, I will also ask your child for his/her consent to participate.

If you agree, your child will be escorted from his/her classroom during regular school hours to a separate room at school. In that setting, I will teach your child various strategies to analyze and solve math division problems and then give him/her an opportunity to practice this method. Depending on the assigned group, some students will be taught the strategy at the outset and some at the end of the study. I expect that this self-regulatory training will improve your child's math division skill and self-confidence. This entire session will take about sixty minutes. Your child's responses will only be seen by me and my advisor and participation is voluntary. ***At any time your child can refuse to answer any question or end his/her participation without any penalty. Participation or non-participation will have no effect on your child's grades.***

The risks from participating in this study are no more than encountered in everyday life. The benefit of your child's participation is that it may allow researchers and teachers to get a better understanding how to include strategy training to improve students' math performance and their accuracy judgments. .

If the results of the study are published, the names of the students or any identifying characteristics will not be used in any of the publications. If you would like a copy of the article, please provide me with your address and I will send you a copy in the future.

If you have any questions about this research, you can contact me at 347-239-9213 or dramdass@gc.cuny.edu, my advisor Professor Barry Zimmerman at (212) 817-8291, or BZimmerman@gc.cuny.edu. If you have questions about your child's rights as a participant in this study, you can contact Kay Powell, IRB Administrator at The Graduate Center at (212) 817-7525, kpowell@gc.cuny.edu.

I give permission for my child to participate in this study. (please circle one): Yes  
No

\_\_\_\_\_  
Child's name

\_\_\_\_\_  
Parent/Guardian's signature

\_\_\_\_\_  
Date

\_\_\_\_\_  
Investigator's signature

\_\_\_\_\_  
Date

Thank you for your allowing your child to participate in the study.

## Appendix G

**STUDENT ASSENT FORM**

My name is Darshanand Ramdass, and I am a Ph.D. student in the Educational Psychology Program at The Graduate Center of the City University of New York (CUNY). I am conducting a study to improve students' accuracy judgments on math tasks and math solving skills. I am asking your parents for their permission to allow you to participate in this study, but I also want your agreement.

If you agree, you will be escorted from your classroom during regular school hours to a separate room at your school. In that setting, I would teach you various methods to help you solve math division problems better. Depending on your assigned group, you will either be taught this strategy at the beginning or at the end of the study. This entire session will take about sixty minutes. Your answers will only be seen by me and my advisor. Your participation is voluntary. ***At any time you can refuse to answer any questions or decide to quit, and that decision will not affect your grades in school.***

There is no risk for you in this research, and I expect that your math division skill will improve as a result of this training. I also expect that the results will help teachers to instruct students better in the future. When I discuss or write about the results of this study, I will not use your name or anything that might identify you to others.

If you have any questions about this research, you can contact me at 347-239-9213 or dramdass@gc.cuny.edu, my adviser, Professor Zimmerman at (212) 817-8291 or BZimmerman@gc.cuny.edu. If you have questions about your rights as a participant in this study, you can contact Kay Powell, IRB Administrator at The Graduate Center (212) 817-7525, kpowell@gc.cuny.edu.

I agree to participate in this study. (please circle one): Yes      No

\_\_\_\_\_  
Student's signature

\_\_\_\_\_  
Date

\_\_\_\_\_  
Investigator's signature

\_\_\_\_\_  
Date

Thank you for your participation in this study.

## References

- A summary of findings from Program for International Student Assessment PISA. (2006). Retrieved December 25, 2007 from <http://nces.ed.gov/surveys/pisa/pisa2006highlights.asp>
- Alvarez, A., & Adelman, H. S. (1986). Overstatements of self-evaluations by students with psychoeducational problems. *Journal of Learning Disabilities, 19*(9), 567-571.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.
- Bandura, A. (1986). *Social foundations of thought and action. A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: New Age.
- Bandura, A., & Adams, N. E. (1977). Analysis of self-efficacy theory of behavioral change. *Cognitive Therapy and Research, 1*, 287-308.
- Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45*, 1017-1028.
- 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.
- Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *Journal of Experimental Education, 69*(2), 133-151.
- 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. *The Journal of Experimental Education, 73*, 269-290.
- Bong, M. (1997). Generality of academic self-efficacy judgments: Evidence of hierarchical relations. *Journal of Educational Psychology, 89*, 696-709.
- Bong, M. (2006). Asking the right question. How confident are you that you could successfully perform these tasks? In F. Pajares, & T. Urdan (Eds.), *Self-efficacy*

- beliefs of adolescents* (pp. 287-303). Greenwich, CT: New Age.
- Bong, M., & Hocevar, D. (2002). Measuring self-efficacy: Multi-trait multi-method comparison of scaling procedures. *Applied Measurement in Education, 15*, 143-171.
- Borkowski, J. G. (1992). Metacognitive theory: A framework for teaching literacy, writing, and math skills. *Journal of Learning Disabilities, 25*, 253-257.
- Bouffard-Bouchard, T., Parent, S., & Larivee, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students. *International Journal of Behavioral Development, 14*, 153-164.
- Brannick, M. T., Miles, D. E., & Kisamore, J. L. (2005). Calibration between student mastery and self-efficacy. *Studies in Higher Education, 30*, 473-483.
- Brown, S. D., & Lent, R. W. (2006). Preparing adolescents to make career decisions: A social cognitive perspective. In F. Pajares, & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 201-223). Greenwich, CT: New Age.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245-281.
- Chen, P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences, 14*(1), 79-92.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collins, J. L. (1982, March). *Self-efficacy and ability in achievement behavior*. Paper presented at the Annual Meeting of the American Educational Research Association, New York.
- Damon, W., & Hart, D. (1982). The development of self-understanding from infancy through adolescence. *Child Development, 53*, 841-864.
- Dempster, A. C. (1995). General Problem Solving Strategies. In E. Wood, V. E. Woloshyn, & T. Willoughby (Eds.), *Cognitive Strategy Instruction for Middle and High Schools* (pp. 117-136). Dexter, Michigan: Thomson-Shore, Inc.
- Everyday Learning Corporation. (2004). *Everyday Mathematics: Teacher's Reference Manual: Grades 4-6*. Chicago, IL: Wright Group/McGraw Hill.
- 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(4), 221-230.
- Garavalia, L., Olson, E., & Comeau, S. (2005). Achievement calibration and causal attributions. *Academic Exchange Quarterly*, 9(2), 99-103.
- Gaskill, P. J., & Murphy, P. K. (2004). Effects of a memory strategy on second-graders' performance and self-efficacy. *Contemporary Educational Psychology*, 29, 27-49.
- Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, 116, 119-136.
- Hacker, D. J., & Bol, L. (2004). Metacognitive theory. Considering the social-cognitive influences. In D. M. McInerney & S. V. Etten (Eds.), *Big Theories Revisited* (pp 275-297). Greenwich, Connecticut: Information Age Publishing, Inc.
- Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92, 160-170.
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy mathematics performance correspondence. *Journal for Research in Mathematics Education*, 20, 261-273.
- Jamit Fractions, Worksheet 6. (<http://jamit.com.au/fractions.htm>). Retrieved 5-7-2008 from <http://jamit.com.au/worksheets/fraction6.php>
- Jöreskog, K. G., & Sörbom, D. (2006). LISREL 8.80 Student Edition. Scientific Software Lincolnwood, IL: International Inc.
- Keren, G. (1991). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica*, 77(3), 217-273.
- Klassen, R. (2002). A question of calibration: A review of the self-efficacy beliefs of students with learning disabilities. *Learning Disability Quarterly*, 25(2), 88-102.
- Klassen, R. M. (2006). Too much confidence? The self-efficacy of adolescents with learning disabilities. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 181-200). Greenwich, CT: New Age.
- Klassen, R. M. (2007). Using predictions to learn about the self-efficacy of early adolescents with and without learning disabilities. *Contemporary Educational Psychology*, 32, 173-187.
- Lee, J., Grigg, W., and Dion, G. (2007). *The Nation's Report Card: Mathematics 2007*

Retrieved December 25, 2007 from  
<http://nces.ed.gov/nationsreportcard/pubs/main2007/2007494.asp>

- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.
- McCombs, B. L. (1989). Self-regulated learning and academic achievement: A phenomenological view. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, Research, and Practice*, (pp. 51-82). New York: Springer.
- Montague, M., & Bos, C. S. (1986). The effect of cognitive strategy training on verbal math problem solving performance of learning disabled adolescents. *Journal of Learning Disabilities*, 19(1), 26-33.
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38, 30-38.
- National Council of Teachers of Mathematics (2008). Principles and standards for school mathematics. Overview: Standards for grades 3-5. Retrieved February 1, 2008 from <http://standards.nctm.org/document.chapter5.index.htm>
- National Mathematics Advisory Panel. (2008). *Foundations for success. The final report of the National Mathematics Advisory Panel*, U.S. Department of Education: Washington, DC. Retrieved March 19, 2008 from <http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf>
- 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(1), 7-28.
- Pajares, F. (1996a). Self-efficacy in academic settings. *Review of Educational Research*, 66(4), 543-578.
- Pajares, F. (1996b). Self-efficacy beliefs and mathematical problem solving of gifted students. *Contemporary Educational Psychology*, 21, 325-344.
- Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology*, 24, 124-139.
- Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development*, 33, 214-221.

- Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental abilities in mathematical problem-solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.
- Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem-solving. A path analysis. *Journal of Educational Psychology, 86*, 193-203.
- Pajares, F., & Miller, M. D. (1997). Mathematics self-efficacy and mathematical problem solving: Implications of using different forms of assessment. *The Journal of Experimental Education, 65*(3), 213-228.
- Pintrich, P. R. & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.
- Schraw, G. (1995). Measures of feeling-of-knowing accuracy: A new look at an old problem. *Applied Cognitive Psychology, 9*(4), 321-332.
- Schraw, G., Potenza, M. T., & Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology, 18*, 455-463.
- Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology, 73*, 93-105.
- Schunk, D. H. (1982). Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *Journal of Educational Psychology, 74*, 548-556.
- Schunk, D. H. (1983). Progress self-monitoring: Effects on children's self-efficacy and achievement. *Journal of Experimental Psychology, 51*, 89-93.
- Schunk, D. H. (1985). Participation in goal-setting: Effects on self-efficacy skills and skills of learning disabled children. *Journal of Special Education, 19*, 307-317.
- Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*(2), 359-382.
- Schunk, D. H. (1998). Teaching elementary students to self-regulate practice of mathematical skills with modeling. In D.H. Schunk & B.J. Zimmerman (Eds.), *Self-regulated learning. From teaching to self-reflective practice* (pp. 137-159). New York, NY: The Guilford Press.
- Schunk, D. H., & Cox, P. D. (1986). Strategy training and attributional feedback with learning disabled students. *Journal of Educational Psychology, 78*, 201-209.
- Schunk, D. H., & Ertmer, P. A. (1999). Self-regulatory processes during computer skill

- acquisition: Goal and self-evaluative influences. *Journal of Educational Psychology, 91*, 251-260.
- Schunk, D. H., & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds.), *Handbook of self-regulation* (pp. 631-649). San Diego, CA: Academic Press.
- Schunk, D. H., & Gunn, T. P. (2001). Self-efficacy and skill development: Influence of task strategies and attributions. *Journal of Educational Research, 79*, 238-244.
- Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology, 77*, 313-322.
- Schunk, D. H., Hanson, A. R., & Cox, P. (1987). Peer model attributes and children's achievement behaviors. *Journal of Educational Psychology, 79*, 54-61.
- Schunk, D. H., & Pajares, F. (2004). Self-efficacy in education revisited. Empirical and applied evidence. In McInerney, D. M., & Etten, S. V. (Eds.), *Big Theories Revisited* (pp 115-138). Greenwich, Connecticut: Information Age Publishing, Inc.
- Schwartz, J. E. (2008). *Elementary mathematics: Pedagogical content knowledge*. Boston, MA: Pearson Education, Inc.
- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*, 240-261.
- Stone, N. J. (2002). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12*, 437-478.
- Trends in International and Science Study TIMSS (2007). Mathematics achievement of fourth- and eighth-graders in 2007. Retrieved January 31, 2009 from [http://nces.ed.gov/timss/results07\\_math07.asp](http://nces.ed.gov/timss/results07_math07.asp)
- Tuckman, B. W., & Sexton, T. L. (1990). The relation between self-beliefs and self-regulated performance. *Journal of Social Behavior and Personality, 5*, 465-472.
- Wilhite, S. C. (1990). Self-efficacy, locus of control, self-assessment of memory skill, and study activities as predictors of college course achievement. *Journal of Educational Psychology, 82*, 696-700.
- Yates, J. F. (1990). *Judgments and decision making*. Englewood Cliffs, NJ: Prentice Hall.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychology, 25*(1), 3-17.

- Zimmerman, B.J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp.202-231). New York, NY: Cambridge University Press.
- Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D.H. Schunk & B.J. Zimmerman (Eds.), *Self-regulated learning. From teaching to self-reflective practice* (pp. 1-19). New York, NY: The Guilford Press.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Self-regulation: Theory, research, and applications* (pp.13-39). Orlando, FL: Academic Press.
- Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement. Theoretical perspectives* (pp. 1-37). Mahwah, NJ: Lawrence Erlbaum Associates.
- Zimmerman, B. J. (2002). Achieving self-regulation: The trial and triumph of adolescence. In F. Pajares & T. Urdan (Eds.), *Academic motivation of adolescents* (vol.2, pp.1-27). Greenwich, CT: Information Age.
- Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal*, 31, 845-862.
- Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29, 663-676.
- 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., & Kitsantas, A. (1996). Self-regulated learning of a motoric skill: The role of goal setting and self-monitoring. *Journal of Applied Sport Psychology*, 8, 60-75.
- Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process goals to outcome goals. *Journal of Educational Psychology*, 89, 29-36.
- Zimmerman, B. J., & Kitsantas, A. (1999). Acquiring writing revision skill: Shifting from process to outcome self-regulatory goals. *Journal of Educational Psychology*, 91, 241-250.

Zimmerman, B. J., Kitsantas, A. (2005). The hidden dimension of personal competence. In A. J. Elliot, & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 509-526). New York, NY: The Guilford Press.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.

Zimmerman, B. J., & Martinez-Pons, M. (1992). Perceptions of efficacy and strategy use in the self-regulation of learning. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 185 – 207). Hillsdale, NJ: Erlbaum.