Effects of evaluative feedback on math self-efficacy, grade self-efficacy, and math achievement of ninth grade algebra students: a longitudinal approach.

Deborah June Burnett Thompson
University of Louisville

Follow this and additional works at: http://ir.library.louisville.edu/etd

Recommended Citation
https://doi.org/10.18297/etd/1434

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.
EFFECTS OF EVALUATIVE FEEDBACK ON MATH SELF-EFFICACY, GRADE SELF-EFFICACY, AND MATH ACHIEVEMENT OF NINTH GRADE ALGEBRA STUDENTS: A LONGITUDINAL APPROACH

By

Deborah June Burnett Thompson
B.S., University of Kentucky, 1978
M.Ed., University of Louisville, 1983

A Dissertation
Submitted to the Faculty of the Graduate School of the University of Louisville in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

College of Education and Human Development
University of Louisville
Louisville, Kentucky

May 2007
EFFECTS OF EVALUATIVE FEEDBACK ON MATH SELF-EFFICACY, GRADE SELF-EFFICACY, AND MATH ACHIEVEMENT OF NINTH GRADE ALGEBRA STUDENTS: A LONGITUDINAL APPROACH

By

Deborah June Burnett Thompson
B.S., University of Kentucky, 1978
M.Ed., University of Louisville, 1983

April 5, 2007

by the following Dissertation Committee:

Dissertation Director, Dr. Namok Choi

Dr. William S. Bush

Dr. W. Blake Haselton

Dr. John L. Keedy

Dr. Joseph M. Petrosko
DEDICATION

This dissertation is dedicated to my husband, Rock, and my children, Christopher, Katie, and Benjamin Thompson. Their unending faith gave me the spirit and drive to accomplish this goal.
ACKNOWLEDGMENTS

I would like to thank my mentor and advisor, Dr. Namok Choi, for her guidance and countless hours of work. Her enthusiasm for my ambitious endeavors was much appreciated. I would also like to thank the members of my committee, Dr. William Bush, Dr. Blake Haselton, Dr. John Keedy, and Dr. Joseph Petrosko, for their valuable insight and comments. Their time and knowledge added significantly to the quality of this study.

I am very grateful to the Scottish Rite Foundation in Kentucky for their confidence in my abilities to be successful and their generous financial support throughout my doctoral program. This study would not have been accomplished without the assistance of many individuals from Oldham County Schools. In particular, I would like to express my thanks to Dr. David Weedman, Carol Jones, and Brent Deaves of Oldham County High School, whose support made this study possible.

I have a deep gratitude to extend to my friends, Diane Belt, Susan Hack, and Andrea Neikirk, who were so supportive, especially near the end when things were overwhelming. I want to extend a very special thanks to Julie Hardin who provided a constant source of comfort and pushed me when I tried to give up.

I would also like to thank my parents, Harold and Marlene Burnett, and my children, Christopher, Katie, and Benjamin, for their help, patience, and understanding throughout this process. Most importantly, I would like to thank my husband, Rock, for his patience with my passion for the pursuit of this goal. Without his support and love, I would not have achieved it.
ABSTRACT

EFFECTS OF EVALUATIVE FEEDBACK ON MATH SELF-EFFICACY, GRADE SELF-EFFICACY, AND MATH ACHIEVEMENT ON NINTH GRADE ALGEBRA STUDENTS: A LONGITUDINAL APPROACH

Deborah J. B. Thompson

May 12, 2007

The double-blind study used two intact ninth grade high school Algebra One classes for a total sample size of 46. Within each of the classes the participants were assigned randomly to one of three treatment groups: self-referenced feedback, social-referenced feedback, and a control group. Self-referenced feedback compared the student’s performance to his or her previous performances. Social-referenced feedback compared the student’s performance to that of the other students in the class. The control group did not receive comparative feedback but innocuous comments such as “study your notes before class.”

Six measurements of math self-efficacy, grade self-efficacy, and math achievement were taken over a five-month interval. Research questions investigated how students’ math self-efficacy, grade self-efficacy, and math achievement changed over time; and whether students’ math self-efficacy, grade self-efficacy, and math achievement varied by treatment group over time. Results of multilevel analysis indicated that all three of the dependent variables changed over time with significant p-values for the intercepts and slopes using a linear unconditional growth model. No
significant differences in math or grade self-efficacy were found by treatment group, however, the self-referenced group exhibited linear growth in math self-efficacy over the last four measurements. The social-referenced feedback group showed a growth pattern similar to that of the control group for math self-efficacy but had the lowest math self-efficacy rating on the last measurement. Mathematical models of grade self-efficacy indicated that the self-referenced feedback group had higher certainties for achieving higher grades than the other two feedback groups even though the results were not significant. The social-referenced feedback group did not exhibit different grade self-efficacy for achieving higher grades in the course from the control group.

Math achievement changed significantly by treatment group. While this study indicated only minimal improvements in the Diagnostic Exam scores, the self-referenced feedback treatment group exhibited a rate of change more than five times that of the control group, and more than one and a half times that of the social-referenced feedback treatment group.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ iv
ABSTRACT ....................................................................................................................... v
LIST OF TABLES .............................................................................................................. ix
LIST OF FIGURES ............................................................................................................ x

CHAPTER

I. INTRODUCTION ........................................................................................................ 1
   Statement of the Problem ......................................................................................... 7
   Research Questions ................................................................................................. 8
   Purpose of the Study ............................................................................................... 8
   Limitations ............................................................................................................. 9

II. LITERATURE REVIEW ............................................................................................ 11
   Social Cognitive Learning Theory .......................................................................... 11
   Self-Efficacy ......................................................................................................... 19
      Self-Efficacy and Academic Achievement ......................................................... 26
   Measurement of Self-Efficacy ............................................................................... 30
   Domain Specific Self-Efficacy and Academic Achievement .............................. 33
   Sources of Self-Efficacy ....................................................................................... 40
   Self-Efficacy and Evaluative Feedback ................................................................. 49
   Self-Efficacy and Goal Setting ............................................................................ 55
   Demographic and Psychological Variables ......................................................... 63
   Summary ............................................................................................................. 69
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>III.</td>
<td>METHODS</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Purpose</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Participants</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Design</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Variables</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Instruments</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Procedure</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Threats to Experimental Validity</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>96</td>
</tr>
<tr>
<td>IV.</td>
<td>RESULTS</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Preliminary Statistical Analysis</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Reliability Analyses</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Descriptive Statistics</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Intercorrelations</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Analyses for Research Questions</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>146</td>
</tr>
<tr>
<td>V.</td>
<td>DISCUSSION</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>Implications</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>Limitations</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>Recommendations</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>161</td>
</tr>
<tr>
<td></td>
<td>APPENDICES</td>
<td>172</td>
</tr>
<tr>
<td></td>
<td>CURRICULUM VITAE</td>
<td>234</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comparison of Reliability Coefficients</td>
<td>99</td>
</tr>
<tr>
<td>2. Test-Retest Reliability Estimates for BSRI and CSES</td>
<td>101</td>
</tr>
<tr>
<td>3. Descriptive Statistics for BSRI, CSES, and CATS Math</td>
<td>103</td>
</tr>
<tr>
<td>4. Descriptive Statistics for Math Self-Efficacy by Treatment Group</td>
<td>105</td>
</tr>
<tr>
<td>5. Descriptive Statistics for the Diagnostic Exams by Treatment Group</td>
<td>107</td>
</tr>
<tr>
<td>6. Descriptive Statistics for Grade Self-Efficacies by Treatment Group</td>
<td>110</td>
</tr>
<tr>
<td>7. Intercorrelations of the Variables</td>
<td>114</td>
</tr>
<tr>
<td>8. Intercorrelations of Femininity and Masculinity with Math Self-Efficacy Ratings and Diagnostic Exams</td>
<td>116</td>
</tr>
<tr>
<td>9. Within-Person Ordinary Least Squares Regression Models for Math Self-Efficacy</td>
<td>123</td>
</tr>
<tr>
<td>10. HLM Results of Unconditional and Conditional Growth Models for Math Self-Efficacy</td>
<td>126</td>
</tr>
<tr>
<td>11. Within-Person Ordinary Least Squares Regression Models for Grade Self-Efficacy</td>
<td>133</td>
</tr>
<tr>
<td>12. HLM Results of Unconditional and Conditional Growth Models for Grade Self-Efficacy</td>
<td>136</td>
</tr>
<tr>
<td>13. Within-Person Ordinary Least Squares Regression Models for Diagnostic Exams</td>
<td>140</td>
</tr>
<tr>
<td>14. HLM Results of Unconditional and Conditional Growth Models for Diagnostic Exam</td>
<td>142</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Illustration of triadic reciprocal interaction, according to Bandura (1977, p. 10).</td>
<td>12</td>
</tr>
<tr>
<td>2. Interaction between self-efficacy perceptions and outcome perceptions</td>
<td>20</td>
</tr>
<tr>
<td>3. Math self-efficacy means by treatment group</td>
<td>106</td>
</tr>
<tr>
<td>4. Diagnostic exam means by treatment group</td>
<td>109</td>
</tr>
<tr>
<td>5. Grade self-efficacy rating means by treatment group</td>
<td>111</td>
</tr>
<tr>
<td>6. Mean grade self-efficacy ratings by treatment group</td>
<td>112</td>
</tr>
<tr>
<td>7. Line graphs for individual students for math self-efficacy ratings</td>
<td>121</td>
</tr>
<tr>
<td>8. Regression models of math self-efficacy</td>
<td>124</td>
</tr>
<tr>
<td>9. Line graphs for individual students for grade self-efficacy ratings</td>
<td>132</td>
</tr>
<tr>
<td>10. Regression models of grade self-efficacy</td>
<td>134</td>
</tr>
<tr>
<td>11. Line graphs for individual students for grade self-efficacy ratings</td>
<td>138</td>
</tr>
<tr>
<td>12. Regression models of diagnostic exam</td>
<td>141</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

Mathematics is an academic filter. Great technological advances in the late 20th and early 21st centuries and economic globalization have increased the need for a more educated workforce, yet American public education has fallen behind other countries on mathematics and science achievement since the 1970s (Alexander & Pallas, 1984). American education has always been under public scrutiny. A prominent example was the publication of the National Commission on Excellence in Education’s report, *A Nation at Risk: The Imperative of Educational Reform* in 1983. The report stated that American high school graduates were not adequately prepared for the job market (Alexander & Pallas, 1984). In response to the report, public school systems have been increasingly held accountable for student achievement with many states implementing high-stakes tests to assess student learning. Indeed, the United States has made little progress in the last few decades in student achievement in mathematics and science despite the alarming tone of *A Nation at Risk*.

Subsequent reports that have appeared since *A Nation at Risk* (National Commission on Excellence, 1983) was published have emphasized the low international ranking of the United States in mathematics and science achievement. In particular, *America 2000: An Education Strategy Sourcebook* (U.S. Department of Education, 1991) stated that the United States ought to be ranked first in the world in mathematics
and science (Goldschmidt & Eyermann, 1999). Disappointingly, however, the United States ranked in the middle of the 38 countries that participated in the third Trends in International Math and Science Study (TIMSS; U.S. Department of Education, 1999) in 1995. At the secondary level of the 1995 TIMSS, for instance, the United States was ranked 19th of 21 participating countries for general math and science knowledge, outperforming only Cyprus and South Africa. For advanced mathematics, the United States ranked 15th out of 16 participating countries, with no countries scoring significantly lower than the United States. The United States fared better at the elementary level in the TIMSS for both 1995 and 2003 (U.S. Department of Education, 1999, 2004).

Even though four core subjects (reading, social studies, math, science) are commonly measured in the United States, math receives the most attention by researchers and critics. By far the most important reason for the attention to math achievement scores is the strong correlation between math and educational attainment and subsequently, to more promising career opportunities (Murnane, Willett, & Levy, 1995). Murnane et al. (1995) found in a longitudinal study of American high school seniors that mathematics scores in high school were more predictive of salaries of 1980 high school graduates than 1972 graduates. Analysis completed in the study indicated that a 1980 high school graduate with strong math skills earned less than a 1972 high school graduate with weak math skills, indicating that to earn the same amount as eight years earlier required higher math skills. This illustrated the growing importance of math skills and potential earnings. The researchers identified a widening gap in basic skills between the college and noncollege groups that they suggested was partially responsible for the wage
gap between the two groups as well (Murnane et al., 1995). It is commonly held that educational attainment is directly correlated with earning potential. If mathematics is an academic filter, then educational attainment may be at risk for students who perform poorly in mathematics.

Not only math achievement but also math self-efficacy has been reported to be significantly related to career choices students make (Betz & Hackett, 1983; Lent, Brown, & Larkin, 1984, 1986, 1987; Lent, Lopez, & Bieschke, 1991, 1993; Lopez & Lent, 1992; Luzzo, Hasper, Albert, Bibby, & Martinelli, 1999). Although biological gender has been often pointed out to be a factor in discrepancy in math performance, math self-efficacy helps explain the difference in college majors and career choices more than biological gender (Betz & Hackett, 1981; Hackett, 1985; Lent, Lopez, & Bieschke, 1991; Pedro, Wolleat, Fennema, & Becker, 1981; Zeldin & Pajares, 2000). In addition to math self-efficacy, prior math exposure is another factor that weakens the biological gender effect. No gender differences in math self-efficacy were found when this difference was controlled for in statistical analyses (Fennema & Sherman, 1977; Sherman & Fennema, 1977). Many college admission initiatives are aimed at encouraging females to enroll in science or math related majors, such as engineering. The gap between math achievement and career choice is of American public interest due to the shortage of qualified individuals for higher-level occupations that require strong mathematics skills.

The public will not disagree that mathematics is an important skill. What remains in debate is what determines whether or not students are successful at mathematics. Many external factors have been shown to contribute to mathematics achievement, some of which are not controlled by the students. School districts decide about spending per
pupil, for example, while teachers manage the classroom environment and implement a curriculum, and family environment provides an overall framework for the individual pupil. It is well established, however, that the pupil’s beliefs about specific abilities actually have the greatest predictive power about achievement (Campbell & Hackett, 1986; Hackett & Betz, 1989; House, 1995; Lent, Lopez, & Bieschke, 1993; Pajares, 1996a; Pajares & Graham, 1999; Pajares & Miller, 1994, 1995). Rodriguez (2004) summarily concluded, based on a comparison of the results of the TIMSS in relation to classroom assessment, that teachers should be encouraged to develop math self-efficacy in their students due to its positive relationship with performance.

In addition to the public need for qualified individuals for the American labor force, the policy of the *No Child Left Behind Act of 2001* (NCLB, 2001) also stimulated the importance of raising achievement levels of all students. The *No Child Left Behind Act of 2001* (NCLB, 2001) stipulates that school systems must decrease the achievement gap between student groups, and in particular, the gaps between minority and nonminority groups, and disadvantaged and nondisadvantaged groups. While academically talented students have made great strides, recent emphasis in education reform has focused on students who do not achieve minimal state standards. This has led to mounting pressure on school administrators and faculty to raise student outcomes to the point that, in some states, sanctions are levied against schools that do not exhibit minimum progress. The ability of a school district to increase student achievement has its limitations. One significant predictor of academic achievement is self-efficacy, defined as the perception one has about his or her capabilities to perform a given task (Bandura, 1977a). Several studies have shown that self-efficacy is the most influential
factor in predicting student achievement (Bandura, 1993; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Bong, 2004; Multon, Brown, & Lent, 1991; Zimmerman, Bandura, & Martinez-Pons, 1992).

As a striking example of self-efficacy as a significant predictor of academic performance, Pajares and Kranzler (1995) found that math self-efficacy was a stronger predictor of math problem solving performance than general mental ability. The effect of self-efficacy on academic performance can be attributed to particular characteristics of levels of self-efficacy. Individuals with low self-efficacy, for example, will tend to be less persistent on difficult tasks, use fewer attempts, and attribute their lack of success to a lack of capabilities. Those with high self-efficacy, on the other hand, are more persistent on difficult tasks, use a variety of strategies to solve the problem, and tend to attribute lack of success to a lack of effort rather than incapability. These characteristics are formed by a variety of factors, including perceptions of past performance accomplishments, outcome judgments, and self-set standards (Bandura, 1977a).

One of the factors affecting self-efficacy in academic settings is feedback. An important component in the formation of one’s self-efficacy for a given task is the judgment of past accomplishments against a standard. The standard can be self-set or based on normative behaviors, but in either case the individual perceives some type of information regarding behavior. The information regarding performance or behavior is called feedback. In a classroom setting there are many types of feedback for students such as verbal comments from peers and teachers and nonverbal items such as grades. Negative feedback, as in failing grades, will eventually lead to a perception of lesser capabilities to perform while positive feedback, as in high grades, will eventually lead to

Studies investigating the effects of different types of feedback show that specific types of evaluative feedback, that is, feedback that evaluates a performance, enhances academic achievement, directly and indirectly through self-efficacy (Schunk, 1982, 1983a, 1984b, Schunk & Cox, 1986; Shih & Alexander, 2000). Furthermore, feedback that compares an individual’s performance to his past accomplishments, called self-referenced feedback, produces greater influence on self-efficacy and academic achievement than feedback that compares an individual’s performance to the accomplishments of his peers (Shih & Alexander, 2000).

In addition to receiving evaluative feedback about performances or behaviors, individuals must have an idea of a goal against which to judge the feedback. Goals can be an effective strategy to increase academic achievement and can be focused on specific tasks or more general such as grade goals (Bandura & Cervone, 1983, 1986; Bandura & Schunk, 1981; Kitsantas, Reiser, & Doster, 2004; Page-Voth & Graham, 1999; Schunk, 1984a, 1985, 1996; Schunk & Gaa, 1981; Schunk & Swartz, 1993; Zimmerman & Bandura, 1994; Zimmerman, Bandura, & Martinez-Pons, 1992). Effective goals are proximal in nature to the students; for example, goals nearer in time rather than long-term are more effective. Specific goals are more effective because they are easier to measure and to judge as to their accomplishment. It is easier to judge, for instance, if 40 homework problems were completed accurately than it is to judge if one did “better” in math class. On the other hand, if goals are not sufficiently difficult, they will not be effective. Setting lower goals does not lead to higher achievement nor does setting nearly
impossible goals. Studies have also shown that goals are more effective in increasing self-efficacy if they are self-set rather than prescribed (Pajares, 2002).

Statement of the Problem

Given the importance of math achievement to successful educational pursuits and subsequent career opportunities, the search for improving math achievement continues to garner interest. Math self-efficacy has been demonstrated as one of the strongest predictors of math achievement; however, the need to sort out causal relationships among psychological and educational variables clearly exists. As the literature repeatedly pointed out, goal setting on the students’ part undoubtedly enhances students’ motivation to perform better. On the teachers’ part, the kind of feedback teachers provide during the learning process also influences students’ efficacy to learn and to perform better. There is growing evidence that self-referenced feedback seems to promote students’ learning more effectively (Shih & Alexander, 2000), although the evidence for the effectiveness of appropriate feedback in high school math classes is minimal. Most of the evidence of the effect of type of feedback on academic achievement is either from an elementary setting (Schunk, 1981, 1982, 1983a, 1983b, 1984b, Schunk & Gunn, 1986; Shih & Alexander, 2000) or from studies on writing essays (Schunk & Swartz, 1993). Furthermore, most studies on the effect of feedback on self-efficacy or on math achievement measured self-efficacy or math achievement at one point in the research process or at most two points. No study has reported on patterns in multiple measurements of self-efficacy and of math achievement or if there is any pattern difference in multiple measurements of self-efficacy or of math achievement given different types of feedback on students’ learning in a high school math class.
Research Questions

Based on a review of the literature, six questions are formulated as shown below

1. How does each student’s math self-efficacy rating change over time?

2. Do changes in students’ math self-efficacy ratings vary by feedback treatment?

3. How does each student’s grade self-efficacy rating change over time?

4. Do changes in students’ grade self-efficacy ratings vary by feedback treatment?

5. How does each student’s math achievement (measured by the Diagnostic Exam) change over time?

6. Do changes in students’ math achievement (measured by the Diagnostic Exam) vary by feedback treatment?

Purpose of the Study

There were three purposes to this study. One purpose was to use goal setting as a constant and manipulate the type of feedback to increase math self-efficacy. Based on findings in the literature, increases in math self-efficacy should be followed by increases in math achievement. This study used a quasi-experimental design approach with a control group and two different types of feedback specifically designed to increase math self-efficacy. Three dependent measures of math self-efficacy, grade self-efficacy, and math achievement were assessed for magnitudes of increase over time. An additional purpose of this study were the multiple measurements of math self-efficacy, grade self-efficacy, and math achievement to examine trends in the longitudinal data, which have not been reported in the literature. Finally, this study aimed to extend the literature by
using a high school sample; unlike most of the samples in the literature that used either
elementary school children or college age participants.

Limitations

This study had some limitations reducing the generalizability of the findings but
could provide direction for future research and investigation. Due to the structure of the
school system, the design was quasi-experimental rather than true experimental. It is not
ethical to assign high school students randomly to a particular treatment. Using two
established classes caused the loss of randomization and reduced the statistical power of
the study. There was not any reason to believe, however, that the two classes were
inherently different from one another and this was verified using tests of homogeneity.
Another limitation to the generalizability of the findings is the small sample size. Using a
small sample reduced the power of the statistical tests but this was an unavoidable
limitation when intact classes were selected. Statistical procedures involving repeated
measures, however, can reduce the effect of the small sample size by increasing power.
Another possible consequence of having a small sample size for a longitudinal study for
six months is the possibility of participant mortality. If students move out of the school
district or drop out of the study, the already small sample size could be further impacted,
but this study only lost one participant halfway through and another participant left three
weeks prior to the end of the study.

Additionally, the literature findings indicated that math self-efficacy cannot
override the lack of basic skills. Ninth-grade Algebra One is the lowest level of
mathematics offered at the high school where the study was conducted and so a lack of
basic skills may have been present among the participants in the study. Increasing math
self-efficacy in an algebra one course is difficult if students lack basic prerequisite skills such as addition and subtraction facts. Also, there was not an available standardized math achievement assessment for Algebra One. To avoid possible bias in using a teacher-made test for assessing math achievement, a newly developed state level diagnostic exam for Algebra One was used.
Several areas of research are relevant to the task of improving student achievement and reducing achievement gaps. These include social learning theory and self-efficacy research. This literature review begins with an overview of social cognitive learning theory and how it relates to self-efficacy and learning in children. The second section covers the literature dealing with different aspects of self-efficacy, including the relationship with academic achievement, sources of self-efficacy, feedback, goal setting, and specificity of self-efficacy. In particular, mathematics self-efficacy and its relation to mathematics achievement is reviewed. Of special interest is the impact of goal setting and feedback on math self-efficacy and mathematics achievement.

Social Learning Theory

As the study of human behavior progressed, a primary goal of predicting the actions of people emerged. One conceptual model proposed was unidirectional where behavior (B) is a function of a person’s cognitive and other perceptions (P) and his external environment (E). This is commonly illustrated as $B = f(P, E)$, which can be interpreted as “Behavior is a function of the Person and his Environment.” This model did not allow for the idea of an interaction between the Person and the Environment, hence, a revised conceptual model was developed. A subsequent conceptual model allowed for some flow between the person and the environment, but was still
unidirectional with regard to the behavior. This model is considered partially bi-
directional and can be illustrated using arrows to represent the interactions:

\[ B = f(P \xrightarrow{} E) \]. Finding this model of human interaction lacking, Bandura (1977b, 1978) proposed a system that differed from previously suggested models of human interaction. Human interaction, according to Bandura (1977b, 1978), involves a triadic reciprocal interaction among the three influences. Bandura (1977b, 1978) defines social cognitive learning theory as the continuous interaction between the cognitive, behavioral, and environmental determinants that explain human behavior. The interactions are reciprocal and continuous among all three components and may be described in terms of the person (P), the behavior (B), and the environment (E) (Bandura, 1977b, 1978).

Bandura (1977b, p. 10) illustrated the relationship between these three determinants as shown in Figure 1. Humans are not totally immune to the effects of their environment nor are they totally free from its effects (Bandura, 1977b). Behavior varies depending upon the individual person and the environmental factors present (Bandura, 1977b, 1978). Referring to this triadic interaction as reciprocal determinism, Bandura (1977b, 1978) defined determinism as the production of effects by events, and suggested that it is the interaction between the self-system of the person, their behavior, and the effects of

---

*Figure 1.* Illustration of triadic reciprocal interaction, according to Bandura (1977b, p.10). B = Behavior, P = Person, E = Environment.
the environment that determines whether a learned behavior is exhibited at a later date. At the center of this theory is the self-system, theorized not as a psychic agent that controls behavior but rather as an accumulation of the cognitive structures, mechanisms, and subfunctions that humans use as they interact with the environment (Bandura, 1978). The human self-system has three specific abilities that allow individuals to gain understanding and knowledge for use in learning new behaviors and the parameters governing the exhibition of learned behaviors. These three key features are the foundations of social cognitive learning theory and are based on the human ability to (a) observe, (b) symbolize, and (c) self-regulate (Bandura, 1977b).

According to Bandura, learning takes place in a number of ways but much social learning is accumulated through observations of modeled behavior (Bandura, 1975, Bandura, Grusec, & Menlove, 1966). Three specific human abilities are necessary for social learning to take place: observational learning, symbolizing information, and self-regulation. Each of these components will be discussed in detail. The ability of humans to use symbols for coding of experiences and observed events allows these events to be stored and used for later recall when evaluating possible outcomes (Bandura, 1977b; Bandura et al., 1966). Self-regulatory processes control whether or not a particular behavior is exhibited as individuals can decide when and how to react to their environment (Bandura, 1977b; Bandura et al., 1966). The importance of the ability to self-regulate manifests itself in altering performance through motivation (Bandura, 1978). In this reciprocally interacting system, external factors exert influence on the operation of the self-system (Bandura, 1978). Self-regulatory functions develop based on interactions with the environment; through interactions with others, processing of information, and
development of personal standards by which to evaluate performance (Bandura, 1978).
Each of these key features of social learning is discussed in the following sections to
establish the framework for later research in the area of self-efficacy and academic
achievement.

*Observational Learning*

Early studies established that several factors influence whether or not behavior is
imitated after being modeled (Bandura & Barab, 1971). According to Bandura (1977b),
four component processes dictate whether observational learning takes place. The four
compONENT processes are: Attentional Processes, Retention Processes, Motor
Reproduction Processes, and Motivational Processes. Individuals cannot learn if they do
not possess adequate attentional processing, that is, the ability to attend to, and perceive
accurately, the behavior that they are intended to model (Bandura, 1977b). Many factors
influence the attention and perception of the observer. Among these is the interpersonal
attraction between the model and the observer (Bandura, 1977b). Applying this to the
academic setting, students are the observers and the teacher is the model. Part of the
attentional processing then, applying this concept developed by Bandura (1977b), would
be the relationship between the students and the teacher. If the model (teacher) is
perceived as lacking in some way, either in substance or in personality, then the observer
(student) will choose to ignore the intended behavior to be learned (the lesson) (Bandura,
1977b). Another factor in attentional processing is the incentive to pay attention to the
modeled behavior and symbolize it for later recall (Bandura et al., 1966). However,
Bandura et al. (1966) found that a treatment group that received an anticipatory incentive
set actually performed imitative behavior less accurately than the treatment group without
the incentive set. The researchers concluded that the anticipatory incentive set instructions might have increased achievement anxieties in some of the participants and thus caused a poorer performance (Bandura et al., 1966).

Symbolizing Information

A second key feature of social cognitive learning theory is the human ability to symbolize information. Observation alone does not produce learning; the behavior must be retained for future use (Bandura, 1977b). The ability of humans to use symbolic coding to retain information for later use is very important to the process of learning. Being able to reproduce a modeled behavior that has been observed and retained depends in part upon the physical capabilities of the individual and is referred to as motor reproduction processes.

The individual must also be able to decode the symbolically coded representation stored from the earlier modeling observation into the appropriate behavior (Bandura, 1977b). If the behavior is complex, it is unlikely that the first response will be error-free and the individual will need to make adjustments in order to match the original behavior modeled. Some of the adjustments will be self-correcting as the individual compares his action to the modeled behavior and other adjustments may require feedback from others. Carroll and Bandura (1982) showed that visual feedback enhanced observational learning but only after the modeled behavior was cognitively represented. This is an important concept to recognize when considering high school students, particularly in regular or lower level classes, since retention of information is key to being able to use or apply the information in later situations.
Retention of modeled behavior involves two types of representations, imaginal and verbal (Bandura, 1977b). This allows humans to retain information in a readily accessible format for future recall and subsequent behavior choice. Some modeled behaviors are stored as images but the majority of cognitive processes are verbal. Verbal representations are stored by use of symbolic coding and several historical studies have shown the importance of symbolic storage of information, which is the ability to use representational forms to preserve experiences for future use (Bandura, 1977b; Bandura & Jeffery, 1973; Bandura, Jeffery, & Bachicha, 1974). Bandura and Jeffery (1973) investigated the effects of symbolic coding and rehearsal patterns on retention of observationally learned behavior.

Participants were randomly assigned to a type of coding or noncoding group. Results indicated that observational learning was significantly improved by symbolic coding (Bandura & Jeffery, 1973). There was not a significant difference between the types of coding, numerical or verbal, however, the numerical coders achieved a significantly higher level of learning than the noncoding participants. Verbal coders exhibited a higher level of learning than the noncoding group but not at a significant level. Analysis of variance of the delayed responses was significant, however, and indicated that retained learning was substantially increased by symbolic coding, by immediate rehearsal, and by the interactive effects of both.

This key feature of social cognitive learning theory, the ability to symbolize information, should be considered when designing studies concerned with increasing a specific type of behavior, such as math achievement. Often times, at the high school level, students are not given an opportunity to “rehearse” new material. For mathematics,
this may lead to poorer performance or achievement because the students cannot recall
the process successfully at a later time.

These results demonstrated the importance of symbolic processes on the level of
learning and retention of learning over time. Furthermore, immediate rehearsal of the
coding greatly improved the retention of the learning (Bandura & Jeffery, 1973).
Additional study was conducted to investigate the effects of code meaningfulness on
memory performance (Bandura, Jeffery, & Bachicha, 1974). Analysis of covariance
using the delayed reproduction scores showed a significant result between type of coding
and retention of learned responses (Bandura et al., 1974). Participants who rehearsed the
codes had higher retention rates than those who did not rehearse. These results indicated
that retention of modeled behavior was increased when memory codes were more
meaningful and retrievable. These are important findings to apply to high school students
and how they learn new material; for example, having the students create their own
method of coding might increase their achievement.

Self-Regulation

Bandura (1991) described the self-regulatory mechanism of human behavior as
operating through three subfunctions: self-monitoring of one’s behavior, judgment of
one’s behavior in relation to personal standards and environmental circumstances, and
affective self-reaction. These subfunctions are the mechanisms through which the self
system acts on human behavior. Self-observation is important in the self-regulation
process because it provides necessary information for evaluating one’s progress toward a
set of goals (Bandura, 1991). In order to accomplish goals, individuals need to pay
attention to their performances and behaviors to gather information about successful and
unsuccessful actions. Several factors influence the effect of self-observation on future behavioral changes including the frequency of the self-monitoring, recognition of behavior patterns, accuracy of the self-monitoring, and personal standards of behavior (Bandura, 1991).

Characteristics that Bandura noted as necessary for self-regulation can be thought of as different forms of feedback to the individual. Once the self-observation is noted, it must be judged against some standard. The standards by which individuals gauge their progress are formed from various sources. The first is the set of personal standards formed through direct and vicarious reactions to their past behavior. Individuals also make social comparisons of their behavior to assess the level of their performance. The importance of the activity influences the judgmental process as well, with nonsignificant activities having little or no impact. Individuals also consider the determinant of their behavior, ascribing more importance to behaviors accomplished through their own abilities as opposed to those behaviors heavily influenced by external factors (Bandura, 1991). Finally, self-reactions people experience affect the level of satisfaction derived and thus influence future incentives to repeat the behavior (Bandura, 1991). If the behavior or performance produced a positive self-reaction, then it is likely to create future incentive to repeat the action. If the behavior or performance produced a negative self-reaction, however, then it is likely that the individual will avoid the behavior in the future. This finding provides a basis for the examination of the effects of feedback on behavior. For example, mathematics students who have experienced numerous failing grades (a form of negative feedback) on math performance in the past will find little motivation to attempt future math performances with much effort, expecting more failure.
Motivational processes determine whether a learned behavior is exhibited (Bandura, 1977b). Individuals engage in behaviors they find self-satisfying and do not reproduce those they dislike, even if they are capable of doing so (Bandura, 1977b). This concept is apparent in most American high schools, where many capable students do not work to their potential. Much attention is currently given to the delivery style of lessons and reforming high schools in America, in effect, to engage all students. Intrinsic motivation to perform well on tasks that do not interest students, such as state assessment testing, is generally lacking. According to the literature, being able to perform a task and choosing to perform the task at a subsequent time is dependent upon individual perceptions and past experiences.

Self-Efficacy

As previously stated in the discussion about social cognitive learning theory, whether or not a learned behavior is exhibited depends on the perceived positive or negative outcomes by the individual (Bandura, 1977b). External reinforcement can provide incentive to some individuals under certain circumstances to perform desired behaviors. In an early study on the impact of reinforcement on behavior, Bandura and Barab (1971) found that discrimination processes play an influential role in nonreinforced imitation and concluded that imitative behavior was controlled primarily through anticipated consequences. Individuals initially respond to reinforcement but eventually develop self-corrective adjustments based on informative feedback. Individuals have expectancies that specific behaviors will lead to certain outcomes and efficacy expectations about whether or not they can successfully execute the required behavior (Bandura, 1977a, 1983).
Bandura (1977a) defined self-efficacy as beliefs in one’s capabilities to organize and execute actions required to manage situations that occur in their environment. An individual’s expectations of their ability to perform the desired behavior affect both the initiation and the persistence of coping behavior (Bandura, 1977a, 1978, 1982a, 1982b). People tend to avoid situations they feel exceed their coping skills yet get involved in activities they judge themselves capable of handling (Bandura, 1977a, 1978, 1980, 1982a, 1982b). The illustration in Figure 2 represents Bandura’s theory about the psychological effects of the interaction between self-efficacy perceptions and outcome perceptions (Bandura, 1982b, p. 140). Outcome judgments and self-efficacy judgments are determined to be either negative or positive by the individual. The interaction between negative and positive outcome judgments and self-efficacy judgments produces the different psychological effects noted in Figure 2. Perceived self-inefficacy can be the

<table>
<thead>
<tr>
<th>Outcome Judgment</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social activism</td>
<td>Assured, Opportune action</td>
</tr>
<tr>
<td>Protest</td>
<td></td>
</tr>
<tr>
<td>Grievance</td>
<td></td>
</tr>
<tr>
<td>Milieu Change</td>
<td></td>
</tr>
<tr>
<td>Resignation</td>
<td>Self-devaluation</td>
</tr>
<tr>
<td>Apathy</td>
<td>Despondency</td>
</tr>
</tbody>
</table>

*Figure 2.* Interaction between self-efficacy perceptions and outcome perceptions according to Bandura (1982b, p. 140).
result of two different situations. One situation is a result of people giving up when they doubt that they possess the necessary skills to accomplish a given task. The other situation occurs not as a result of a perceived lack of skills, but rather when people perceive that their efforts will not produce the desired effects. The first situation calls for a development of self-efficacy while the second situation requires a change in the social environment so that the perceived skills can be used. People with high self-efficacy levels will work toward changing the environment through social activism or protest while those with low self-efficacy will tend to give up more readily (Bandura, 1982b). It has been demonstrated that these influences of self-efficacy affect the behavior choices that individuals make (Bandura, 1993). Individuals with low self-efficacy will, for instance, avoid tasks and situations that they believe are beyond their capabilities for success while individuals with high self-efficacy will choose activities that challenge them and assist them in developing skills and competencies in a variety of areas (Bandura, 1993).

Using these perceptions of controllability and outcome expectancies of situations, Bandura and Wood (1989) conducted an experiment in the form of a simulation to observe the effect of self-regulation mechanisms that govern performance attainments. In the simulation, the participants managed an organization and were randomly assigned to one of four treatment conditions: low-controllability of an organization with a high-performance standard condition, low-controllability of an organization with a low-performance standard condition, high-controllability of an organization with a high-performance standard condition, and high-controllability of an organization with a low-
performance standard condition. Management of the organization was simulated through a computer program and included 18 trials.

Effects of the simulation produced significant main effects results for performance standards and trial blocks, and for the interaction between controllability and trial blocks (Bandura & Wood, 1989). Findings indicated that viewing organizations as controllable increased managerial self-efficacy but the strength of perceived self-efficacy across trial blocks varied as a function of controllability and performance standards, producing a marginally significant three-way interaction effect. A three-way interaction effect means that there is a joint effect among the three variables of controllability, management self-efficacy, and performance standards that is stronger than the individual main effects of each variable considered independently. Subjects who had been assigned to the low-performance standard condition exhibited an increasing level of managerial self-efficacy while those assigned to the high-performance standard condition exhibited a decreasing managerial self-efficacy the longer they performed the task. However, subjects who were assigned to the low-controllability exhibited a low sense of self-efficacy regardless of the performance standard condition. The subjects adopted more challenging goals for the organization when they perceived it as controllable. This difference in level of goals became more pronounced as the number of trials increased. The results indicated that perceived self-efficacy, which was affected by prior accomplishments, influenced subsequent organizational performance through its effects on analytic strategies.

The combined set of explanatory variables of prior performance, perceived self-efficacy, personal goals, and analytic strategies in the conceptual model accounted for a
major share of the variance in organizational achievements in the second block of trials \( R^2 = .66, p < .001 \) and the third block of trials \( R^2 = .77, p < .001 \) (Bandura & Wood, 1989). These \( R^2 \) values indicate that between 66% and 77% of the variability in the simulated organizational achievements was explained by the relationship among the explanatory variables of prior performance, perceived self-efficacy, personal goals, and analytic strategies and the response variable of organizational achievement. Past performance and personal goals influenced self-efficacy and ultimately the analytic strategies. Indeed, this is an important finding as this study shows that even when the standards were low or easy to attain, there was a sense of inability to affect change in the organization when low levels of self-efficacy were noted. This finding is common among American high school students, particularly in math classes. Many students who are failing math classes are cognitively capable of learning the material but do not attempt to perform given tasks, perhaps due to a low self-efficacy based in part on a sense of low-controllability.

Individuals with strong self-efficacy also cognitively visualize successful solutions to problems while those with weaker self-efficacy visualize failure and dwell on negative outcomes (Bandura, 1980, 1989). Human motivation is based on cognitive functioning when people exercise forethought as they develop their actions. This is accomplished through setting goals and planning courses of action. As evidenced in the Bandura and Wood (1989) study just discussed, the greater persons’ beliefs in their abilities, the greater and more persistent their efforts (Bandura, 1977a, 1980, 1982a, 1982b, 1989; Bandura & Cervone, 1983). This shows the powerful influence of self-beliefs.
Other studies on the conditions affecting the perception of self-efficacy have been conducted using not only college students but high school students as well (Cervone & Peake, 1986). In their study, two experiments were conducted involving self-efficacy and judgmental processes, one with college students and then replicating the experiment with a small group of high school students. The researchers manipulated the questionnaire about self-efficacy, making the task seem relatively easy or very difficult by the use of an anchoring value. Anchor was operationalized for the experiment as an initial value adjusted upward or downward by the participant, often in the direction of the original initial value. The hypothesis was that the perception of the task would be biased in the direction of the manipulated difficulty anchor.

It was expected that the participants’ judgments of self-efficacy would affect their behavior accordingly (Cervone & Peake, 1986). Results indicated that subjects receiving the high anchor value displayed higher self-efficacy than the low anchor control group and the no anchor group (Cervone & Peake, 1986). Judgments of self-efficacy were significantly affected by the anchoring value. No main effects for task or gender and no interactions between anchoring and task or anchoring and gender were found. In regression analyses, task persistence was regressed on the level of the anchor value and self-efficacy. Self-efficacy was a significant predictor of task persistence, while the anchor value was not a significant predictor. Similar results were found with the high school students (Cervone & Peake, 1986). These studies illustrate the indirect influence of self-efficacy on task performance and help corroborate Bandura’s theory about the self-regulatory function in the self-system and whether or not behavior is exhibited. In
other words, there appears to be some judgment or self-evaluation between one’s capabilities and the expected outcome that affects task initiation and persistence.

Similarly to the study just discussed, Cervone (1989) conducted another experiment designed to study the effects of self-efficacy by investigating the manner in which individuals could hinder or facilitate their performance on an upcoming activity. Subjects were led to focus on the difficult factors of the upcoming task, on the easy factors of the upcoming task, or not focused at all on the upcoming task. Cervone (1989) found similar results to his earlier experiments (Cervone & Peake, 1986) with the type of factors (difficult, easy, none) focused on prior to the experiment significantly affecting self-efficacy judgments. Difficult factors subjects judged they could solve significantly fewer tasks than easy factors subjects. These were the same results from previous studies done by Cervone and Peake (1986). In addition to the self-efficacy judgments, the researcher found a significant difference in task persistence (Cervone, 1989). Difficult factors subjects abandoned the task quicker than easy factor subjects and control factors subjects. This finding supports previously discussed theories about task persistence (Bandura, 1977a, 1980, 1982a, 1982b, 1989; Bandura & Cervone, 1983). As in the earlier experiments (Cervone & Peake, 1986), there was no significant difference between the easy and control factors subjects.

In addition to task persistence, the direct effects of perceived self-efficacy on self-efficacy judgments were also assessed in the same study (Cervone, 1989). Type of factors (difficult, easy, none) significantly affected self-efficacy judgments ($R = .314$) and self-efficacy judgments strongly predicted persistence ($R = .418$). When both perceived self-efficacy and types of factors were used in a regression with task persistence, self-
efficacy judgments significantly predicted persistence ($R = .384$), but types of factors did not ($R = .066$) (Cervone, 1989). Cervone (1989) concluded that focusing on factors that impair or facilitate performance on an upcoming event affects both self-efficacy judgments as well as subsequent behavior. Higher perceived self-efficacy increased task persistence, which holds some implication in the area of performance motivation (Cervone, 1989). While these findings support theories of the influence of self-efficacy, it should be noted that the $R^2$ values for this study range from .10 to .17, leaving a large amount of the variance in self-efficacy judgment and task persistence unexplained. Even though the $R^2$ values are small here, the study gives some support to the influence of self-efficacy judgments on the amount of task persistence. These results strongly support Bandura’s theory of those with high self-efficacy exhibiting more effort and persistence than those with low self-efficacy (Bandura, 1977a, 1978, 1982a, 1982b).

*Self-Efficacy and Academic Achievement*

Cognitive skill is not the sole determinant of academic achievement. Motivational, affective, and selection influences on perceived self-efficacy also contribute to academic achievement (Bandura, 1993). In a meta-analysis of self-efficacy studies, Multon, Brown, and Lent (1991) found evidence for positive and statistically significant relationships between self-efficacy beliefs and academic performance and persistence outcomes. Multon et al. (1991) used a sample of 38 studies from 1977 to 1988 that contained measures of self-efficacy, academic performance or persistence, and enough information to calculate effect size estimates. Results of the meta-analysis effect sizes for the 4,998 subjects were .38 for academic performance and .34 for persistence (Multon et al., 1991). According to Cohen (1977) as cited in Stevens (1996, p. 177),
effect sizes greater than .14 are considered large. These effect sizes showed a positive relationship between self-efficacy and academic performance and persistence. The meta-analysis effect size for the persistence measurement was explained in large part by the way persistence was operationalized in the study. Persistence was defined differently among the studies. Studies that used the number of tasks completed as the measure of persistence showed stronger effect sizes than studies that measured persistence as the time on task (Multon et al., 1991). With regard to this discrepancy, it is important to note that Bandura (1977a, 1978, 1980, 1982b) described those with low self-efficacy as giving up more readily than those with high self-efficacy, which implies task persistence to mean time on task rather than the number of tasks completed.

Further analysis of academic performance in Multon et al.’s (1991) study revealed significant heterogeneity among the effect size estimates. This indicated a large variability in the performance variable among the types of students, measures and study designs (Multon et al., 1991). The researchers indicated four conditions that appeared to influence the effect sizes: the time period during which the self-efficacy and performance were assessed, the students’ achievement status, the type of performance measure, and the student age. For example, the timing of the performance assessment was different in a number of studies. Studies of low-achieving students often used posttreatment scores while the majority of the studies of normal-achieving students used pretreatment or correlational data. Stronger effect sizes were noted when the measurements were from posttreatment rather than pretreatment or correlational data (Multon et al., 1991). Because the studies were not similar in the timing of the performance assessment, the effect sizes should be interpreted with caution.
In addition to the timing condition, stronger relations between self-efficacy and performance were found among low-achieving students rather than normal-achieving students with unbiased effect size estimates of .56 and .33, respectively. The researchers suggested one explanation for this difference between the two groups could be attributed to the type of data collection in the particular study. The strongest effect sizes were found in the studies that used basic skills measures. The next strongest effect sizes were in studies of classroom-based performances, with standardized achievement tests having the weakest effect size. Finally, a fourth condition suggested to explain the heterogeneity among the studies in the academic performance variable was the age of the participants. Performance results varied, with high school and college student samples showing stronger effect sizes than elementary samples (Multon, Brown, & Lent, 1991). The researchers suggested that the larger effect size might be due to older students possessing more accurate self-efficacy perceptions as well as more school experience than the younger elementary students.

Other research in the area of self-efficacy and academic achievement focused on developing models that explained how self-efficacy beliefs affect academic achievement (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Bong, 2004; Zimmerman et al., 1992). Zimmerman et al. (1992) used path analysis to develop a causal model of students’ self-efficacy beliefs and academic goals in self-motivated academic attainment. Students completed questionnaires at the beginning of a social studies course that assessed their efficacies for self-regulated learning and academic achievement, and their grade goal for the course. Parents were also asked for a grade goal for their children for the course. Parents’ grade goals were significantly higher than their children’s grade
goals. Prior grades in the previous social studies course were obtained as well as the final grade for the current social studies course. A path analysis using students’ self-efficacy scores for self-regulated learning and academic achievement, students’ grade goal for the course, parents’ grade goal for the course, and prior social studies grades as the explanatory variables indicated that about 31% of the variability in the final grade was explained by the self-efficacy for academic achievement and student goals \(R = .56, p < .01, R^2 = .31\) (Zimmerman et al., 1992). The path coefficient between self-efficacy for academic achievement and final grade in the course was \(\beta = .21, p < .05\). The path coefficient between student goals and final grade in the course was \(\beta = .43, p < .05\). Self-efficacy for academic achievement influenced personal goals, exhibited by the path coefficient of \(\beta = .36, p < .05\) and the researchers concluded that perceived efficacy to achieve motivates academic attainment directly as well as indirectly by influencing personal goals. There was a combined direct and indirect effect of students’ self-efficacy for academic achievement through student goals of \(\beta = .37, p < .05\). Much of the variance, however, was not explained and it was suggested that there were likely other factors of influence not yet identified (Zimmerman et al., 1992).

Another path analytic study also examined the link between self-efficacy and academic achievement, but with additional factors (Bandura et al., 1996). As in previous self-efficacy and academic achievement studies, students completed several sets of scales including measurements of academic efficacy, self-regulatory efficacy, social efficacy, and academic aspirations. Similar to the study by Zimmerman et al., (1992), parents rated their children on academic efficacy, academic aspirations, and prosocial behavior but this study also included additional factors such as the socioeconomic status of the
family, based on the father’s occupation, and teachers’ ratings of the students on prosocial behavior, problem behavior, depression, and academic achievement (Bandura et al., 1996). The dependent variable was academic achievement and results showed the entire set of factors accounted for 58% of the variance in academic achievement (Bandura et al., 1996), which is larger than the path model found by Zimmerman et al. (1992). Children’s academic self-efficacy and children’s self-regulatory self-efficacy both had direct influence on academic achievement with each path coefficient equal to \(\beta = .11\). 

**Measurement of Self-Efficacy**

Specificity of the measurement of self-efficacy has been recurrent throughout the literature (Bong, 2002; Bong & Hocevar, 2002; Kranzler & Pajares, 1997; Pajares, 1996a; Pajares & Miller, 1994, 1995; Smith & Fouad, 1999). It has been shown that measuring self-efficacy prior to performance has no effect on level of performance or fear arousal (Bandura, 1982a), but Pajares and Miller (1995) conducted a study to determine if the format of the subsequent assessment would influence students’ self-efficacy judgments. Pajares and Miller (1995) used the three subscales of the Mathematics Self-Efficacy Scales (MSES) and a mathematics problem solving performance in a study of middle school gifted students. As was expected, the results indicated that the three math self-efficacy subscales were strongly related to each other, but the problem solving subscale had a stronger relationship with the problem solving performance. The researchers stated that this confirmed previous remarks in the literature that self-efficacy needs to be measured at the corresponding level of specificity of the subsequent task (Pajares & Miller, 1995).
More specifically, Pajares and Miller (1995) examined the format of measuring math self-efficacy. Middle school students were randomly assigned within their math class to one of four groups, two with multiple-choice math problem assessments and two with open-ended math problem assessments. Results indicated no differences in self-efficacy based on the form of the questions presented in the self-efficacy instrument. However, multiple-choice assessment scores were higher than open-ended assessment scores. The researchers suggested a possible explanation was the familiarity of the students with the more traditional assessment format of multiple choice. Multiple regression analyses, controlling for the influence of class level (algebra, prealgebra), gender, assessment format, self-efficacy format, and the interaction between self-efficacy and assessment format, produced a significant model with $R^2 = .57$ (Pajares & Miller, 1995). In addition to the determination of the influence of the assessment type on self-efficacy, the study also included statistical analysis of the calibration between the self-efficacy judgments and the performance on the assessments (Pajares & Miller, 1995). Calibration scores for the open-ended assessment groups were lower than for the multiple-choice assessment groups. The researchers concluded that understanding how self-efficacy and the performance tasks are measured is important for valid results (Pajares & Miller, 1995).

It has been noted numerous times in the literature that measurements of self-efficacy must be specific to the performance task of interest rather than a global assessment and that this is not often followed (Pajares, 1996a; Pajares & Miller, 1994, 1995). For example, Pajares and Graham (1999) found that the only predicting variable in their study of mathematics performance of middle school students was students’ task-
specific self-efficacy. Task-specific math self-efficacy measured in the study accurately predicted mathematics performance at both the beginning and end of the school year. The researchers attributed this finding to the specificity of the self-efficacy instrument used. The problems on the math self-efficacy instrument were based on the subsequent performance task (Pajares & Graham, 1999).

More recent studies examined domain-specific self-efficacy at varying levels of specificity. Bong (2002) asked a sample of Korean female high school students about their subject-, task-, and problem-specific self-efficacy perceptions in English and mathematics and compared them to the students’ midterm and final exam scores in both English and mathematics, used to measure delayed performance. Multivariate analyses of variance were conducted and showed statistically significant differences in both self-efficacy for English and mathematics. Multivariate analyses of variance for differences in achievement were also statistically significant for both English and mathematics. Confirmatory factor analysis and structural equation modeling were used to analyze the data and supported the hypothesis that self-efficacy beliefs measured at different levels of specificity form correlated but separate factors. However, the results did not support the hypothesis that within each academic domain problem-specific self-efficacy would produce stronger relationships with performance scores than task-and subject-specific self-efficacies. It was noted that each achievement domain was predicted by its specific self-efficacy, that is, English self-efficacy predicted only English achievement and mathematics self-efficacy only mathematics achievement (Bong, 2002). Bong (2002) concluded that more specific measures of self-efficacy were somewhat more accurate than more general measures of self-efficacy in predicting outcomes.
The generality of academic self-efficacy and its relation to domain-specific academic self-efficacy is an important issue in self-efficacy research. Bong (1997, 1999, 2001, 2004) is one of the self-efficacy researchers who addressed this issue in several empirical studies. Bong (1997) used selected problems from the Scholastic Aptitude Test preparatory booklets representing English, Spanish, U.S. history, algebra, geometry, and chemistry to form self-efficacy scales measuring each subject. Self-efficacies were collected on each of the subjects from a large sample of American high school students. Correlations among the six subjects ranged from .11 to .92. Using the six high school subjects as first-order factors and a second-order general academic self-efficacy factor, more than 50% of the variance in five of the six subjects was explained (Bong, 1997). However, less than 5% of the variance was explained in Spanish, which indicated that a general academic self-efficacy with one factor was not an adequate representation of the data. Subsequent statistical analyses indicated that the six first-order factors of the high school subjects was best represented by two second-order factors, verbal and quantitative, with chemistry loading on both second-order factors. Findings illustrated that higher order factors were inadequate estimations of subject-specific academic self-efficacy. These findings are important as they provide justification for the development and use of specific academic self-efficacy measures to match particular performances of interest (Bong, 1997). Additional research showed similar results with moderate correlations among subject self-efficacies (Bong, 2001, 2004).

Predictive capability of domain-specific self-efficacy consequently has been investigated for its ability to predict corresponding academic performances at various
grade levels in verbal and quantitative domains by numerous researchers (Campbell & Hackett, 1986; Hackett & Betz, 1989; Lent, Brown, & Gore, 1997; Lent, Lopez, & Bieschke, 1993; Pajares, 1996a; Pajares & Johnson, 1996; Pajares & Miller, 1994, 1995; Pajares & Valiante, 1997; Zimmerman & Bandura, 1994). For example, Zimmerman and Bandura (1994) investigated the predictiveness of academic achievement in a college writing course based on self-efficacy and self-regulatory influences. Students were measured at the beginning of the course on different self-regulatory factors by responding to four questionnaires: self-regulatory efficacy for writing, self-efficacy for academic achievement, grade goals, and self-evaluative standards (Zimmerman & Bandura, 1994). Students were asked to identify their grade goal using a 12-point scale ranging from F to A and to rate how satisfied they would be for each of the 12 grade levels using a 7-point scale with 1 as very dissatisfied and 7 as very satisfied. SAT verbal aptitudes were obtained and final grades for the course at the end of the quarter were used as the measure of writing attainment.

The researchers developed a causal model and conducted a path analysis using verbal aptitude, self-regulatory efficacy for writing, self-efficacy for academic achievement, grade goals, self-evaluative standards, class membership (regular, advanced), and final course grade. Self-regulatory factors mediated the influence of verbal aptitude and added substantially to the writing attainment final course grades. Students’ perceived self-regulatory efficacy for writing predicted both self-efficacy for academic achievement and self-evaluative standards. Path analysis revealed that perceived self-efficacy also directly affected the writing attainment of the participants. Perceived self-efficacy and personal goal setting accounted for 35% of the variance in
writing grades. Self-efficacy for academic achievement affected writing attainment final course grades directly and indirectly through goal setting (Zimmerman & Bandura, 1994).

Based on these results, Zimmerman and Bandura (1994) suggested that instructors could assist students in areas of weakness identified through self-regulatory efficacy scores by providing strategies for improvement or modeling successful behaviors, a significant finding for future studies as it suggests using efficacy scores to impact behavior. These findings are important because the two studies used students with different educational backgrounds, different academic subjects, and different ethnic compositions. The measures of self-efficacy and goal setting were tailored to the specific subject matters and populations (Zimmerman & Bandura, 1994).

Another line of self-efficacy research employed path analysis to determine the direct and indirect influences of writing self-efficacy, writing apprehension, and writing aptitude on essay-writing performance (Pajares & Johnson, 1996; Pajares & Valiante, 1997). Pajares and Johnson (1996) produced a model in their study that accounted for 53% of the variance in essay writing performance while a model for the data in the study by Pajares and Valiante (1997) accounted for 64% of the variability in essay writing performance. This is a considerable amount of variance explained in writing performance. The model produced in the study by Pajares and Johnson (1996) demonstrated direct influences between writing aptitude and writing self-efficacy on writing performance with path coefficients of .378 and .395, respectively. Writing self-efficacy also showed a direct influence on writing apprehension with a path coefficient of -.517 (Pajares & Johnson, 1996). This result is not surprising because, according to
social learning theory, as self-efficacy increases, apprehension decreases. Also not surprising was the strength of the direct influence of writing aptitude on writing self-efficacy with a path coefficient of .464 (Pajares & Johnson, 1996). Self-efficacy is based in part on performance accomplishments so it is to be expected that as one’s aptitude increases, so does one’s self-efficacy.

Similar path coefficient values were noted in the model from a study by Pajares and Valiante (1997). Writing aptitude showed a stronger direct influence on writing performance with a path coefficient of .601 while the direct influence of writing self-efficacy on writing performance was about the same, with a path coefficient of .356 (Pajares & Valiante, 1997). The direct influence of writing aptitude on writing self-efficacy was a little weaker with a path coefficient of .311, but the direct influence of writing self-efficacy on writing apprehension was similar with a path coefficient of -.452 (Pajares & Valiante, 1997). Neither study found a difference in performance between boys and girls (Pajares & Johnson, 1996; Pajares & Valiante, 1997).

**Math Self-Efficacy and Math Achievement**

As a more domain specific efficacy construct, math self-efficacy also has been examined for its direct and indirect influences on other variables such as math anxiety, math confidence, and task persistence (Campbell & Hackett, 1986; Hackett & Betz, 1989; Lent, Lopez, & Bieschke, 1993; Pajares, 1996a; Pajares & Kranzler, 1995; Pajares & Miller, 1994, 1995; Siegel, Galassi, & Ware, 1985; Trice, Elliott, Pope, & Tryall, 1991). For example, a more recent study conducted by Trice et al. (1991) examined the effects of math self-efficacy on task persistence. The effect of math self-efficacy on task persistence was shown to indirectly influence mathematics achievement, in keeping with
social learning theory (Trice et al., 1991). Trice et al. (1991) found that college students with low math self-efficacy responded with low performance on a subsequent task after experiencing high failure. The findings also showed that college students with high math self-efficacy did not perform any differently or actually increased their effort at a subsequent task after experiencing failure.

Researchers have also used math self-efficacy to predict varying levels of mathematics achievement, from specific math problem-solving ability within a specific strand of mathematics such as algebra, to a more general math achievement. For example, math self-efficacy has been shown to be a strong independent predictor of problem solving ability (Pajares, 1996a; Pajares & Kranzler, 1995; Siegel et al., 1985). Siegel et al. (1985) compared two theoretical models for the explanation of the mathematics final exam performance of a group of college students. Results of the comparison indicated that the social learning theory model proposed by Bandura accounted for significantly more of the variance in performance than the math aptitude-anxiety model (Siegel et al., 1985). Multiple regression findings showed that the social learning model accounted for a large proportion of the variance in math performance ($R^2 = .547, F(5, 137) = 33.16, p < .0001$) while the math aptitude-anxiety model accounted for a smaller proportion of the variance in math performance ($R^2 = .164, F(5, 5.39) = 33.16, p < .001$) (Siegel et al., 1985).

More evidence of the influence of math self-efficacy was provided by Pajares and Kranzler (1995). The researchers (Pajares & Kranzler, 1995) found a path analytic model using math self-efficacy, general mental ability, math anxiety, gender, and math level that accounted for 60% of the variance in math problem-solving performance of high school
students. The strongest direct effects on math problem-solving performance were from self-efficacy and general mental ability, with general mental ability having a strong direct effect on self-efficacy. This mediated the indirect effect of general mental ability and math level on performance, with path coefficients indicating that self-efficacy ($\beta = .349$) was slightly stronger than general mental ability ($\beta = .324$). Also noted in the study were the effects of self-efficacy on anxiety ($\beta = .394$) and the effects of general mental ability on self-efficacy ($\beta = .397$). These findings illustrate that anxiety is related to self-efficacy as a reaction to self-evaluation and expectancy outcomes. Math self-efficacy was also found to be a strong independent predictor of problem solving ability among gifted students when the effects of math anxiety, cognitive ability, prior mathematics achievement, self-efficacy for self-regulated learning, and sex were controlled for (Pajares, 1996a). Gifted students exhibited higher math self-efficacy and self-efficacy for self-regulated learning and lower math anxiety than regular students (Pajares, 1996a). The results also indicated that both groups of students were overconfident (Pajares, 1996a), a result reported in other studies in the literature (Hackett & Betz, 1989).

Not only was math self-efficacy found to influence specific math problem solving performances, it has been a significant predictor of more general measures of math achievement in other investigations (Hackett & Betz, 1989; Lent, Lopez, & Bieschke, 1993; Pajares & Miller, 1994, 1995). Hackett and Betz (1989) used college undergraduates to examine the correspondence between mathematics self-efficacy and mathematics performance. Mean scores on the Dowling (1978) Mathematics Confidence Scale were used to measure mathematics performance. The results of the study showed an overall correlation coefficient between self-efficacy and performance of .44, which the
authors indicated as a moderately strong positive relationship. The subscale correlation coefficients ranged from .36 to .49 and were all found to be significant ($p < .001$). While gender differences were not significant, the men in the study tended to be overconfident when comparing the self-efficacy and performance scores and the women tended to be congruent or underconfident when comparing self-efficacy to performance. Overall, the judgments about performance were inaccurate, prompting the authors to suggest that perhaps mathematics performance is more complex than other areas of achievement (Hackett & Betz, 1989).

By contrast, however, Lent et al. (1993) found that math self-efficacy contributed significantly to the prediction of college math course grades with $R^2$ change = .10. The authors noted that math self-efficacy, in addition to having a direct influence on grades, also exhibited an indirect influence on grades through interests (Lent et al., 1993). Other factors influencing grades in the study were past performance and interest in the subject matter. It was suggested that math self-efficacy mediated the effects of past performances on interest and that interest then mediates the effects of math self-efficacy on math course selection (Lent et al., 1993). These same results were found by Lopez, Lent, Brown, and Gore (1997) in their study of a large group of high school students in advanced algebra and geometry. The path analyses used in the study indicated a model in which ability helped determine self-efficacy, which, in turn, affected outcome expectations and subject matter interest (Lopez et al., 1997). Past performance produced the largest path coefficient in the model to self-efficacy with a value of .34 for the advanced algebra sample and .55 for the geometry sample, and the researchers suggested that the evaluation of objective ability could affect self-efficacy judgments. The stronger
path coefficient for the geometry sample is important to note because it illustrates the strength of past performance accomplishments on self-efficacy in a sample of students who are not achieving at an advanced level.

Additional research along these same lines included several independent variables. Math self-efficacy, math self-concept, perceived usefulness of mathematics, prior experience with mathematics, and gender were tested in a path analysis for predictiveness of mathematics problem solving (Pajares & Miller, 1994). Math self-efficacy was more predictive of mathematics problem solving than the other variables measured. The effects of gender and prior experience on math self-concept, usefulness of mathematics, and problem solving performance were mediated by mathematics self-efficacy. The path analysis showed that math self-efficacy had a stronger direct effect on problem solving, and math self-concept and prior experience had modest direct effects (Pajares & Miller, 1994).

Sources of Self-Efficacy

There are four sources of self-efficacy expectations: performance accomplishents, vicarious experience, verbal persuasion, and physiological states (Bandura, 1977a, 1982b; Bandura & Adams, 1977; Bandura, Adams, & Beyer, 1977). These four different sources of personal efficacy expectations were identified by Bandura (1977a, 1982b) as having different levels of influence on self-efficacy in that the more dependable the source of personal efficacy, the greater the changes in self-efficacy. Bandura (1977a, 1978, 1980, 1982b) further theorized that expectations of personal efficacy determine whether coping behavior will be initiated, how much effort will be expended, and how long the effort will be sustained in the face of difficulty. Each of
these will be discussed briefly and pertinent empirical studies that support each source of self-efficacy will be provided, however, some of the studies examine one or more sources of self-efficacy and therefore, were placed in the most appropriate discussion based on the findings of the study.

Performance Accomplishments

Performance accomplishments are based on personal mastery experiences. A review of the literature indicates that performance accomplishments are also referred to as mastery learning, enactive performance, and enactive mastery. Increased self-efficacy can lead to generalization in other situations and studies in this area show that successes raise mastery expectations and repeated failures lower them, however, people are more influenced by their perceptions of successful performances rather than the successes themselves (Bandura, 1982b). Bandura, among others, conducted early historical studies and several are reported here.

Bandura, Jeffery and Wright (1974) conducted a study on participant modeling as a function of the amount of response induction aids. They found significant behavioral changes when a wide array of performance aids was available. One of the supportive aids was using subtasks that gradually increased in difficulty. This strategy can be easily applied to a classroom setting where difficult concepts could possibly be presented as a series of smaller subtasks in order to increase successes in the students. Bandura et al. (1974) concluded that widespread changes were best achieved by supplementing the aided behavior with independent self-directed practice to increase the sense of efficacy, that is, rehearsal influences learning of new material, a notion suggested by Bandura and
Jeffery (1973). Other studies support these findings as well (Bandura, 1983; Bandura, Adams, Hardy, & Howells, 1980).

One classic study by Bandura and Adams (1977) indicated that enhancement of self-efficacy could affect behavioral changes. The tests were conducted on severe snake phobics who received standard desensitization treatment until the anxiety reactions were completely extinguished (Bandura & Adams, 1977). Desensitization was operationalized as the process whereby aversive stimuli are presented at graduated levels so that anxiety reactions are minimized, similar to the use of subtasks suggested by Bandura et al. (1974). The process is continued until the anxiety reactions to the aversive stimuli are eliminated. The first study investigated changes in avoidance behavior through enhanced personal efficacy expectations and the second study investigated the process of efficacy and behavioral change using participant modeling as the treatment. Self-efficacy was shown to be an accurate predictor of behavioral change following complete desensitization (Bandura & Adams, 1977).

Later studies indicated that performance accomplishment is determined by the quality of analytic thinking and self-beliefs of efficacy can enhance or impair performance on the cognitive functioning of an individual (Bandura, 1989). Self beliefs of efficacy can affect analytic thinking, anticipatory cognitive simulations, cognitive motivation, intrusive affective arousal, assessment of self-percepts of efficacy, and active producers of performance accomplishments (Bandura, 1989). People must make inferences about conditions and situations by drawing on their knowledge and ability to make judgments about possible results of their actions. People with positive judgments about past performance tend to believe they have the skills to accomplish a task while
those with negative experiences tend to avoid the task, believing they lack the necessary skills to be successful. Performance accomplishment for high school students is based on previous experiences and is a large determinant in whether or not they begin a task, the task persistence exhibited, and how much effort they spend on solving the task. Additional early studies focused on comparisons of effective sources of self-efficacy and these are discussed in the remaining sections on sources of self-efficacy.

Bandura’s four hypothesized sources of self-efficacy (performance, vicarious, verbal, emotional) have been examined for influences on mathematics self-efficacy (Lent et al., 1991; Lopez & Lent, 1992). Lent et al. (1991) found support for the relations between the four hypothesized sources of efficacy and mathematics self-efficacy. Performance accomplishments exhibited the strongest influence of the four sources on mathematics self-efficacy (Lent et al., 1991). Lopez and Lent (1992) used the Sources of Math Efficacy Scale (SMES) to measure the sources of math self-efficacy, a scale consisting of 40 items corresponding to the four sources of efficacy (performance, vicarious, verbal, emotional) identified by Bandura (1977a). They obtained similar results to earlier research (Lent et al., 1991), with performance accomplishments exhibiting the strongest influence on math self-efficacy (Lopez & Lent, 1992).

Using different latent structure models to fit the data, Lent, Lopez, Brown, and Gore (1996) provided additional evidence for the four sources of math self-efficacy. The research was conducted on two groups of participants, one at the college level and the other at the high school level. Statistical analyses showed that a four-factor model fit both groups relatively well, but a five-factor model fit the high school group better. Three factors common to both groups included personal performance, social persuasion,
and emotional arousal. Vicarious learning was the fourth factor for the college group but was replaced in the high school group by adult modeling and peer modeling. The researchers suggested that the interdependence among the factors might be explained by several efficacy sources occurring at the same time. For instance, doing well on an exam (personal performance) might encourage praise (verbal) and lessen future anxiety about testing (emotional arousal) (Lent et al., 1996).

Sources of self-efficacy information were shown to make unique contributions to math self-efficacy in a study by Matsui, Matsui, and Ohnishi (1990). Questionnaires were completed by college freshmen about four sources of self-efficacy (performance accomplishments, vicarious, verbal persuasion, emotional arousal) described by Bandura with regard to their high school math experiences. The researchers determined the content validity of the efficacy information by using principal factor analysis. Similar to other results, Matsui et al. (1990) found that the math self-efficacy was slightly higher for men than women but found no significant differences among the four sources of efficacy between the men and women. Results also indicated that the four sources of efficacy were significantly related to math self-efficacy for the men and all but vicarious were significantly related to math self-efficacy for the women. The authors noted that the participants used for the study had been accepted at a highly selective university and therefore may have exhibited higher scores, possible reducing some effects had the sample been from a different population (Matsui et al., 1990).

Different results for advanced versus regular high school students were obtained in a study of the sources of math self-efficacy and math ability on math course grades (Lopez, Lent, Brown, & Gore, 1997). In the path analysis, there was a significant path
coefficient of .22 ($p < .05$) from social persuasion to math self-efficacy for the advanced algebra group but the path coefficient was not significant for the geometry group ($\beta = -.07$, $p < .05$). The geometry class was a required course but the advanced algebra course was an elective. This is an important finding as it indicates that the advanced algebra course is affected by social persuasion, perhaps in the form of competing for grades and norm-referencing one’s ability with others. It also indicates that direct experience is more important than social persuasion, such as verbal praise or peer comparison, for the students not in the advanced algebra course (Lopez et al., 1997).

**Vicarious Experience**

Learning by observing the actions of others is called vicarious experience. Observing a variety of models can help individuals increase their self-efficacy if the models do not experience negative consequences. However, there is more benefit to individuals’ self-efficacy from observing models overcoming difficulties through perseverance rather than achieving a relatively simple task (Bandura, 1977a; Bandura & Adams, 1977; Bandura et al., 1977). Observing successful models with a wide variety of characteristics can also increase an individual’s self-efficacy. Vicarious experience does rely on an individual’s inferences of modeled behavior and therefore is a less dependable source of efficacy expectations than personal accomplishments (Bandura, 1977a; Bandura & Adams, 1977; Bandura et al., 1977).

Several early experiments on the effects of different experiences and subsequent behavior were conducted using snake phobics who responded to newspaper advertisements (Bandura, 1977a; Bandura & Adams, 1977; Bandura et al., 1977). Bandura et al. (1977) conducted an experiment to test the theory that psychological
procedures affect changes in behavior by altering the level and strength of self-efficacy. The participants either experienced handling the snakes (performance accomplishments) or watched others handle the snakes (vicarious experience). Results showed that the subjects who handled the snakes demonstrated higher, more generalized, and stronger efficacy expectations than did those who observed others handling the snakes, which, in turn, exceeded the control group’s expectations (Bandura et al., 1977). Past performance of the participants in the experiments proved to be a less accurate predictor of behavioral change than self-efficacy judgments. The researchers concluded that people process and integrate diverse sources of information about their self-efficacy and regulate their behavior and effort accordingly (Bandura, et al., 1977). People develop different efficacy expectations even though they participate in similar experiences, in part, because of differences in cognitive processing of efficacy information and also because of exposure to a variety of experiences that have changed individual levels of efficacy (Bandura, 1977a).

Additional studies conducted on the effects of peer models on self-efficacy showed that the type of modeled behavior affected achievement. (Schunk, Hanson, & Cox, 1987). Children in their study who observed peer models gradually learning to work fraction problems attained higher self-efficacy than children who observed a mastery peer model. Additional study comparing self-modeling with peer modeling showed an increase in self-efficacy (Schunk & Hanson, 1989). The participants were videotaped solving fraction problems and then allowed to view their tapes. Using the videotapes as feedback effectively conveyed a sense of progress development in children who previously doubted their progress. Schunk and Hanson (1989) encouraged the use
of videotape for students who had previously experienced difficulty. This study indicates
that vicarious experience can enhance the self-efficacy for individual students but the use
of self-videotapes appears to be evidence of performance accomplishment rather than
vicarious experience. This strategy could be used to convince students with extremely
weak self-efficacy of their performance accomplishment.

Verbal Persuasion

Verbal persuasion is widely used to influence human behavior because it is
readily available (Bandura, 1977a; Bandura & Adams, 1977; Bandura et al., 1977). This
produces a weaker efficacy expectation since there is not a corresponding experience for
the observer. As might be expected based on the performance accomplishment research
already discussed, negative experiences will quickly dispel any verbal suggestions that
the individual can perform the behavior (Bandura, 1977a). Even though previous
research stated that verbal persuasion was less influential on self-efficacy than either
performance accomplishments or vicarious learning, if students are not motivated to
learn, they will not experience either source of self-efficacy. Verbal feedback and its
influence on self-efficacy have been studied extensively (Schunk, 1981, 1982, 1983a,
1984b; Schunk & Cox, 1986; Schunk & Gunn, 1986). In an early study, Schunk (1981)
looked at the type of modeled behavior and effort feedback. The results indicated that the
students who received the cognitive modeling as opposed to the didactic increased both
developmental skills and self-efficacy while the students who received no instruction did
not show any significant changes in skills or self-efficacy and became less persistent at
working the problems, which supports earlier research by others (Bandura, 1977a, 1978,
1980, 1982). An important finding was that effort attribution failed to show any effect on
self-efficacy and the author suggested this might be due to an already positive environment so that the attribution was not effective. Schunk (1981) cautioned the use of effort attribution in that an increase in effort will not override any lack of basic skills necessary to be successful. Later studies in other areas of self-efficacy and academic achievement actually excluded any participants who did not have basic skills. In the current educational arena, however, these are precisely the students (those who lack basic skills) who must be addressed in order to meet state and national goals.

**Physiological States**

Not as strongly as the other three sources, but physiological states seem to influence self-efficacy. Several experiments were conducted to test the theory that psychological procedures affect changes in behavior by altering the level and strength of self-efficacy. Perceived self-efficacy affects emotional arousal through physiological changes that individuals experience during times of anxiety and stress (Bandura, 1977a; Bandura & Adams, 1977; Bandura et al., 1977). High levels of anxiety and stress usually have debilitating effects on performance and individuals expect less success in such situations (Bandura, 1977a; Bandura & Adams, 1977). Affective influences of self-efficacy can cause stress and anxiety arousal when the task is perceived to be threatening or beyond one’s capabilities. Such reactions often cause the individual to experience physical symptoms such as increased heart rate and increased blood pressure (Bandura, 1993). The ability to control disturbing thoughts is part of perceived coping efficacy and individuals who lack this ability can experience stress and depression (Bandura, 1993). Research conducted on mood indicates that there is an effect on self-efficacy through affective influences (Forgas, Bower, & Moylan, 1990; Kavanaugh & Bower, 1985;
Physiological states include emotional arousal and some research has been conducted to study its effects on self-efficacy. Lopez and Lent (1992) found that while performance accomplishments exhibited the strongest influence on math self-efficacy, emotional arousal added significantly to the prediction of math self-efficacy. However, it was suggested that emotional arousal might have functioned as a suppressor variable in the regression due to the incongruence between the negative zero-order correlation with the math self-efficacy and the positive beta weight in the regression equation (Lopez & Lent, 1992). A suppressor variable is a variable that enhances the effects of other variables in the set of independent variables and in so doing, increases the multiple $R^2$ of the relationship (Tabachnick & Fidell, 2001).

**Self-Efficacy and Evaluative Feedback**

Feedback is the information about a performance or behavior that an individual uses to make decisions about future performance or behavior in task initiative and persistence. As previously noted in the self-regulation of social learning theory discussion, outcome judgment affects the perceptions of self-efficacy and several studies have focused on the effects of feedback on self-efficacy and the development of skills or behaviors of the participants (Bandura & Barab, 1971; Bandura & Perloff, 1967; Cervone & Palmer, 1990; Schunk, 1981, 1982, 1983a, 1984a; Schunk & Cox, 1986; Schunk & Gunn, 1986; Schunk & Rice, 1987). Additional studies investigated the effects of the
sequence and timing of using different types of feedback on self-efficacy of children (Schunk, 1984b; Schunk & Cox, 1986).

In a very early study, Bandura and Perloff (1967) tested two types of feedback on behavioral control, a self-monitored reinforcement system and an externally imposed reinforcement system. The self-monitored reinforcement system is composed of four elements: a self-prescribed standard of behavior, a social comparison process, reinforcers under the person’s own control, and the person serves as his own reinforcing agent (Bandura & Perloff, 1967). The majority of the children chose the most number of responses for a single reward. Some children altered their performance standard to a higher level without an increase in reward but a few lowered their standard and then did not reward themselves for the lower achievement. Results of their study indicated that there was no significant difference between the self-monitored and external reinforcement systems; however, boys appeared to be more responsive to the external reinforcement (Bandura & Perloff, 1967). Girls exhibited equal efficacy under both systems. Both systems showed higher sustained behavior than the groups that received noncontingent rewards or no rewards at all (Bandura & Perloff, 1967). This was an important experiment because it pointed subsequent research in a direction of establishing the conditions for self-reinforcement and investigation of subsequent alterations to self-set standards.

A subsequent study by Bandura and Barab (1971) found that discrimination processes have an influential role in nonreinforced behavior. The researchers also found that external reinforcement influenced initial imitative behavior but individuals later developed self-corrective adjustments based on informative feedback. Self-corrective
adjustments are important in education, especially at the high school level. High school students typically experience less verbal reinforcers than elementary and middle school students and must develop self-corrective skills in order to achieve at a higher level.

Quite similar to these studies, Schunk (1983b) investigated the effects of self-monitoring on self-efficacy and skill with similar results. Children in the study were randomly assigned within sex and school to one of three treatment groups: self-monitoring, external monitoring, or no monitoring. Results indicated that monitoring was highly effective in promoting self-efficacy and achievement, however, the method of the monitoring was not important. Students in the no monitoring group were left to assess their own progress and even if skills were developed, they may have been unaware of their capabilities (Schunk, 1983b).

Subsequent research focused on the effects of different types of feedback in educational settings. Schunk (1982) tested the effects of effort attributional feedback given in the context of competency on children’s perceived self-efficacy and achievement. Effort attributional feedback was operationalized as feedback (verbal in this study) that attributed achievement to the effort of the participant. Four conditions were tested: past attribution, future attribution, no feedback with monitoring, and no feedback without monitoring. Past attribution is feedback that relates information based on performances or tasks that have already taken place. Future attribution is feedback that relates information to performances or tasks that are yet to happen. For example, students in the past attribution group received feedback such as “you’ve been working hard” while students in the future attribution group received feedback such as “you need to work hard.” The remaining groups received no feedback. Results showed that using
the past attribution feedback promoted task involvement, skill development, and increased self-efficacy. However, the future attribution feedback did not show any increase in achievement behavior beyond that of the training group that received no feedback or monitoring. Schunk (1982) stated that this finding supported Bandura’s theory of performance accomplishments as information for self-assessment. Furthermore, effort attribution may actually decrease self-efficacy because the feedback is not related to competence or skill, as in accomplishing a given task successfully (Schunk, 1982).

Building on this study, Schunk (1983a) conducted additional research to examine the differentiation between two types of feedback. The study compared the effects of effort attributional feedback, operationalized as in the previous study (Schunk, 1982), to the effects of ability attributional feedback. Ability attributional feedback was operationalized as providing feedback related to the skills of the participant. Students were randomly assigned within sex to one of four treatment groups: ability attributional feedback given, effort attributional feedback given, both ability and effort attributional feedback given, or no attributional feedback given. The children in the ability only attributional feedback showed the highest skill development and self-efficacy while the remaining two treatment groups were not different from each other but did outscore the no attributional feedback group (Schunk, 1983a). This finding supports Bandura’s theory that performance accomplishment is the most influential source of self-efficacy and gives support for the use of specific ability feedback with students in the classroom.

Subsequent examination of providing both effort and ability feedback in a different sequence showed that the group who received the ability feedback early
exhibited higher self-efficacy and higher skill level attainment than did the group receiving the effort feedback early (Schunk, 1984b). It was suggested that the group who received the effort feedback early, upon hearing their success was due to effort, held the subsequent ability feedback with little regard. A similar study conducted by Schunk and Cox (1986) used effort feedback in conjunction with a verbalization process to influence the self-efficacy and skill performance of learning disabled students. The children were randomly assigned within sex to one of nine conditions: one of three verbalization conditions (continuous verbalization, discontinued verbalization, no verbalization) combined with one of three effort feedback routines (first half of training, second half of training, no effort feedback). Continuous effort feedback was found to lead to higher self-efficacy and skill development than the other two verbalization conditions and effort feedback increased the skill development more than no feedback (Schunk & Cox, 1986). Additional research using specific task strategies, such as the verbalization in this study, provided similar results (Schunk & Gunn, 1986; Schunk & Rice, 1987).

Schunk and Rice (1987) found that providing multiple sources of strategy value feedback information enhanced remedial readers’ self-efficacy. Cognitive strategies were operationalized as systematic plans toward improving performance and students who received multiple sources of strategy information had higher judgments of self-efficacy. It should be noted that in all of the studies conducted in this area, teachers referred the students for participation (Schunk, 1981, 1982, 1983a, 1984b; Schunk & Cox, 1986; Schunk & Gunn, 1986; Schunk & Rice, 1987). While this method provides specific target groups, more randomized selection procedures are preferred, as the results could be more generalizable.
Other types of variables were also examined in relation to performance feedback and the effects on self-efficacy. Cervone and Palmer (1990) examined the effects of anchoring biases on judgments of self-efficacy and perceived self-efficacy levels as participants received feedback about their performance. Anchor was operationalized as in previously discussed studies; an initial value adjusted upward or downward by the participant, often in the direction of the original initial value. The subjects were undergraduate students randomly assigned to one of four conditions: low anchor, intermediate anchor, high anchor, or no anchor control condition. The anchors were intended to bias the self-efficacy judgments of the participants in the direction of the anchor (Cervone, 1989; Cervone & Peake, 1986). Self-efficacy judgments were strongly affected by anchoring biases with repeated-measures results showing the anchoring by trial interaction to be significant. The sex by trial interaction was also significant meaning that there is a joint effect between sex and trial that is stronger than the individual main effects of each variable considered independently.

Further analyses of variances showed that anchoring significantly affected initial judgments of self-efficacy. Significant differences were also found in sex and initial level of self-efficacy. Initial level of self-efficacy significantly predicted subsequent efficacy judgments in all groups but the control (low anchor, \( r = .70 \), intermediate anchor, \( r = .48 \), high anchor, \( r = .38 \)) (Cervone & Palmer, 1990). Results corroborated earlier findings (Cervone & Peake, 1986) that anchoring biases affect judgments of self-efficacy (Cervone & Palmer, 1990). The performance feedback about successes and failures affected the self-efficacy perception but initial levels of self-efficacy persisted. Results from this experiment indicated that the performance experiences did not completely
eliminate the initial self-efficacy beliefs. This is an important finding when considering the effects of self-efficacy of failure-experienced students, particularly in mathematics. While performance accomplishments are the most influential source of self-efficacy, this study indicates that single successful experiences are not adequate to increase self-efficacy, a phenomenon also noted by Bandura (1989, 1993). These findings have important implications for high school education. Most performance feedback at the high school level is in the form of assessments such as quizzes or exams, and not very frequent within a semester course. For students with high self-efficacy, this amount and type of feedback may be appropriate. However, these findings might be useful in attempts to increase achievement of students with low self-efficacy.

**Self-Efficacy and Goal Setting**

Goal setting is another major variable that has been examined in relation to self-efficacy (Bandura & Cervone, 1983; Bandura & Schunk, 1981; Schunk, 1984a, 1985, 1996; Zimmerman & Bandura, 1994). Self-efficacy has been found to affect performances cognitively both directly and indirectly through goal setting (Bandura, 1993). Prior to task performance, individuals mentally choose a course of action based on previous experience and evaluation of their ability to be successful. Bandura and Schunk (1981) tested the effects of the types of goals on self-directed learning. The researchers investigated the effects of proximal goals, distal goals, or no goals on self-directed learning. Proximal goals were suggested to the students in the form of setting a goal of completing at least 6 pages of the problem set during each session. Seven sessions were scheduled using a total of 42 pages of problems. The distal goals suggested to the students were to complete the entire 42 pages by the end of the seven
sessions. Results indicated that the participants in the proximal goals group achieved higher mastery of mathematical operations than the other two groups. In addition, the participants progressed rapidly and developed a sense of personal efficacy and intrinsic interest in previously uninteresting mathematical activities. The researchers attributed proximal goals as having at least three psychological effects on behavior. The first is motivational in that the proximal goals are attainable and provide immediate incentives for behavior. The second psychological effect is the development of personal self-efficacy, which helps determine the amount of effort and persistence that is applied to activities, a notion previously noted in the literature (Bandura, 1977b; Bandura et al., 1977). The third psychological effect is the standards that subgoals provide for people to judge themselves against. Achieving smaller subgoals as part of an overall larger goal enhances self-efficacy as people have a sense of accomplishment (Bandura & Schunk, 1981). These findings have implications for using proximal goals to increase the self-efficacy of high school math students, particularly those who have a high number of failure experiences.

Schunk (1985) also reported that student participation in goal setting led to the highest self-efficacy and skill enhancement when compared to groups with proximal goals or no goals. Teachers identified learning disabled children for participation in the study where they were randomly assigned to one of three treatment groups: participation in goal setting, proximal goal setting, or no goals. All groups received the skills training and results indicated that the children who participated in the goal setting established higher initial expectations than the proximal goals group. Schunk (1985) suggested that this difference may have lowered the initial expectation of the proximal group and...
affected their subsequent skills development. These results conflict with earlier findings by Bandura and Schunk (1981) regarding self-efficacy and behavior but may be attributed to the different populations sampled here (learning disabled versus nondisabled), according to Schunk (1985).

Perceived self-efficacy was also found to influence academic achievement through its impact on goal setting (Schunk, 1984a, 1985, 1986; Zimmerman & Bandura, 1994). Combining proximal goal setting with rewards for performance, Schunk (1984a) tested the effects of these variables on self-efficacy and skill performance. Elementary school children were identified based on their lack of division skills and randomly assigned to one of three treatment groups: rewards only, goals only, or rewards and goals. The order of the treatment was the same in each of the three schools (goals only, rewards only, rewards and goals) to avoid any discouragement for students not receiving rewards. The results showed that combining the rewards for performance with the proximal goals increased self-efficacy and performance skills more than the treatments separately. Schunk (1984a) suggested that future research with these variables measure the initial sense of self-efficacy to determine if the proximal goals and rewards for performance enhance skills through a common informational process of making the children aware of their initial sense of self-efficacy.

Other research has examined the influence of goal setting on academic performance, however, not specifically mathematics performance (Kitsantas, Reiser, & Doster, 2004; Page-Voth & Graham, 1999; Schunk & Gaa, 1981; Schunk & Swartz, 1993; Zimmerman & Kitsantas, 1997). Schunk and Gaa (1981) identified several influences of goal setting on academic performance. Students who set goals are more
likely to expend more effort and persist which increases their time on task. These characteristics have also been associated with students who have high self-efficacy. Students use goals to compare and adjust their progress toward accomplishment of a specific task. Goals are most effective when they are specific, difficult but attainable, and close in time. In addition, two types of goal setting were identified, process and product goals. Product goals are those in which a product is created or achieved, such as a project. Process goals are tasks that must be accomplished in order that a product goal is attained, such as turning in an outline of the project on time. These process goals are often generalizable skills to other product goals (Schunk & Gaa, 1981).

A study by Schunk and Swartz (1993) on a writing strategy provides additional support for using process rather than product goals. The researchers found that the writing strategy implemented was more beneficial to the quality of writing than the paragraph goal. In addition to this finding, the use of feedback in conjunction with the strategy goal increased achievement outcomes and transfer of knowledge. It is possible that the students perceived the strategy as more useful, seeing its potential for application in other settings. This finding is similar to that of Zimmerman and Kitsantas (1997). Using process goals, product (outcome) goals, and a shift from process to product goals, Zimmerman and Kitsantas (1997) found that the group who shifted from process to product goals showed the greatest self-efficacy; followed by the group who focused on process goals, and then the group that focused only on the product goals.

While performance can be increased by the use of certain types of goals, self-efficacy is not necessarily increased (Page-Voth & Graham, 1999). Page-Voth and Graham (1999) found that the use of goals when writing papers produced higher quality
work than those without goals but did not find any effect on writing self-efficacy. The reverse has also been found, as exhibited by the findings of Button, Mathieu, and Aikin (1996). Even if self-efficacy is increased, performance is not always increased. Their study showed that assigned goals and self-efficacy had no effect on performance or personal goals (Button et al, 1996). In addition to this work, West and Thorn (2001) also found that goal setting did not affect performance but for the younger adults, the goal setting did increase self-efficacy. These findings contradict those of Page-Voth and Graham who found an effect on performance but not on self-efficacy.

When goal setting is combined with self-evaluation, motivation and outcomes can be affected. Schunk (1996) randomly assigned students within sex, ethnic background, and classroom, to one of four conditions: learning goal with self-evaluation, learning goal without self-evaluation, performance goal with self-evaluation, or performance goal without self-evaluation. The learning goal was defined as how to solve problems and the performance goal was merely to solve the problems. The two groups with self-evaluation and the learning goal without self-evaluation led to higher self-efficacy, skill, and motivation than the remaining group of performance goal without self-evaluation (Schunk, 1996).

More recent work in the area of goal setting examined the effects of goal setting, self-evaluation, and organizational signals on ability and self-efficacy (Kitsantas et al., 2004). As with previous work in this area, the students in the process goals group had a higher self-efficacy than the students in the product goals group. The self-evaluation was also shown to have influence on skills, but this was greater for the students in the product goals group. The last variable, organizational signals, showed no effect on skill or self-
efficacy (Kitsantas et al., 2004). The implications of goal setting for teaching are important to note here. Goals appear to increase achievement toward a particular skill, if they are specific enough and focus on a process rather than a product (Schunk, 2003). Teachers can use models in the classroom and increase self-efficacy by providing successful experiences upon which the students can build. They can also help students set appropriate goals and self-evaluation techniques to assess progress toward those goals (Schunk, 2003).

**Self-Efficacy, Evaluative Feedback, and Goal Setting**

The effectiveness of combining evaluative feedback and goal setting was also shown in different settings. For example, Bandura and Cervone (1983) examined the mediating effects of self-efficacy on goal setting with performance motivation. The researchers used four treatment groups: goals and performance feedback, goals alone, feedback alone, and no goals or feedback. The group that set goals and received performance feedback showed the largest motivational impact. Neither the goals alone nor feedback alone impacted motivational levels. Bandura and Cervone (1986) conducted further research in this area, examining the relationship between accomplishments and perceived self-efficacy and personal goal setting. The results of this research indicated that the discrepancy between a self-set standard and attainment influences motivation. Motivation was best maintained by strong self-efficacy and further observations indicated that most of those who fell short of a self-set standard continued to hold to that standard or one slightly lower. However, no increase in effort was observed with individuals who experienced both low self-efficacy and failure to meet low self-set standards (Bandura & Cervone, 1986).
There is converging evidence that self-efficacy and personal goals enhance performance and motivation through a series of meta analyses, contrary to beliefs that self-efficacy and personal goals are self-debilitating (Bandura & Locke, 2003). Many studies concerning self-efficacy, goals setting, feedback, and academic performance have been conducted, but of particular interest is the research in the area of mathematics achievement at the high school level. Malpass, O’Neil, and Hocevar (1999) studied the effects of several variables, including learning goal orientation, self-efficacy, self-regulation, and math achievement, on an Advanced Placement calculus exam of a group of mathematically gifted high school students. A structural equation model was used to analyze the intercorrelations among the variables. Learning goal orientation was operationalized as a focus on progress and mastery through effort or as a focus on ability. Results showed a significant positive relationship between learning goal orientation and self-regulation ($\beta = .448$), but not between learning goal orientation and self-efficacy or math achievement. These findings were expected for the relationship between learning goal orientation and self-regulation in that gifted students are more learning goal oriented, however, the lack of a significant relationship between learning goal orientation and self-efficacy was not expected. The researchers expected a positive relationship between these two variables as well, since self-regulation is key to high self-efficacy.

Other research in this area has been conducted using classroom strategies to influence course grades. Teachers provided daily feedback to failure-experienced math students using a 0 to 4 rating scale for four performance areas: conduct, understanding, math, and persistence). Results indicated that the feedback directly affected the student’s approach to learning mathematics, which indirectly led to better learning and
understanding of mathematics but no control group was used in the study (Tzur & Movshovitz-Hadar, 1998). With the current emphasis on reducing the achievement gap between lower and upper level students, these results with failure-experienced students are encouraging, but the study has several limitations. First, the study was conducted in Israel and therefore, not necessarily generalizable to American high school students. Second, it is not possible to establish a causal relationship between use of feedback and the increase in math achievement without a control group. Also, the study did not provide any inferential statistics for comparison to other research. However, the study does offer some information about the use of feedback with failure-experienced students that can be used in future research.

Using both goal setting and feedback as variables to predict fraction skill, Shih and Alexander (2000) conducted an experiment using fourth-grade Taiwanese children. Students were randomly assigned to one of four treatment groups: goal setting with self-referenced feedback, goals setting with social-referenced feedback, self-referenced feedback only, and social-referenced feedback only (Shih & Alexander, 2000). The students were randomly assigned within and across three classrooms and pretest measures showed no significant differences among data based on gender or classrooms. Measures of goal orientation, self-efficacy, and fraction skill were included on the pretest. Subjects in the two goal setting groups were asked to set a goal for themselves. One of the groups was given information about their previous performance (self-referenced feedback) while the other group was given information about the average of a group of similar students (social-referenced feedback). The remaining two groups received the information (self-referenced or social-referenced feedback) but were not
asked to set any goals (Shih & Alexander, 2000). Shih and Alexander (2000)
hypothesized that children in the goal-setting groups would demonstrate higher fraction
skill and self-efficacy than children in the no goal groups and that the children in the self-referenced group would demonstrate higher fraction skill and self-efficacy than the children in the social-referenced group.

Contrary to their expectations, findings from their study indicated that goal setting had no effect on fraction skill or self-efficacy (Shih & Alexander, 2000). These results differed from those found by Bandura & Wood (1989). The lack of evidence to corroborate previous experiments could have been due to a cultural difference since the study was conducted in a non-Western country (Taiwan). It is possible that the children in the no-goal group set their own goals even though they were not asked to do so and that this is due to a cultural difference (Shih & Alexander, 2000). Results did exhibit a significant difference at the posttest, however, for the children in the self-referenced feedback groups (M = 96.42) with higher fraction skill and self-efficacy than children in the social-referenced feedback groups (M = 91.38) (Shih & Alexander, 2000). This finding supports earlier ideas that performance feedback provides a motivational source for students to engage in the tasks (Bandura, 1977). Such engagement could likely have led to greater levels of self-efficacy as well as higher fraction skills (Shih & Alexander, 2000).

Demographic and Psychological Variables

Other variables affecting self-efficacy and academic achievement are described in this section. Of particular interest is the possible influence of some of these variables on math achievement indirectly through influences on math self-efficacy (Campbell &
Pajares and Johnson (1996) suggested in their study of self-efficacy and writing performance that future studies include socioeconomic status (SES) as a variable. They indicated a belief that SES would be a strong predictor of academic performance. However, using path analysis on a study of 279 children, Bandura, Barbaranelli, Caprara, and Pastorelli (2001) found that a family’s SES had no direct effect on children’s perceived self-efficacy, academic aspirations and achievement, or perceived occupational efficacy. The researchers did find that the influence of SES was entirely mediated through the perceived efficacy of the parents to promote the academic development of their children and their educational aspirations for their children. Furthermore, the researchers found that the perceived efficacies of the parents and their academic aspirations for their children had no direct affect effect on the children’s career aspirations but parental academic aspirations were positively linked to their children’s perceived self-efficacy (Bandura et al., 2001).

Other measured variables demonstrated stronger path coefficients than SES in this study (Bandura et al., 2001). For example, parental academic aspirations had a direct influence on academic achievement with a path coefficient of .35 and also a direct influence on children’s academic aspirations with a path coefficient of .54. As with Zimmerman et al. (1992), parental factors were influential on academic achievement, however, SES was entirely mediated through parental academic aspirations and prosocial behavior. While the inclusion of additional factors in this study seemed to explain more of the variance in academic achievement, the implications for practical use need further
exploration. In the current American public educational setting, increasing students’ academic achievement is a primary goal and identifying factors over which the school system has influence would be more beneficial. While identification of SES could be used as an indicator for identifying students with potentially low academic achievement, discovery of strategies for affecting change in academic self-efficacy to influence academic achievement would be more beneficial and practical.

Several studies have found links between math anxiety and math achievement (Betz, 1978; Betz & Hackett, 1983; Hackett & Betz, 1989). Math anxiety can be defined as “feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems in a wide variety of ordinary life and academic situations” according to Richardson & Suinn (1972, p. 551). Math anxiety was shown to be moderately negatively related to achievement, in that high achievement in math was related to lower levels of math anxiety (Betz, 1978; Betz & Hackett, 1983). Hackett and Betz (1989) viewed math anxiety as a result of low mathematics self-efficacy. Cooper and Robinson (1991) found that math self-efficacy, mathematics ability, mathematics anxiety, and level of support from teachers and parents were significantly related to mathematics performance. Results were similar to those reported by others, with mathematics self-efficacy having a correlation with mathematics anxiety of $r = -.41$ (Cooper & Robinson, 1991). Multiple regression analyses showed that ACT scores accounted for 48% of the variance in mathematics performance and mathematics anxiety accounted for an additional 21%, a significant amount of the total variance explained (Cooper & Robinson, 1991).
Few, if any, studies found significant differences between boys and girls upon examination of previous math achievement or math level. Various studies examined math achievement data or self-efficacy for sex-related differences in addition to other variables of interest but few found significant differences between boys and girls (Bong, 1999; Fennema & Sherman, 1977; Hackett & Betz, 1989; Junge & Dretzke, 1995; Randhawa, Beamer, & Lundberg, 1993). In a study of high school students, researchers gathered data on two math attitude measures, three math self-efficacy scales, and a mathematics achievement test (Randhawa et al., 1993). Utilizing a structural equation modeling approach, they indicated that while the same pattern of relationships existed among the variables for both boys and girls, the strength for the boys was greater than that of the girls but suggested that this needed further investigation to be conclusive. The path model showed significantly stronger variance in math achievement explained for the boys ($R^2 = .51$) than for the girls ($R^2 = .42$).

This contradicts early research by Fennema and Sherman (1977), who investigated sex-related differences in mathematics performance using several variables, some cognitive and some affective. The participants were high school students in grades 9 through 12. The results indicated that the mathematical background of the students, for example, how many math courses taken, explained any differences due to sex. By contrast, Junge and Dretzke (1995) found that males had a stronger self-efficacy than females in their study. The researchers reported that males were higher on almost all of the items of mathematics self-efficacy, even when the items were not significant. The math self-efficacy ratings were not compared to any performance outcome for accuracy. The authors noted that even though the male scores were higher, overall both sets of
scores were relatively high due to the gifted/talented nature of the sample (Junge & Dretzke, 1995). Bong (1999) attributed difference in academic self-efficacy based on biological gender to the specificity or global nature of the instrument used to measure self-efficacy. When problem-specific self-efficacy measures are used, there is not a statistical difference between males and females (Bong, 1999).

Overall, there is a lack of evidence that biological gender is a significant explanatory variable in observed differences between males and females. This lack of explanatory power of biological gender perhaps could be due to a more psychological foundation. Several studies demonstrated that gender role rather than biological gender is an alternate explanation for the observed differences between males and females. Gender role is defined as the extent to which individuals use culturally defined standards for masculinity and femininity to evaluate their personality and behavior (Bem, 1981). Traditionally, gender role has two dimensions, Masculinity and Femininity. Masculine traits are more strongly associated with characteristics such as ambitious, self-reliant, and independent while Feminine traits are more strongly associated with characteristics such as gentle, affectionate, and understanding. Theoretically, Masculinity and Femininity are independent of each other, indicating that a person can have both traits (Constantinople, 1973). This conceptualization allows a classification of people into four categories: Masculine, Feminine, Androgynous (persons with high scores on both), and Undifferentiated (persons with low scores on both) (Bem, 1981).

Some empirical research has established that self-efficacy has a stronger relationship with Masculinity than Femininity (Choi, 2003). Choi (2003) indicated the relation between the goal-orientations and achievement-orientations of self-efficacy and
masculine traits is stronger than for the feminine traits. Choi (2004) also found that the association between Masculinity and self-efficacy becomes stronger as the self-efficacy becomes more general. The variance with self-efficacy and masculinity was over 30%, much stronger than the 5% variance between Femininity and self-efficacy (Choi, 2004). These studies indicate that a difference in self-efficacy is related to a difference in psychological gender rather than biological gender.

As previously noted, one of the most powerful sources of self-efficacy is performance accomplishments (Bandura, 1982b; Lent et al., 1991; Lopez & Lent, 1992). Effects of previous achievements were studied in different domains, for example in writing self-efficacy where previous writing achievement was measured as writing aptitude and showed a strong direct effect on performance (Pajares & Johnson, 1996; Pajares & Valiante, 1997). Expectedly, various forms of previous math achievement were used as independent variables in numerous studies to either control for differences in groups or to help predict math self-efficacy (Hackett & Betz, 1989; Lent et al., 1991, 1993; Pajares, 1996a; Pajares & Graham, 1999; Pajares & Kranzler, 1995; Pajares & Miller, 1994; Siegel et al., 1985). For example, Pajares (1996a) used the first-semester grade point average (GPA) in algebra class as one of the variables in a path analysis model predicting mathematics problem-solving performance. Path coefficients of .293 and .249 from math GPA to mathematics performance, respectively, for both regular and gifted students, were significant. While path coefficients from math self-efficacy to math performance were also significant for both regular and gifted students, .387 and .455, respectively, path coefficients from math GPA to math self-efficacy were not. There was a significant direct effect from math GPA to math self-efficacy with a path coefficient of
-.389 for regular students, however, there was not a significant direct effect observed for the gifted students. These findings suggest that the influence of prior math achievement can be mediated by math self-efficacy (Pajares, 1996a). These same conclusions were found in other studies (Lent et al., 1991, 1993; Pajares & Kranzler, 1995; Pajares & Miller, 1994).

Previous math performance also seems to mediate gender effect on self-efficacy (Lent et al., 1991). In a statistical analysis of the relationships between sources of self-efficacy information to math self-efficacy, Lent et al., (1991) found that when self-efficacy on the source variables was regressed with gender, gender was no longer a significant predictor of math self-efficacy. The researchers suggested that the effects of gender on self-efficacy were mediated by differences in past performance (Lent et al., 1991). Other areas of self-efficacy, such as writing and vocational interests, have shown similar influences (Lent, Brown, & Larkin, 1984, 1986; Pajares & Valiante, 1997). Statistical analyses by Pajares and Valiante (1997) revealed that sex differences in writing performance were a function of writing aptitude, a measure of previous achievement. When men and women had comparable prior coursework, gender differences in academic self-efficacy were nonsignificant (Lent et al., 1984, 1986).

Summary

Math achievement is considered to be an academic filter, keeping students who lack mathematics skills from entering lucrative career paths. This mathematics skills deficit is creating a vast job market crisis as well, with less available and less qualified people within certain occupations in the United States, for example, engineering. This review of the literature, beginning with Bandura’s social cognitive learning theory and
specifically regarding the findings of math self-efficacy as it relates to math performance, provides a framework for additional research on the relationship between math self-efficacy and mathematics achievement. A review of the literature shows that in addition to the direct sources of self-efficacy, other variables can influence self-efficacy directly and indirectly. These variables include goals setting, feedback, previous academic achievement, and gender role orientation. Some of these variables can be manipulated or controlled in order to increase self-efficacy, and ultimately academic achievement, while others cannot be changed. Gender role orientation and prior math achievement, for example, are related to math self-efficacy. The specific nature of the relationship is based in part on performance accomplishments, as in the case of using prior math achievement in the formation of one’s math self-efficacy, however, changing prior math achievement is not possible. It is possible, on the contrary, to create positive, successful performance accomplishments that could be used to form a higher math self-efficacy. In the case of gender role orientation, higher Masculinity scores exhibit some characteristics of higher self-efficacy, such as confidence and achievement, and are reported to be positively correlated with self-efficacy scales (Hackett & Betz, 1989; Pajares, 2002).

Self-efficacy is the perceived capabilities one possesses to perform a given task or exhibit a behavior. People make judgments about their capabilities to accomplish a particular task and, based on what the perceived outcome is, choose to exhibit behavior accordingly. Four sources of self-efficacy have been identified: performance accomplishments, vicarious experiences, verbal persuasion, and physiological responses. Depending upon these sources, self-efficacy is developed within specific domains of human behavior. In particular, there is an abundance of research with regard to academic
self-efficacy and academic achievement. It is well established in the literature that self-efficacy is more predictive of academic achievement than other variables, such as cognitive ability. This is true in specific domains of academic achievement, particularly in mathematics. Specifically, it has been shown that math self-efficacy is more predictive of math performance than general mental ability (Pajares & Kranzler, 1995).

Goal setting has been shown to be effective in increasing math self-efficacy if they are specific, proximal, and difficult enough to be challenging (Pajares, 2002). In addition, it has been found that students who participated in setting their own goals experienced a greater increase in math self-efficacy. Setting goals that are challenging without being too difficult takes practice. The danger in setting goals that are too difficult is the failure to achieve them, perhaps eventually lowering one’s self-efficacy. On the other hand, setting goals that are too easily accomplished lowers the ultimate accomplishment value, also resulting in a lowered self-efficacy. Having students set goals has a positive potential influence on math self-efficacy in that the goals force more objective evaluation against a set standard. This aids students in developing more realistic and more self-regulatory functions as setting goals becomes more automatic.

Specific types of feedback, for example self-referenced feedback, have increased math self-efficacy of regular and low level achieving students. Self-referenced feedback has been shown to increase math self-efficacy more than social-referenced feedback, however, there has not been research to establish that feedback is better than nothing at all. There has not been research, for example, that uses two treatment groups receiving self-referenced and social-referenced feedback compared to a control group with no feedback in a high school math course. Neither has there been a study conducted over a
period of time with repeated measures of math self-efficacy to establish a causal relationship between type of feedback and increases in math self-efficacy.

While the link between math self-efficacy and math achievement has been established, using the relationship to influence math self-efficacy in order to increase math achievement has not been studied. Pajares (1996b) suggested “. . . more experimental designs are required in which self-efficacy beliefs are altered and the effects of these changes on academic attainments measured.” He recommended, among other methods, the use of goals, performance feedback and social comparison information (Pajares, 1996a, 1996b). Pajares (1996a, 1996b) further encouraged the use of longitudinal or repeated measures designs that the causal influence of self-efficacy might be established as well allow for interpretations of the reciprocal influence of self-efficacy and academic achievements. Clearly there is a need for more rigorous study design in which an independent variable is manipulated with more than one measurement of self-efficacy in order to assess the differential patterns in self-efficacy across different treatment groups. To address this call for a longitudinal study, the following research questions are proposed.

Research Questions

1. How does each student’s math self-efficacy rating change over time?
2. Do changes in students’ math self-efficacy ratings vary by feedback treatment?
3. How does each student’s grade self-efficacy rating change over time?
4. Do changes in students’ grade self-efficacy ratings vary by feedback treatment?
5. How does each student’s math achievement (measured by the Diagnostic
Exam) change over time?

6. Do changes in students’ math achievement (measured by the Diagnostic Exam) vary by feedback treatment?
CHAPTER III

METHOD

This chapter contains the purpose of the study, the measurement of the variables involved, the design of the study, description of the participants, and the procedures for implementation of the study.

Purpose

This study aimed to serve three purposes. The first purpose was to extend previous research in a different setting by using high school participants. The second purpose was to demonstrate that the use of specific types of feedback could increase math self-efficacy that would, in turn, increase math achievement. Third, this study sought to examine trends in self-efficacy and its effects on math achievement over a longer time interval than what was often reported in the literature.

Research Questions

The study was designed to answer these research questions, previously noted.

1. How does each student’s math self-efficacy rating change over time?
2. Do changes in students’ math self-efficacy ratings vary by feedback treatment?
3. How does each student’s grade self-efficacy rating change over time?
4. Do changes in students’ grade self-efficacy ratings vary by feedback treatment?
5. How does each student’s math achievement (measured by the Diagnostic
Exam) change over time?

6. Do changes in students’ math achievement (measured by the Diagnostic Exam) vary by feedback treatment?

Participants

The target population consisted of ninth-grade students in lower level mathematics entering suburban high schools in the United States. A sample of two ninth-grade Algebra One classes at a suburban high school in Kentucky was used for the study. In this school system, the eighth-grade math teacher recommended the level of ninth-grade math class to which students were assigned. The options were, in order from least to most difficult, Algebra One, Freshman Geometry, and Advanced Freshman Geometry. The students were assigned randomly to the two Algebra One classes by the school computer program and the classes contained 22 and 24 students for a total sample size of 46. At this school, the total number of students was approximately 1300 students with 340 assigned to the ninth grade. The average age of the ninth-grade students was 14.5 years old, ranging from 13.1 to 15.7 years old. Socioeconomic status of a school is often described by the percent of students participating in the free lunch program. For the 2005-2006 school year, this school had approximately 16.7 percent of its student population participating in the free lunch program. The gender composition of the ninth grade class was approximately 52.5 percent male and 47.5 percent female. The ethnicity of the ninth grade class for the 2006-07 year consisted of slightly over 90% White/Caucasian students, approximately 4% each Hispanic and African-American students, and less than 1% each Native American students and other ethnicities not listed.
Design

This research study was quasi-experimental and used repeated measures with a control group and two treatment groups. Two intact classes of ninth-grade Algebra One students were used for the research. Most of the ninth-grade students recommended for Algebra One in the ninth grade had taken Prealgebra in the eighth grade while a small number of ninth-graders had taken Algebra One in the eighth grade and were told to repeat Algebra One in the ninth grade. The school used a block schedule so the Algebra One course used for the study began in August 2006 and was completed by the middle of the school year (January, 2007).

One purpose of this study was to increase math self-efficacy. The Algebra One course was chosen since this was a group of students who were more likely to have lower levels of self-efficacy. The Freshman Geometry and Advanced Freshman Geometry classes were designed for the upper level and gifted students. Typically, students taking classes at that level have higher self-efficacy than those students taking the lower level courses such as Algebra One.

One third of each of the Algebra One classes was used as a control and the remaining two thirds of each class as the treatment groups. There was a constant treatment of goal setting for all students and one independent variable of type of feedback. The independent variable had 3 conditions, self-referenced feedback, social-referenced feedback, and no feedback. One third of the students in each of the classes were assigned randomly to receive self-referenced feedback and another third in each of the classes were assigned randomly to receive social-referenced feedback. The remaining one third of each class received no feedback but did receive innocuous comments.
Variables

There were three dependent variables, math self-efficacy, grade self-efficacy, and math achievement. Grade self-efficacy was a dependent variable added based on suggestions from a well-respected researcher in the field of academic self-efficacy, F. Pajares (personal communication, July 3, 2006, see Appendix S). The independent variable manipulated was type of feedback. In addition, a constant treatment was goal setting. Additional measurements were general self-efficacy and two covariate measurements, previous math achievement and gender sex role. Each of the variables and measurements are described here.

Type of feedback. The variable of feedback was manipulated as a method of increasing students’ performance accomplishments, one of the most influential sources of self-efficacy (Bandura, 1977b). The three conditions of feedback were self-referenced feedback, social-referenced feedback, and no feedback. Self-referenced feedback was operationalized for this research study as providing information to the student compared to his or her previous performance or behavior (see Appendix M). Social-referenced feedback was operationalized for this research study as providing information to the student compared to the performance or behavior of others (see Appendix M). A placebo effect was implemented for the no feedback control group to make it appear that all students were receiving similar communication from the teacher. Written comments for the no feedback control group provided nonjudgmental remarks that did not reflect any comparison to previous performance or behavior (see Appendix M). The list of feedback comments in Appendix M was validated by asking several teachers and human resource individuals to mark the items as self-referenced feedback, social-referenced feedback, or
neither (see Appendix N). Classifications by this group agreed with 100 percent accuracy. The teacher for the study wrote one of the assigned types of feedback two to three times per week in an individual student notebook kept in the classroom, appropriately selected based on the performance of the student on the assigned task.

*Goal setting.* Goal setting has been shown to be an effective strategy for increasing academic achievement if it is specific, proximal, and of sufficient difficulty level. Furthermore, students who set their own goals have exhibited higher levels of achievement. Goal setting provides a standard by which individuals make judgments regarding their accomplishments. To provide this self-set standard, the students were asked to set a grade goal, from 0 to 100 percent, regarding the 5-minute checks for understanding. These grade goals were written and graphed by the students for each 5-minute check for understanding (see Appendix O). Actual scores were graphed next to the goals for comparison. The weekly graphs were printed on index cards and kept in a pocket in the student’s classroom notebook.

*Previous math achievement.* All students in public middle schools in Kentucky were required to take a state math assessment in the eighth grade. Previous math achievement was measured using the standard scale score of this assessment. The standard scale scores were obtained from the district office. The scale of the scores was continuous, with possible scores from 325 to 800, however, a score of 325 was considered a nonresponse and scores above 584 were considered a distinguished response on the assessment. Ninth-grade students who did not attend public middle school in the same school district as the site of the study did not have a score for this variable.
Instruments

The measurement of some of the variables used in this study required instruments that are described here. Some of the instruments are newly developed and others are well established in the literature.

*Kentucky High School Diagnostic Mathematics Assessments for Algebra One*

To assess math achievement, it is desirable for the instrument to be more generalizable rather than teacher-constructed or textbook-specific. For this reason, the instrument to assess math achievement for this research study was a random sampling of the multiple-choice items on the *Kentucky High School Diagnostic Mathematics Assessments for Algebra One* (Diagnostic Exam hereafter). The Diagnostic Exam consisted of five clusters with each cluster having eight or twelve multiple-choice questions and two or three open response questions. The five clusters were: 1) One-Variable Expressions, Equations, and Inequalities; 2) Properties of Lines and Graphing Lines; 3) Linear Functions; 4) Polynomial Expressions; and 5) Systems of Equations.

The entire Diagnostic Exam would require three class periods and approximately 4 hours to administer. Therefore, a random selection of the multiple-choice items was used to create a 20-item instrument for assessing math achievement in this study. A random selection of problems from each of the five clusters was conducted, selecting five of twelve multiple choice items from One-Variable Expressions, Equations, and Inequalities, and five of twelve from Properties of Lines and Graphing Lines. Four of eight multiple choice items were selected from Linear Functions, and three of eight from Polynomial Expressions and three of eight from Systems of Equations. Three versions of the Diagnostic Exam were available and the different versions contained the same type of
problem in the same order. For example, question 5 on all three versions would be the same question but with altered values in the problem. Therefore, once the random questions were selected for version one, the same questions were used from both the second and third versions so the resulting 20-item instruments contained the same content but with altered values (see Appendices A through E).

Math achievement was assessed monthly during the course of the semester for a total of six math achievement assessment scores using the 20-item Diagnostic Exams. The first and last assessments were version one (see Appendix A) and the remaining four assessments were different versions of the 20-item Diagnostic Exam. Therefore, a total number of versions needed for the study were five but only three versions existed so two additional versions of the assessment were created by altering the same problems from version one (see Appendices A through E). While the problems are very similar on each of the 20-item Diagnostic Exam versions, the values are altered for each problem in addition to appearing in a random order. The first and last administrations of the 20-item Diagnostic Exam were version one for consistency of measurement. Due to some typographical errors, however, versions two, three, and four presented to the students contained problems for which there were no correct answers. The scores reported for the Diagnostic Exam are percentages of the problems answered correctly.

 Development of the Diagnostic Exam. A panel of mathematics teachers and university professors from the state who were familiar with the state core content developed the questions. The Diagnostic Exam was designed to give teachers an indication of the depth of knowledge of mathematics for students as they progress through a sequence of math courses in preparation for the state assessment at the end of
their junior year. All three versions of the original full-length Diagnostic Exam were posted on the state education website for use by all teachers in the state.

Reliability of the Diagnostic Exam. The Diagnostic Exam was used in a pilot study with a large sample of Kentucky high school students but is still in its final stages of revision. Reliability information has not been published but the results from this study contain internal consistency reliability and test-retest reliability.

Validity of the Diagnostic Exam. The Diagnostic Exam seemed to have a satisfactory level of content validity. Close inspection of the key standards used to develop the Diagnostic Exam and the Algebra One framework at the high school for the site of the study indicated that, for the most part, there was an alignment between the two. Thirteen of the fifteen key skills identified in the key standards are specifically addressed in the Algebra One framework at the high school. The fact that the Algebra One framework aligned so well supported the content validity of the Diagnostic Exam.

Measurement of Math Self-Efficacy

There are several instruments that were developed to measure math self-efficacy. One of the most widely used instruments is the Math Self-Efficacy Scale (MSES; Betz & Hackett, 1983). In spite of its satisfactory psychometric properties, one problem of using the MSES relates to the disalignment between the specificity level of self-efficacy measured by such an instrument and the specificity level of criterial task performance. From a social cognitive perspective, in order for math self-efficacy to have good predictive capability requires that the self-efficacy be measured at a content specific level that matches the specificity level of the criterion variable (Pajares, 1996a; Pajares & Miller, 1994, 1995; Pajares & Urdan, 2006). Assessing global self-efficacy, for example,
by asking “How confident are you in solving math problems?” will not result in an accurate assessment of efficacy when the criterial task is to solve polynomial expressions. Both Bandura and Pajares repeatedly pointed out that the mismatch between the specificity level of the predictor variable and the specificity level of the criterion variable largely explain nonsignificant predictors that are prevalent in educational research.

To measure very content specific math self-efficacy, a group of researchers (Bong, 1997, 1998, 1999, 2002, Bong & Hocevar, 2002; Nielson & Moore, 2003; Pajares, 1996a; Pajares & Graham, 1999; Pajares & Miller, 1997) have used a different method of measuring math self-efficacy in which the level of specificity would correspond closely if not perfectly to the level of the criterion variable (typically math problem-solving performance). They simply showed the math problems that served as the criterion variable to the students and asked them to indicate their level of self-efficacy on each problem. For instance, Pajares (1996a) used 19 algebra problems and asked students to assess their degree of confidence to solve each of the 19 problems. Once math self-efficacy was assessed in that manner, the same 19 math problems were used to assess the students’ math performance. Similarly, Bong (1997, 1998, 1999, 2002) also assessed self-efficacy to solve certain problems on which students’ academic performance was subsequently assessed.

In the current study this new measurement method was adopted to measure math self-efficacy. Participants in this study were asked to rate their confidence on a 6-point scale with 1 indicating no confidence at all and 6 indicating complete confidence, preventing students from choosing a “middle” value which would cause them to not adequately discriminate their confidence level. The possible sums of the confidence
scores ranged from 20 to 120. The math self-efficacy measurement was given prior to each of the 20-item Diagnostic Exams for a total of six math self-efficacy measurements. To prevent response bias, there were five versions of the math self-efficacy measurement, with the reference problems matching each of the five 20-item Diagnostic Exams created to measure math achievement (see Appendices A through E). Each math self-efficacy measurement used the same confidence scale but problems appear in a different order and with altered values for each problem to match the Diagnostic Exam immediately following the math self-efficacy measurement (see Appendices F through J).

Reliability and validity of math self-efficacy measurement. The new measurement method seems to produce scores that are reliable. Pajares (1996a), for instance, reported an internal consistency reliability coefficient of .94 for solving algebra problems, which is more than acceptable. The high degree of reliability appears to be consistent across different test formats of math problems (e.g., multiple choice versus open ended). Pajares and Miller (1997) reported internal consistency reliability coefficients of .92 for each format.

Measurement of Grade Self-Efficacy

The inclusion of the measurement of grade self-efficacy was to strengthen the study by adding another dimension to the self-efficacy ratings. The rating measurement was constructed in a similar manner to that in the study by Zimmerman and Bandura (1994), where the students were asked to rate their certainty of achieving each of the grades listed, from F to A+, using a rating scale with 1 indicating no certainty and 6 indicating high certainty. This grade self-efficacy was measured at each of the six math self-efficacy measurement points (Appendix R) and the sum of the ratings was
calculated. Prior to completing the rating scale, the example to explain the scale was if you were highly certain you could do one sit-up, you would mark a “6” and if you were not certain at all that you could do 1000 sit-ups, you would mark a “1.”

*Children’s Self-Efficacy Scale*

Children’s Self-Efficacy Scale (CSES, Bandura, 1990) was used by permission (see Appendix W) to measure academic self-efficacy. Formerly referred to as the Multidimensional Scales of Perceived Self-Efficacy, the CSES was developed by Bandura (1990) to assess self-efficacy in academic domains of functioning. The current CSES contains 55 items and has nine domains, comprising nine subscales, which include self-efficacy in enlisting social resources, self-efficacy for academic achievement, self-efficacy for self-regulated learning, self-efficacy for leisure time skills and extracurricular activities, self-regulatory efficacy, self-efficacy to meet others’ expectations, social self-efficacy, self-assertive efficacy, and self-efficacy for enlisting parental and community support. The number of items for each subscale varies; there are four items for self-efficacy in enlisting social resources, nine items for self-efficacy for academic achievement, ten items for self-efficacy for self-regulated learning, eight items for self-efficacy for leisure time skills and extracurricular activities, eight items for self-regulatory efficacy, four items for self-efficacy to meet others’ expectations, four items for social self-efficacy, and eight items for self-assertive efficacy. Students respond using a 7-point scale with 1 indicating *cannot do at all*, 4 indicating *moderately can do*, and 7 indicating *highly certain can do* (see Appendix K).

*Reliability of the CSES.* The subscales of the CSES have been shown to be highly reliable. For example, Zimmerman et al. (1992) found Cronbach alpha coefficients of .87
for the self-regulated learning scale and .70 for the self-efficacy for academic achievement scale. Additional analysis of the CSES was conducted in more recent studies (Miller, Coombs, & Fuqua, 1999; Choi, Fuqua, & Griffin, 2001), providing somewhat similar internal consistency reliability coefficients for each of the nine dimensions. Specifically, the internal consistency reliability coefficients for the earlier (Miller et al., 1999) and later studies (Choi et al., 2001) were .60 and .63 for self-efficacy in enlisting social resources, .74 and .72 for self-efficacy for academic achievement, .87 and .86 for self-efficacy for self-regulated learning, .74 and .76 for self-efficacy for leisure time skills and extracurricular activities, .79 and .81 for self-regulatory efficacy, .73 and .81 for self-efficacy to meet others’ expectations, .83 and .76 for social self-efficacy, .84 and .79 for self-assertive efficacy, and .70 and .79 for self-efficacy for enlisting parental and community support. The two studies were conducted with samples from different populations, high school students and college freshmen, another indication of the reliability of the instrument. Good reliability signifies that the results from the CSES will be consistent when administered in different settings or at different times with the same group.

Validity of the Children’s Self-Efficacy Scale. The original scale contained items that were designed from learning strategies reported by high school students during structured interviews conducted by Zimmerman and Martinez-Pons (1986). Using factor analyses of teachers’ ratings and students’ interviews in a subsequent study, the items reported from the students were validated as a theoretical construct (Zimmerman & Martinez-Pons, 1988). Later psychometric work on the CSES confirmed the nine domains using factor analysis (Choi et al., 2001). Results of three different rotational
approaches in both the principal component and principal axis factor solutions were nearly identical, indicating that almost all of the items aligned with theoretically derived dimensions (Choi et al., 2001). Using a cutoff value of 1.00 for a minimum eigenvalue, the retained factors in the analysis explained approximately 58% of the variance.

This analysis indicates that the CSES has good construct validity, which indicates that this instrument is a good measure of what it purports to measure. This measure of general academic self-efficacy provides additional information that may explain individual differences (Choi, 2003). The CSES is a general self-efficacy scale measure that is highly correlated with more problem-specific self-efficacy scores, such as the math self-efficacy measurements described earlier, even though the general measures have not demonstrated the same strength of prediction in performance (Bong, 2002; Choi, 2005). The CSES provides evidence for the construct validity of the math self-efficacy measurements used in the study.

_Bem Sex Role Inventory (BSRI)_

The short form of the BSRI was used by permission (see Appendix X) to measure gender sex role. The BSRI is an inventory that was developed in 1974 to measure gender roles. The original version was based on a set of 60 personality characteristics, with 20 stereotypically feminine characteristics such as gentle, understanding, and affectionate; 20 stereotypically masculine characteristics such as ambitious, independent, and assertive; and 20 neutral items such as happy, truthful, and conceited, considered neither feminine nor masculine. Some of the masculine personality characteristics such as ambitious, self-reliant, and competitive correlate highly with higher self-efficacy ratings (Choi, 2004). Scores on the BSRI are based on ratings of each of the 60 items using a 7-
point scale (1=Never or almost never true to 7=Always or almost always true. Two subscale scores comprise the BSRI score, one for Femininity and one for Masculinity, since this inventory is based on two dimensions rather than one dimension. Two dimensions allow individuals to score independently on each of the subscales. One can receive high scores, for example, on both Femininity (F) and Masculinity (M), unlike earlier measures of gender roles where a high score on one dimension created a low score on the other dimension (Bem, 1974). Scores from the BSRI can be used to classify individuals into one of four categories using a median-split: Feminine (high F-low M), Masculine (low F-high M), Androgynous (high F-high M), or Undifferentiated (low F-low M).

A short form of the BSRI was developed after early criticism that not all of the items loaded on the expected factors (Pedhazur & Tetenbaum, 1979). The short form of the BSRI was developed using factor loadings analysis on the original form (Bem, 1981). It consists of half as many items in each subscale, 10 items each for Femininity, Masculinity, and neutral (see Appendix L for sample items).

Reliability of the BSRI. Bem reported the BSRI as having high internal consistency reliability. Coefficient alphas for two different samples of undergraduate females for the Femininity score were .75 and .78 for the original form and .84 for the short form and .78 for both samples of undergraduate males for the original form and .87 for the short form. Coefficient alphas for two different samples for females for the Masculinity score were .87 and .86 for the original form and .84 and .86 for the short form and .86 and .87 for both samples of the males for the original form and .85 for the short form. A test-retest reliability with a 4-week interval produced product-moment
correlations ranging from .76 to .94 for the original BSRI and from .76 to .91 on the short BSRI. The evidence shows that the short form exhibits higher reliability than the original BSRI.

**Validity of the BSRI.** According to Choi and Fuqua (2003), there has been an ongoing debate in terms of what the BSRI actually measures since its publication in 1974. One major issue with regard to validity of the long form BSRI, according to Choi and Fuqua (2003), concerns the dimensionality of sex role operationalized by the BSRI. They point out that many empirical factor analytic studies reported more than two factors indicating that there may be more than two dimensions of sex role. More recent study (Choi, Fuqua, & Newman, in press), in fact, supports the hierarchical nature of sex role suggested by Marsh and Myers (1986) who had delineated several subconstructs (such as shy, compassionate, dominant, decisive, athletic, self-sufficient, and interpersonal affect) under two global Masculine and Feminine constructs. The short form BSRI is said to be conceptually purer (Bem, 1981) as a measure of narrower constructs, such as instrumentality and expressiveness.

**Procedure**

*Prior to First Day*

One third of each of the two intact Algebra One classes was assigned randomly to receive self-referenced feedback, another third of each of the two intact Algebra One classes was assigned randomly to receive social-referenced feedback, and the remaining one third was the control group that received neutral comments instead of specific feedback (see Appendix M). The students were numbered in alphabetical order within each class and a random number process from a computer was used to randomly assign
the treatments. Within each class, the first eight numbers selected received the self-referenced feedback, the next eight numbers selected received the social-referenced feedback, and the remaining eight numbers were the control group. Eight digit identification numbers were created for each student and recorded in a safe place for compiling information after the study was completed. The identification numbers were placed on index cards and sealed in envelopes containing the student name on the outside.

The school uses a computerized scheduling system to assign students to classes so there is no researcher bias present in the assignment of students to the intact classes. Ninth grade students at this high school all had their electives scheduled for the same block of the day so there was less risk of the same type of student ending up being together by forced scheduling conflicts. If a particular elective, like band or agriculture, for example, were offered only one block of the day, this might force all of the students taking those particular electives into the same Algebra One class. This might produce inherently different groups. Since the freshmen all took electives the same block, there was minimal chance of this type of bias being present in the composition of the intact classes.

The International Review Board at the University of Louisville granted approval for the study (Appendix P) and the site of the study granted permission as well (see Appendix Q). The parents (or guardians) of the students gave informed consent for their children to participate in the study (see Appendix U). The informed consent letter was on colored paper attached to the individual student schedules at registration on August 4, 2006. A return box for signed consent letters was available at the registration and the
researcher was available for questions, located in another area of the building. The parent (or guardian) of any student who did not attend registration was contacted by mail between school registration and the first day of classes, approximately two weeks later by a third party. Students were asked to sign the assent form during the open house for freshmen on August 8, 2006 (see Appendix V) and these were collected by a third party. The study was described in general terms with explanation of the possible benefits to the students. Any student who did not attend the open house for freshmen or did not return the signed form was contacted by mail prior to the beginning of school. Missing parental consent and assent forms were sent home with the student by a third party and students were asked to return the forms.

First Day of the Study

On the first three days of school, a third party came to the classes and collected signed assent and consent forms. The study was described in general terms and started on the fourth day of the course. All students in both classes received a class notebook that was kept in crates in the classroom. Students were allowed to select their own color of notebook so the teacher did not know which notebook belonged to which student. Students were instructed to develop some sort of symbol or character they would recognize for identifying their notebook. No names were on any notebook and the envelopes containing the identification number to be used for the study were disseminated. The index card with the identification number was kept in the pocket inside the notebook for the duration of the semester. After selecting the notebook and putting the index card in the pocket, the empty envelopes were collected and students were cautioned about keeping the identification number and symbol on the notebook.
confidential. The class notebooks were kept near the teacher’s desk in plastic containers identified by class period.

After the notebooks were prepared, both classes of students were given the CSES (see Appendix K), followed by the BSRI (see Appendix L for sample items). Verbal instructions prior to the CSES used the example of tossing a paper wad into the garbage can and rating confidence for varying distances using the same 7-point scale of 1 as Cannot do at all to 7 as Highly certain can do. Students were asked to write their identification number at the top of the Children’s Self-Efficacy Scale instrument and not to put their name or any other identifying information on the instrument before rating their degree of confidence on each of the items.

After completing the CSES, they were given the BSRI and instructed to write their identification number at the top but no other identifying information. An example inventory was explained to the students using the rating scale of 1 as Never or almost never true to 7 Always or almost always true. Students were asked to rate each of the personality characteristics for themselves and not to leave any characteristic unmarked.

Second Day of Study

All students in both classes were asked to complete the math self-efficacy measurement for version one of the 20-item Diagnostic Exam on the following day of class (see Appendix F). The students were told that they were going to be asked to answer the problems after completing the self-efficacy measurement. Each of the complete problems was shown on the overhead projector for about 10 to 20 seconds and students asked to rate their confidence in solving the problem. Bong (1997; 1998; 1999; 2002) used this method in many studies of academic self-efficacy to prevent students
from trying to work the problem and focusing instead on their confidence in answering
the problem. The 6-point rating scale had values from 1 as no confidence at all to 6 as
complete confidence. Both classes were given version one of the 20-item Diagnostic
Exam (see Appendix A) to complete after returning the math self-efficacy measurement
for version one (see Appendix F). Students wrote their identification numbers on each of
the instruments.

Assessment Days

Four different versions of the 20-item Diagnostic Exam were given approximately
monthly throughout the course (see Appendices B through E). A comparable math self-
efficacy measurement to match the problems in each of the different versions was
completed prior to taking each 20-item Diagnostic Exam (see Appendices G through J).
The procedure for each math self-efficacy measurement was the same as previously
explained, students were asked to write their identification number at the top of the math
self-efficacy measurement and the 20-item Diagnostic Exam and told they would be
given the 20-item Diagnostic Exam after completing the math self-efficacy measurement.
Each problem was shown on the overhead projector for approximately 10 to 20 seconds
for the math self-efficacy measurement and students were asked to rate their confidence
for solving each problem. At the end of the course, all students completed the original
math self-efficacy measurement for version one of the 20-item Diagnostic Exam (see
Appendix F) prior to solving the problems on version one of the Diagnostic Exam. They
also completed a second CSES and BSRI.
Typical Day of Class

There were 5-minute checks for understanding two to three times per week consisting of two to five problems pertaining to the previous day’s lesson that could be completed in five minutes. There were not 5-minute checks on days of unit tests or Diagnostic Exams or if the lesson carried over to the next day. Students wrote their answers in the classroom notebook kept in the classroom in a specified plastic container for each class. Graphs were printed on index cards each week and used to graph goals and scores for the 5-minute checks. Students wrote only their identification numbers on these index cards and they were kept in the pocket in the notebook. Students wrote their goals for each 5-minute check in the notebooks and the teacher provided a written comment in the notebook for each 5-minute check completed by the student. The written comment was self-referenced feedback, social-referenced feedback, or a neutral innocuous comment (see Appendix M), according to the randomly assigned treatment.

On the days of 5-minute checks, the students each retrieved their classroom notebook from the specified container and read the written feedback (or comment) from the teacher. Next, the students located their score on the previous 5-minute check in their notebook and recorded this percentage on the weekly bar graph, using blue ink (see Appendix O). The weekly graph was printed on an index card and kept in the pocket inside the notebook. Before continuing, the students corrected the previous 5-minute check in their notebook. The students corrected their work by examining the solution guide on the overhead projector and comparing it to their own work. They were given four to five minutes to do this.
After correcting the previous 5-minute check, the students turned the page, wrote the date and their self-set goal for the 5-minute check for the day on the left side of the notebook. The students recorded this goal for the 5-minute check for the day on the weekly graph, using green ink. The index card was returned to the pocket inside the notebook. While the students were writing and graphing their goals, the teacher put the new 5-minute check set of problems for the day on the overhead projector. Once it was determined that all of the index cards were secured in the pockets of the notebooks, the 5-minute check set of problems for the day were presented on the overhead projector. On the right page of the notebook, the students worked the 5-minute check problems for the day from the overhead projector. The students returned their notebooks to the appropriate crate.

Daily lessons introduced the topic and included a variety of common educational practices such as lecture notes and examples on the overhead projector, manipulative activities, worksheets, small group work, partner work, guided independent practice, and assessments such as quizzes, open responses, or end of the unit exams. The two classes received the same instruction, including assignments and activities, for Algebra One by the same teacher. The researcher was a mathematics instructor certified in Kentucky to teach mathematics in grades eight through twelve and was the teacher for the two classes. Before the next 5-minute check, the teacher scored the 5-minute check problems and wrote a comment from the appropriate list (see Appendix M), according to the randomly assigned treatment, self-referenced feedback, social-referenced feedback, or neutral innocuous comment (control group).
Threats To Experimental Validity

Quasi-experiments in educational settings present difficult issues that must be addressed carefully. The purpose of the study was to use a long-term treatment to attempt to increase math self-efficacy in order to increase math achievement. Previous studies of academic self-efficacy have not been as long term and therefore several issues have not been addressed.

Threats to Construct Validity

*Threats due to experimenter effects.* In view of the fact that the researcher was the instructor for the two classes, the opportunity for bias toward the different treatment groups was present. The study was carried out in a double-blind manner, however, with neither the students nor the teacher having knowledge of specific treatments or owners of the notebooks. This reduced opportunities for potential experimenter bias effects. In addition, the BSRI, CSES, math self-efficacy, grade self-efficacy, and Diagnostic Exams were not scored until after the completion of the study to prevent any alteration in the delivery of the instruction.

*Threats due to reactive effects.* Other potential threats to the design of the study involved the participants. Students expect to receive feedback in the form of grades and also to receive written comments regarding their progress. Keeping the monthly math self-efficacy and math achievement scores unknown until the end of the course reduced the reactive effects from the participants. In other words, this prevented the participants from skewing their performances in a particular direction. By providing all students with written comments, feedback or neutral, the students all appeared to be receiving the same treatment, in effect reducing the possibility of the placebo effect. In addition, the
researcher used feedback comments from the list in Appendix M and this provided more uniform comments and of the correct nature, self-referenced, social-referenced, or neutral.

**Threats to Internal Validity**

_Threats due to extraneous variables._ It is possible other variables affecting long-term studies such as this one reduced the generalizability of the results. The personality of the instructor and the classroom environment, for example, are issues that have not been addressed in the literature due to the fact that most studies were short term and these variables did not cause an impact on the results. In the present study, the researcher is a veteran teacher and may have other qualities such as good classroom control or an effective teaching style that confounded the results of the manipulation of the independent variable. Future replications could implement strategies or controls to address the personality of the instructor or other such variables. For the purpose of this study, having the same instructor who was familiar with potential threats to the design was important. With regard to feedback, for instance, the researcher knew to avoid verbal comments containing comparisons to performance to keep from confounding the results. This might have been difficult to control if the research study were implemented by using other classroom teachers.

**Summary**

The study conducted was a rigorous double-blind design with one manipulated variable of feedback treatment and three dependent variables of math self-efficacy, grade self-efficacy, and math achievement. The results of the study are analyzed in the next chapter and include descriptive statistics, intercorrelation coefficients among the major
variables, and the Cronbach Alpha coefficients and test-retest reliability coefficients of the scores from the BSRI, CSES, math self-efficacy scales, grade self-efficacy scales, and the Algebra One Core Content Diagnostic Exam assessments.

Each of the six research questions was analyzed using multilevel modeling. The three questions concerning how the dependent variables change over time, for example, were answered using unconditional growth models in a Hierarchical Linear Model (HLM). A detailed description and additional information regarding HLM can be found in Appendix T. The remaining research questions involve the changes in each of the dependent variables by treatment group. Conditional growth models are used in HLM to analyze the effects of the feedback treatment on math self-efficacy, grade self-efficacy, and math achievement (measured by Diagnostic Exam).
CHAPTER IV

RESULTS

Results of the study are presented in this chapter in two sections. The first section provides general findings about the study, with preliminary statistical analyses regarding reliability, descriptive statistics, and correlations among the variables. The second section provides results with respect to the six research questions. Statistical analyses were performed using SPSS 15.0 and HLM 6.03.

Preliminary Statistical Analyses

Reliability Analyses

Internal consistency reliability. The internal consistency reliability for each of the scales used in this study was measured by obtaining Cronbach’s alpha. Scales used in this study include the Bem Sex Role Inventory (BSRI), the Children’s Self-Efficacy Scale (CSES), math self-efficacy scales, grade efficacy scales, and diagnostic exams. Table 1 contains the coefficient alpha estimates and the number of items for each of the scales. All coefficient alpha estimates ranged from .73 to .95, and the majority of the values were either slightly less than or greater than what were reported in the literature (Choi, Fuqua, & Griffin, 2001; Miller, Coombs, & Fuqua, 1999; Pajares, 1996a; Pajares & Miller, 1997), or from the developer (Bem, 1981). The estimates reflected were adequate, indicating that the students were adequately consistent in their responses.
Table 1

*Comparison of Reliability Coefficients*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of items</th>
<th>Current Study</th>
<th>Other Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BSRI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Femininity</td>
<td>10</td>
<td>.86</td>
<td>.84</td>
</tr>
<tr>
<td>Masculinity</td>
<td>10</td>
<td>.84</td>
<td>.87</td>
</tr>
<tr>
<td><strong>CSES overall</strong></td>
<td>55</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td><strong>CSES by subscales</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Resources</td>
<td>4</td>
<td>.73</td>
<td>.60/.63</td>
</tr>
<tr>
<td>Academic Achievement</td>
<td>9</td>
<td>.90</td>
<td>.74/.72</td>
</tr>
<tr>
<td>Self-Regulated Learning</td>
<td>10</td>
<td>.89</td>
<td>.87/.86</td>
</tr>
<tr>
<td>Leisure Time</td>
<td>8</td>
<td>.82</td>
<td>.74/.76</td>
</tr>
<tr>
<td>Self-Regulatory</td>
<td>8</td>
<td>.89</td>
<td>.79/.81</td>
</tr>
<tr>
<td>Others’ Expectations</td>
<td>4</td>
<td>.81</td>
<td>.73/.81</td>
</tr>
<tr>
<td>Social Self-Efficacy</td>
<td>4</td>
<td>.82</td>
<td>.83/.76</td>
</tr>
<tr>
<td>Self-Assertive Efficacy</td>
<td>4</td>
<td>.83</td>
<td>.84/.79</td>
</tr>
<tr>
<td>Parental Support</td>
<td>4</td>
<td>.80</td>
<td>.70/.79</td>
</tr>
<tr>
<td><strong>Math Self-Efficacy</strong></td>
<td>6</td>
<td>.93</td>
<td></td>
</tr>
<tr>
<td>Math SE #1</td>
<td>20</td>
<td>.92</td>
<td>.92/.94</td>
</tr>
<tr>
<td>Math SE #2</td>
<td>20</td>
<td>.94</td>
<td>.92/.94</td>
</tr>
<tr>
<td>Math SE #3</td>
<td>20</td>
<td>.92</td>
<td>.92/.94</td>
</tr>
</tbody>
</table>
Test-Retest. Both the BSRI and the CSES were administered twice, once at the beginning of the Algebra One course and a second time at the end of the course in order to estimate test-retest reliability coefficients. Test-retest correlation coefficients with a five-month interval illustrated acceptable stability, with significant test-retest reliability coefficients for both instruments. As can be seen in Table 2, test-retest reliability coefficients for the BSRI subscales of Femininity \( (r = .67, p < .01) \) and Masculinity \( (r = .61, p < .01) \) were significant. Bem (1981) reported test-retest reliabilities with a four-week interval of .85 and .91 for college age females and males, respectively, on the Femininity subscale, and .91 and .76 for females and males, respectively, on the Masculinity subscale. Another study reported .56 for Masculinity and .68 for Femininity with a four-year interval (Yanico, 1985). Even though the test-retest reliability coefficients for the current study were lower than those reported by Bem (1981), they
Table 2

*Test-Retest Reliability Estimates for BSRI and CSES*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reliability Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSRI</td>
<td></td>
</tr>
<tr>
<td>Femininity</td>
<td>.67</td>
</tr>
<tr>
<td>Masculinity</td>
<td>.61</td>
</tr>
<tr>
<td>CSES subscales</td>
<td></td>
</tr>
<tr>
<td>Social Resources</td>
<td>.62</td>
</tr>
<tr>
<td>Academic Achievement</td>
<td>.57</td>
</tr>
<tr>
<td>Self-Regulated Learning</td>
<td>.66</td>
</tr>
<tr>
<td>Leisure Time</td>
<td>.37</td>
</tr>
<tr>
<td>Self-Regulatory</td>
<td>.41</td>
</tr>
<tr>
<td>Others’ Expectations</td>
<td>.39</td>
</tr>
<tr>
<td>Social Self-Efficacy</td>
<td>.26</td>
</tr>
<tr>
<td>Self-Assertive Efficacy</td>
<td>.44</td>
</tr>
<tr>
<td>Parental Support</td>
<td>.59</td>
</tr>
</tbody>
</table>

show respectable stability. Additionally, no test-retest reliability coefficients of the BSRI scores obtained from ninth-graders have been reported in the literature. Therefore, the reliability coefficients with such a long time interval obtained in the current study makes a unique contribution to the sex roles literature.

Test-retest reliability coefficient estimates for the nine subscales of the CSES ranged from .26 to .66. Overall, these results were lower than expected. One reason for
these lower results may be the long time interval of five months between administrations of the instrument. Another reason for the low reliability estimates may be related to the nature of the constructs being measured. Unlike masculine and feminine personality traits, self-efficacy is situation and domain specific, which means that it can fluctuate depending on the circumstance (Bandura, 1986).

**Descriptive Statistics**

Means and standard deviations of the major variables appear in Table 3. Most of the variables showed comparable means and medians. This indicates that the frequency distributions for each variable were somewhat symmetric. Means are not resistant measures of central tendency, being drawn toward extreme values, while medians are resistant to extremes. Frequency distributions with a mean close in value to the median indicate a lack of extreme values in the distribution. The BSRI and CSES variables and subscales appeared to be symmetrically distributed based on the means and medians with the exception of the Self-Regulatory subscale of the CSES. The presence of lower minimum values is indicated since the mean (45.73) is less than the median (49.00).

The same is true for the measure of previous math achievement, the CATS Math scale score. The mean (548.03) is slightly less than the median (551.00), indicating a slightly negatively skewed distribution. It is important to note that the CATS Math scale score used in this study is from a state-mandated test given to all public school students in Kentucky, with this particular score from the eighth grade assessment. Eight of the participants in the current study either did not attend public schools in the eighth grade or did not attend the eighth grade in the school district of the site of the study. Since CATS Math scale scores could not be obtained for these eight students and the distribution was
somewhat symmetric to begin with, the missing scores were replaced with the mean value of the collected data. Descriptive statistics are reported in the table for the original data and for the data with the missing values replaced by the mean.

Table 3

*Descriptive Statistics for BSRI, CSES, and CATS Math*  \((N = 46)\)

<table>
<thead>
<tr>
<th>Variable</th>
<th>(M^a)</th>
<th>(Mdn)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BSRI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Femininity</td>
<td>48.13</td>
<td>50.00</td>
<td>9.41</td>
</tr>
<tr>
<td>Masculinity</td>
<td>49.07</td>
<td>49.00</td>
<td>9.76</td>
</tr>
<tr>
<td><strong>CSES</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Resources</td>
<td>21.00 (20.90)</td>
<td>21.00</td>
<td>4.17</td>
</tr>
<tr>
<td>Academic Achievement</td>
<td>48.82 (46.91)</td>
<td>51.00</td>
<td>9.27</td>
</tr>
<tr>
<td>Self-Regulated Learning</td>
<td>48.78 (50.06)</td>
<td>49.00</td>
<td>9.62</td>
</tr>
<tr>
<td>Leisure Time</td>
<td>40.47 (39.24)</td>
<td>41.00</td>
<td>7.52</td>
</tr>
<tr>
<td>Self-Regulatory</td>
<td>45.73 (45.51)</td>
<td>49.00</td>
<td>9.37</td>
</tr>
<tr>
<td>Others’ Expectations</td>
<td>21.98 (21.26)</td>
<td>22.00</td>
<td>4.19</td>
</tr>
<tr>
<td>Social Self-Efficacy</td>
<td>24.47 (23.19)</td>
<td>25.00</td>
<td>3.84</td>
</tr>
<tr>
<td>Self-Assertive Efficacy</td>
<td>23.31 (21.74)</td>
<td>25.00</td>
<td>4.37</td>
</tr>
<tr>
<td>Parental Support</td>
<td>18.76 (20.55)</td>
<td>19.00</td>
<td>5.39</td>
</tr>
<tr>
<td><strong>CATS Math Scale Score</strong></td>
<td>548.03</td>
<td>551.00</td>
<td>23.63</td>
</tr>
<tr>
<td><strong>CATS Math with Replaced Values</strong></td>
<td>548.03</td>
<td>548.03</td>
<td>21.43</td>
</tr>
</tbody>
</table>

\(^a\)Values in parentheses are means obtained from the Choi, Fuqua, and Griffin (2001) study.
The Femininity and Masculinity subscales of the BSRI had similar means to each other (Femininity, $M = 48.13$, Masculinity, $M = 49.07$), but Masculinity had a slightly larger standard deviation (Femininity, $SD = 9.41$; Masculinity, $SD = 9.76$). The difference between the two subscales was not found to be statistically significant, $t(88) = -.47, p > .05$. Bem (1981) reported means of 53.80 for Femininity and 48.30 for Masculinity. The sample for Bem’s work was college students; that may account for the difference in means since the sample here was high school students. The mean in this study showed a lower value for Femininity ($M = 48.13$) but a slightly higher mean for Masculinity ($M = 49.07$).

There was a noticeable difference among the means of the nine CSES subscales. Four subscales, Academic Achievement, Self-Regulated Learning, Leisure Time and Extracurricular Activities, and Self-Regulatory Efficacy have means higher than 40 while the remaining five subscales, Enlisting Social Resources, Meets Others’ Expectations, Social Self-Efficacy, Self-Assertive Efficacy, and Parental and Community Support have means less than 25. Mean results for each of the nine subscales in the current study, however, are quite similar to means obtained by Choi, Fuqua, and Griffin (2001) for the same items in each subscale, reported in parentheses in Table 3.

Descriptive statistics for math self-efficacy ratings are reported by treatment group in Table 4. All three of the treatment groups showed increases in the ratings over time with minimum means occurring on math self-efficacy rating #1. The control group means for the math self-efficacy ratings increased overall, from math self-efficacy rating #1 ($M = 68.92$) to math self-efficacy rating #5 ($M = 120.25$), with a small decrease at math self-efficacy rating #6 ($M = 116.33$). There was not much change between math
Table 4

*Descriptive Statistics for Math Self-Efficacy by Treatment Group*

<table>
<thead>
<tr>
<th>Feedback Treatment Group:</th>
<th>Control ($N = 12$)</th>
<th>Social ($N = 12$)</th>
<th>Self ($N = 11$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math Self-Efficacy #1</td>
<td>68.92 (18.40)</td>
<td>80.00 (18.83)</td>
<td>77.82 (19.40)</td>
</tr>
<tr>
<td>Math Self-Efficacy #2</td>
<td>105.67 (29.13)</td>
<td>119.25 (19.59)</td>
<td>121.82 (29.42)</td>
</tr>
<tr>
<td>Math Self-Efficacy #3</td>
<td>105.50 (18.41)</td>
<td>110.67 (18.55)</td>
<td>113.45 (28.68)</td>
</tr>
<tr>
<td>Math Self-Efficacy #4</td>
<td>112.33 (18.91)</td>
<td>110.33 (18.35)</td>
<td>117.09 (32.84)</td>
</tr>
<tr>
<td>Math Self-Efficacy #5</td>
<td>120.25 (15.58)</td>
<td>125.58 (15.25)</td>
<td>123.73 (27.25)</td>
</tr>
<tr>
<td>Math Self-Efficacy #6</td>
<td>116.33 (19.08)</td>
<td>111.00 (21.52)</td>
<td>127.18 (32.54)</td>
</tr>
</tbody>
</table>

*M (SD)*

self-efficacy rating #2 ($M = 105.67$) and math self-efficacy rating #3 ($M = 105.50$). The social-referenced feedback group was more inconsistent than the control group, showing a large increase between math self-efficacy rating #1 ($M = 80.00$) and math self-efficacy rating #2 ($M = 119.25$), but a decrease between math self-efficacy ratings #2 ($M = 119.25$) and #3 ($M = 110.67$). For the social-referenced feedback group, there was not much change between math self-efficacy ratings #3 ($M = 110.67$) and #4 ($M = 110.33$) and also a decrease at math self-efficacy rating #6 ($M = 111.00$). The decrease between math self-efficacy ratings #5 ($M = 125.58$) and #6 ($M = 111.00$) was larger for this group than the decrease for the control group between #5 ($M = 120.25$) and #6 ($M = 116.33$). These results are illustrated in Figure 3, indicating the overall trend of each of the groups.

In contrast to the control and social-referenced feedback groups, there was no decrease between math self-efficacy ratings #5 ($M = 123.73$) and #6 ($M = 127.18$) for the
self-referenced feedback group. As with the other two groups, the self-referenced feedback group showed the largest increase between math self-efficacy ratings #1 ($M = 77.82$) and #2 ($M = 121.82$), however, this increase was larger than the increase shown by the other two groups for the same two ratings. Similar to the social-referenced feedback

![Estimated Marginal Means of Math Self-Efficacy Ratings](image)

**Figure 3.** Math self-efficacy means by treatment group.

group, the self-referenced feedback group also exhibited a decrease between math self-efficacy ratings #2 ($M = 121.82$) and #3 ($M = 113.45$) of the same magnitude as the decrease for the social-referenced group for the same two ratings. The overall increases for each of the groups was similar, but the social-referenced feedback group showed the
least increase of the three treatment groups while the self-referenced feedback group showed the most increase of the three groups. Graph 3 shows that two math self-efficacy ratings, #4 and #6, for the social-referenced group are lower than both of the other groups. One explanation for the decreases for the third and fourth math self-efficacy ratings may be the level of difficulty of the course being taught during that time period. Students may have reflected their self-efficacy for the course in these ratings rather than the specific questions presented in the rating instrument.

Comparable results were obtained for math achievement. Descriptive statistics are reported in Table 5 for the Diagnostic Exams, the measure of math achievement for this study. Means for the Diagnostic Exams for the control group were inconsistent in that they fluctuated and gained two percent overall from 38.33% for Diagnostic Exam #1.

<table>
<thead>
<tr>
<th>Feedback Treatment Group:</th>
<th>Control (N = 12)</th>
<th>Social (N = 12)</th>
<th>Self (N = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic Exam #1</td>
<td>38.33 (9.62)</td>
<td>29.17 (9.50)</td>
<td>28.64 (14.33)</td>
</tr>
<tr>
<td>Diagnostic Exam #2</td>
<td>33.50 (8.06)</td>
<td>36.50 (8.33)</td>
<td>40.18 (14.18)</td>
</tr>
<tr>
<td>Diagnostic Exam #3</td>
<td>29.00 (10.99)</td>
<td>35.17 (14.70)</td>
<td>36.45 (13.40)</td>
</tr>
<tr>
<td>Diagnostic Exam #4</td>
<td>36.92 (17.32)</td>
<td>32.75 (9.56)</td>
<td>35.91 (17.11)</td>
</tr>
<tr>
<td>Diagnostic Exam #5</td>
<td>35.42 (9.16)</td>
<td>41.67 (6.16)</td>
<td>45.00 (12.04)</td>
</tr>
<tr>
<td>Diagnostic Exam #6</td>
<td>40.42 (16.02)</td>
<td>36.25 (10.47)</td>
<td>45.91 (17.15)</td>
</tr>
</tbody>
</table>

$M (SD)$
to 40.41% for Diagnostic Exam #6, dipping down to 29.00% for Diagnostic Exam #3. The social-referenced feedback group had more gain overall than the control group, starting with the lowest mean of 29.17% for Diagnostic Exam #1 and reaching 41.67% for Diagnostic Exam #5. While the mean percentages did not consistently increase each time for the social-referenced feedback group, the lowest mean occurred on Diagnostic Exam #1. The self-referenced group also had their lowest mean on Diagnostic Exam #1 but their highest mean on Diagnostic Exam #6. This self-referenced feedback group showed the most increase from the lowest to the highest means and was more consistent in the overall pattern as illustrated in the graph in Figure 4. The graph indicated that the self-referenced feedback and the social-referenced feedback groups followed trends that were somewhat alike until Diagnostic Exam #6. One explanation for this difference may have been the timing of the winter break for the school and the continuation of the course for three weeks into January. Diagnostic Exam #6 was given during the last week of the course, two days prior to the final exam. It is important to consider that these different results may have been influenced by the feedback treatment.

Grade self-efficacy ratings were also collected with each math self-efficacy. The participants rated their “certainty” in achieving each of the grades from F to A+ for the Algebra One course (Appendix R). Ratings of high certainty in achieving each of the grade levels, F through A+, would have a grade self-efficacy rating of 60 while ratings of no certainty in achieving each of the grade levels, F through A+, would have a grade self-efficacy of 10. The mean of the certainty for achieving each grade for each of the treatment groups was calculated. For example, the mean certainty rating for achieving the lowest grade F was found for each of the treatment groups. Means shown
by treatment group in Table 6 indicated that participants had some certainty of achieving each of the grade levels for each of the grade self-efficacy ratings. The mean grade self-efficacy ratings by treatment group at each of the six measurement points illustrated in Figure 5 indicated great differences between the three treatment groups. For example, the control group and the social-referenced feedback group both showed a decline in grade self-efficacy between ratings #5 and #6 while the self-referenced feedback group showed an increase.

*Figure 4.* Diagnostic exam means by treatment group.


<table>
<thead>
<tr>
<th>Feedback Treatment Group</th>
<th>Control (N = 12)</th>
<th>Social (N = 12)</th>
<th>Self (N = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Self-Efficacy #1</td>
<td>41.73 (15.28)</td>
<td>41.00 (8.81)</td>
<td>47.55 (11.60)</td>
</tr>
<tr>
<td>Grade Self-Efficacy #2</td>
<td>42.64 (15.37)</td>
<td>45.00 (9.35)</td>
<td>48.18 (9.98)</td>
</tr>
<tr>
<td>Grade Self-Efficacy #3</td>
<td>44.09 (9.03)</td>
<td>40.09 (6.91)</td>
<td>47.18 (8.85)</td>
</tr>
<tr>
<td>Grade Self-Efficacy #4</td>
<td>45.91 (11.35)</td>
<td>38.55 (12.16)</td>
<td>48.55 (10.72)</td>
</tr>
<tr>
<td>Grade Self-Efficacy #5</td>
<td>46.73 (8.14)</td>
<td>42.55 (8.38)</td>
<td>44.91 (10.54)</td>
</tr>
<tr>
<td>Grade Self-Efficacy #6</td>
<td>43.27 (8.33)</td>
<td>39.36 (11.31)</td>
<td>47.09 (12.00)</td>
</tr>
</tbody>
</table>

To assess the nature of the change in the grade self-efficacy ratings, the mean values for each of the rating items were calculated and then plotted along a curve for each of the six measurement times and the results for the three groups are presented in Figure 6. Comparisons of the three treatment groups at each measurement point was possible since the grade self-efficacy rating items appeared in rank order on the instrument, from F to A+. All three groups were similar in their certainty of achieving higher grades on grade self-efficacy rating #1. With grade self-efficacy rating #2, the self-referenced feedback group began to show higher certainty in achieving higher grades, as evidenced by the shallower slope of the line on the right side of the graph in Figure 5. This trend for the self-referenced feedback group continued but with grade self-efficacy rating #3, the social-referenced feedback group began to show less certainty to achieve higher grades not only as compared to the self-referenced feedback group, but also less certainty than
Figure 5. Grade self-efficacy rating means by treatment group.

the control group. This pattern is more evident in the graph for grade self-efficacy rating #4 where the social-referenced feedback group still showed less certainty of achieving the higher grades than the control group and both of these groups showed much less certainty than the self-referenced feedback group to achieve higher grades for the course. The self-referenced feedback group showed a little more certainty to achieve higher grades on grade self-efficacy rating #4 with an even shallower line on the right side of the graph.
For grade self-efficacy ratings #5 and #6, both the social-referenced feedback and control groups exhibited more certainty to achieve higher grades, but the social-referenced

![Graphs showing mean grade self-efficacy ratings](image)

**Figure 6.** Mean grade self-efficacy ratings by treatment group.
feedback group was still lower than the control group and the self-referenced feedback group maintained its higher level of certainty to achieve higher grades. The last of the grade self-efficacy ratings showed that the social-referenced feedback group was only slightly less certain than the control group on achieving higher grades for the course while the certainty of the self-referenced group to achieve at higher grades stayed at a high level. The mean grade self-efficacy ratings correspond to the results for the Diagnostic Exams. Students increased their certainty to achieve higher grades overall as the course progressed and the Diagnostic Exams showed similar characteristics for the self-referenced feedback group and lower mean Diagnostic Exam scores for the social-referenced feedback group on two of the Diagnostic Exams. This finding is important to note as it indicates that social-referenced feedback may have a more negative impact on math achievement than no feedback at all.

**Intercorrelations of the Variables**

The intercorrelations among the major variables appear in Table 7. There was no significant correlation between Femininity and Masculinity, which was to be expected since the instruments measure different traits. Significant correlations existed between the Femininity subscale and each of the nine subscales of the CSES, ranging from $r = .31, p < .05$, for the Self-Regulatory Efficacy subscale, to $r = .61, p < .01$, for the Enlisting Social Resources subscale. These significant positive correlations indicated that as the Femininity score increased, so did the CSES subscale score. The larger correlations for Femininity than Masculinity with the CSES subscales were unexpected and contradict earlier findings by Choi (2004). Masculinity, however, was significantly correlated with expected subscales, which reflect competitive and aggressive traits, such as academic
### Table 7

**Intercorrelations of the Variables**

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
<th>11.</th>
<th>12.</th>
<th>13.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Femininity</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Masculinity</td>
<td>.24</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Social Resources</td>
<td>.61**</td>
<td>.21</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Acad Achievement</td>
<td>.44*</td>
<td>.37*</td>
<td>.39**</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Self-Reg Learning</td>
<td>.44**</td>
<td>.00</td>
<td>.59**</td>
<td>.45**</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Leisure Time</td>
<td>.44**</td>
<td>.41**</td>
<td>.50**</td>
<td>.46**</td>
<td>.39**</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Self-Reg Efficacy</td>
<td>.31*</td>
<td>-.18</td>
<td>.24</td>
<td>.14</td>
<td>.45**</td>
<td>-.01</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Others’ Expectations</td>
<td>.39**</td>
<td>.10</td>
<td>.40**</td>
<td>.32*</td>
<td>.67**</td>
<td>.22</td>
<td>.58**</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Social Self-Efficacy</td>
<td>.54*</td>
<td>.48**</td>
<td>.44**</td>
<td>.50**</td>
<td>.43**</td>
<td>.48**</td>
<td>.34*</td>
<td>.52**</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Self-Assertive Eff</td>
<td>.48**</td>
<td>.60**</td>
<td>.43**</td>
<td>.55**</td>
<td>.38*</td>
<td>.55**</td>
<td>.14</td>
<td>.42**</td>
<td>.65**</td>
<td>_____</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Parental Support</td>
<td>.42**</td>
<td>.16</td>
<td>.48**</td>
<td>.48**</td>
<td>.71**</td>
<td>.45**</td>
<td>.19</td>
<td>.56**</td>
<td>.38**</td>
<td>.51**</td>
<td>_____</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Mean MathSE</td>
<td>.20</td>
<td>.06</td>
<td>.24</td>
<td>.41**</td>
<td>.15</td>
<td>.29</td>
<td>.12</td>
<td>.13</td>
<td>.29</td>
<td>.11</td>
<td>.02</td>
<td>_____</td>
<td></td>
</tr>
<tr>
<td>13. Mean Diagnostic</td>
<td>.09</td>
<td>-.11</td>
<td>-.03</td>
<td>.11</td>
<td>.06</td>
<td>.09</td>
<td>.27</td>
<td>.05</td>
<td>-.03</td>
<td>-.10</td>
<td>-.07</td>
<td>.20</td>
<td>_____</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01
achievement subscale, leisure time, and self-assertive. It is important to note that the participants in this previous study were college students and that the current study involved high school freshmen.

There were no significant correlations between Femininity and the mean math self-efficacy rating or Femininity and the mean Diagnostic Exam score. Aggregating information can mask significance, so the intercorrelations between Femininity and each of the six math self-efficacy ratings and Femininity and each of the six Diagnostic Exams were examined (Table 8). No significant correlations were found between Femininity and any of the individual math self-efficacy ratings or Femininity and the individual Diagnostic Exams.

Masculinity produced different results from Femininity with the CSES subscales with significant correlations for the Academic Achievement \( (r = .37, p < .05) \), Leisure Time and Extracurricular Activities \( (r = .41, p < .01) \), Social Self-Efficacy \( (r = .48, p < .01) \), and Self-Assertive Efficacy \( (r = .60, p < .01) \) subscales and nonsignificant correlations for the remaining five subscales, shown in Table 7. These significant positive correlations indicated that as the Masculinity score increased, so did the CSES subscale score. Similar to the results for Femininity, however, Masculinity did not have a significant correlation with either the mean math self-efficacy rating or the mean Diagnostic Exam. As with Femininity, examination of the intercorrelations between Masculinity and each of the six individual math self-efficacy ratings and six Diagnostic Exams showed no significant correlations (Table 8).

As shown in Table 7, the nine subscales of the CSES showed significant correlations with each other, except for the Self-Regulatory Efficacy subscale. These
### Table 8

**Intercorrelations of Femininity and Masculinity with Math Self-Efficacy Ratings and Diagnostic Exams**

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
<th>11.</th>
<th>12.</th>
<th>13.</th>
<th>14.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Femininity</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Masculinity</td>
<td>.24</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. MathSE#1</td>
<td>.23</td>
<td>.06</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. MathSE#2</td>
<td>.11</td>
<td>.10</td>
<td>.78**</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. MathSE#3</td>
<td>.17</td>
<td>.08</td>
<td>.76**</td>
<td>.88**</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. MathSE#4</td>
<td>.01</td>
<td>.04</td>
<td>.66**</td>
<td>.69**</td>
<td>.79**</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. MathSE#5</td>
<td>.20</td>
<td>-.06</td>
<td>.65**</td>
<td>.64**</td>
<td>.20**</td>
<td>.79**</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. MathSE#6</td>
<td>.25</td>
<td>.24</td>
<td>.58**</td>
<td>.59**</td>
<td>.72**</td>
<td>.72**</td>
<td>.66**</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Diagnostic#1</td>
<td>-.01</td>
<td>-.08</td>
<td>.14</td>
<td>.02</td>
<td>-.05</td>
<td>.11</td>
<td>.04</td>
<td>.00</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Diagnostic#2</td>
<td>-.14</td>
<td>-.20</td>
<td>.08</td>
<td>.21</td>
<td>.04</td>
<td>.04</td>
<td>.18</td>
<td>.07</td>
<td>.03</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Diagnostic#3</td>
<td>.08</td>
<td>-.08</td>
<td>.09</td>
<td>.04</td>
<td>-.01</td>
<td>.07</td>
<td>.22</td>
<td>.05</td>
<td>.05</td>
<td>.39**</td>
<td>____</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Diagnostic#4</td>
<td>.30</td>
<td>.14</td>
<td>.12</td>
<td>.04</td>
<td>.03</td>
<td>.11</td>
<td>.27</td>
<td>.12</td>
<td>.34*</td>
<td>.14</td>
<td>.54**</td>
<td>____</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Diagnostic#5</td>
<td>.13</td>
<td>-.15</td>
<td>.54**</td>
<td>.41**</td>
<td>.41**</td>
<td>.47**</td>
<td>.52**</td>
<td>.45**</td>
<td>.13</td>
<td>.14</td>
<td>.32*</td>
<td>.51**</td>
<td>____</td>
<td></td>
</tr>
<tr>
<td>14. Diagnostic#6</td>
<td>-.01</td>
<td>-.04</td>
<td>.27</td>
<td>.21</td>
<td>.09</td>
<td>.14</td>
<td>.22</td>
<td>.12</td>
<td>.32*</td>
<td>.41**</td>
<td>.49**</td>
<td>.53**</td>
<td>.46**</td>
<td>____</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01
significant correlations ranged from $r = .32, p < .05$, to $r = .71, p < .01$. The Self-Regulatory Efficacy subscale showed significant correlations with only three of the other subscales, Self-Regulated Learning ($r = .45, p < .01$), Meets Others’ Expectations ($r = .58, p < .01$), and Social Self-Efficacy ($r = .34, p < .05$). This means that as one of the CSES subscales increases, so does the other CSES subscale. None of the significant correlations were negative, which would have indicated that as one of the CSES subscales decreased, the other CSES subscale increased.

All of the individual math self-efficacy ratings showed significant correlations with each other, ranging from $r = .58, p < .01$, to $r = .88, p < .01$. These significant positive correlations suggested fairly strong linear relationships among the six math self-efficacy ratings. The individual Diagnostic Exams showed significant correlations among the latter four; Diagnostic Exams #3, #4, #5, and #6 were all significantly correlated with one another, ranging from $r = .32, p < .05$, to $r = .54, p < .01$. Diagnostic Exam #6 was also significantly correlated with Diagnostic Exam #1 with $r = .32, p < .05$, and Diagnostic Exam #2 with $r = .41, p < .01$. Diagnostic Exam #1 was only significantly correlated with one other Diagnostic Exam, #4 with $r = .34, p < .05$. Diagnostic Exam #2 was also only significantly correlated with one other Diagnostic Exam, #3 with $r = .39, p < .01$. The significant correlations among the latter Diagnostic Exams suggest that students were scoring more in the same way on those exams than they did on the exams near the beginning of the course. It is to be expected that as the students are exposed to more content, their scores on the Diagnostic Exams would become more consistent with each other.
Only one of the Diagnostic Exams was significantly correlated with any of the math self-efficacy ratings, #5. It exhibited significant positive correlations with each of the six math self-efficacy ratings, ranging from $r = .41, p < .01$, to $r = .54, p < .01$. There is a positive, somewhat linear relationship between each of the math self-efficacy ratings and Diagnostic Exam #5, meaning that as the scores on the math self-efficacy ratings increased, so did the scores on the Diagnostic Exam #5. Diagnostic Exam #5 was administered shortly before the midyear winter break from school. The algebra one course was not completed until three weeks after the winter break, and the last math self-efficacy rating and Diagnostic Exam #6 were administered two days before the end of the course.

Analyses for the Research Questions

Results for each of the research questions for this study are examined in this section. The research questions were:

1. How does each student’s math self-efficacy rating change over time?
2. Do changes in students’ math self-efficacy ratings vary by feedback treatment?
3. How does each student’s grade self-efficacy rating change over time?
4. Do changes in students’ grade self-efficacy ratings vary by feedback treatment?
5. How does each student’s math achievement (measured by the Diagnostic Exam) change over time?
6. Do changes in students’ math achievement (measured by the Diagnostic Exam) vary by feedback treatment?
Previous math achievement and the Masculinity subscale scores of the BSRI were checked for differences among the three treatment groups. Based on a review of the literature, these two variables were positively correlated with math self-efficacy. Analysis was conducted for possible inclusion of these variables as covariates in the statistical investigation of the effect of the feedback treatment on both math self-efficacy ratings and math achievement, as measured by the Diagnostic Exams. One way Analysis of Variance (ANOVA) was conducted on the mean Masculinity subscale scores for the three treatment groups of self-referenced feedback, social-referenced feedback, and no feedback (control). The means were not significantly different among the three treatment groups, F(2, 43) = 1.24, p > .05. ANOVA was also conducted on the measure of previous math achievement, comparing the mean CATS Math scale scores for the three treatment groups of self-referenced feedback, social-referenced feedback, and no feedback (control). The means were not significantly different among the three treatment groups, F(2, 35) = .66, p > .05. The analysis was repeated using the mean score for the missing values and there was not a significant difference among the mean CATS Math scale scores of the three treatment groups, F(2, 43) = .69, p > .05.

Due to the nonsignificance in the two variables across the three groups and also due to the small sample size, these variables were not included in subsequent analyses. This is important because if the treatment groups were inherently different, it would not be possible to distinguish the influence of the treatment on the dependent variable. Group differences could confound the effect of the treatment making it difficult to assess any true impact on the outcomes of the study. No group differences were found so the effect of the feedback treatment on math self-efficacy, grade self-efficacy, and math
achievement was analyzed by exploration of the change over time of the participants within the treatment groups. Due to the longitudinal nature of the measurements within the participants and participants nested within treatment groups, a multilevel modeling approach was used. Multilevel modeling allows for analysis of not only individual variations over time but also between individual variations over time. In other words, one level of the multilevel modeling approach examines the within-person change and another level examines the between-person differences in change. For example, overall, individuals may have increased their math self-efficacy ratings or their scores on the Diagnostic Exams but additional information is necessary concerning the differences in how the individuals within each feedback treatment group changed.

Longitudinal data, like that collected for this study, can be analyzed using a growth curve modeling approach in HLM. A description of Hierarchical Linear Modeling (HLM) can be found in Appendix T. Two-level hierarchical linear modeling uses two questions, one at each level of the model. The level-1 question is about within-person change and the level-2 question concerns between-person differences in change. Each pair of the research questions in this study formed the two levels necessary when building a hierarchical model. Beginning with the first outcome variable of math self-efficacy, research question one, “How does each student’s math self-efficacy rating change over time?” is a level-1 question because it is about within-person change. Research question two, “Do changes in students’ math self-efficacy ratings vary by feedback treatment?” is a level-2 question for math self-efficacy since it concerns between-person differences in change. The equations for each model are described in detail, with explanations of each subscript. Subsequent explanations for the remaining research questions are implied.
Prior to building a hierarchical model, it is advisable to look at the individual data as line graphs to observe an overall trend (Singer & Willett, 2003). Figure 7 contains the line graphs of the math self-efficacy ratings for each student, with interpolated missing values and observed data. Overall trends appeared to increase from the initial measurement to the last measurement, with some decreases intermediated. As previously noted, the last math self-efficacy rating was administered after the midyear winter break, shortly before the course ended. This may explain some of the decrease noted between the last two measurements on some of the students. Individual regression

Figure 7. Line graphs for individual students for math self-efficacy ratings.
was conducted on the data for each student and is in Table 9. Time was measured in
approximately four-week intervals, with time equal to zero for the first math self-efficacy
ingrating administered on the fifth day of the course. Time equal to one represents the
second administration of math self-efficacy, approximately four weeks after the first.
The estimates of the initial status, the first math self-efficacy rating, ranged from 35.95 to
131.52. This difference in individuals is very large, indicating that the initial math self-
efficacy rating for the regression of math self-efficacy rating on time was very low for
some students and very high for others. The slope of the change in math self-efficacy
rating for the regression of math self-efficacy rating on time was also widely varied
among the students, ranging from -1.49 to 18.49. This indicated that for some students,
the average change in math self-efficacy decreased approximately 1.49 points for each
subsequent math self-efficacy rating while students at the other end of the spectrum
gained about 18.49 points in math self-efficacy for each subsequent math self-efficacy
rating. A graph of the regression equations is presented in Figure 8 and the variation in
intercepts and slopes are easily seen.

The R-squared values also varied greatly, ranging from 0.00 to 0.97. This
indicated the proportion of variation in math self-efficacy ratings that was explained by
the least squares regression of math self-efficacy on time varied from 0% to 97%. For
some students, there was very little explained variation while for others, almost all of the
variation in math self-efficacy ratings was explained by the least squares regression of
math self-efficacy on time. The data for math self-efficacy indicate a great deal of
variation within the students. Hierarchical Linear Modeling (HLM) is an appropriate
statistical tool to use for this study because of the variability nested within the students.
Table 9.

**Within-Person Ordinary Least Squares Regression Models for Math Self-Efficacy.**

<table>
<thead>
<tr>
<th>ID</th>
<th>Initial Status</th>
<th>Rate of Change</th>
<th>Residual Variance</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>14519287</td>
<td>123.35, 17.62</td>
<td>1.60, 5.81</td>
<td>580.17</td>
<td>0.03</td>
</tr>
<tr>
<td>21106350</td>
<td>80.10, 15.18</td>
<td>7.83, 5.02</td>
<td>440.20</td>
<td>0.38</td>
</tr>
<tr>
<td>21266749</td>
<td>97.81, 16.03</td>
<td>5.34, 5.29</td>
<td>490.32</td>
<td>0.20</td>
</tr>
<tr>
<td>22354166</td>
<td>69.95, 7.90</td>
<td>6.89, 2.61</td>
<td>119.28</td>
<td>0.64</td>
</tr>
<tr>
<td>23104156</td>
<td>59.34, 11.31</td>
<td>8.69, 3.37</td>
<td>167.81</td>
<td>0.77</td>
</tr>
<tr>
<td>23439138</td>
<td>129.29, 17.53</td>
<td>8.89, 5.79</td>
<td>586.44</td>
<td>0.37</td>
</tr>
<tr>
<td>24162729</td>
<td>105.76, 12.36</td>
<td>4.23, 4.08</td>
<td>291.60</td>
<td>0.21</td>
</tr>
<tr>
<td>26091435</td>
<td>86.00, 6.71</td>
<td>9.00, 5.20</td>
<td>54.00</td>
<td>0.75</td>
</tr>
<tr>
<td>26111533</td>
<td>102.05, 13.21</td>
<td>5.91, 4.36</td>
<td>333.18</td>
<td>0.32</td>
</tr>
<tr>
<td>29095097</td>
<td>94.38, 11.93</td>
<td>8.51, 3.94</td>
<td>271.68</td>
<td>0.54</td>
</tr>
<tr>
<td>29339060</td>
<td>110.05, 14.47</td>
<td>13.71, 4.78</td>
<td>399.48</td>
<td>0.67</td>
</tr>
<tr>
<td>34808964</td>
<td>115.00, 15.06</td>
<td>2.20, 4.98</td>
<td>433.20</td>
<td>0.05</td>
</tr>
<tr>
<td>35063844</td>
<td>83.81, 8.76</td>
<td>4.54, 2.89</td>
<td>146.42</td>
<td>0.38</td>
</tr>
<tr>
<td>36505675</td>
<td>35.95, 7.98</td>
<td>18.49, 2.64</td>
<td>121.68</td>
<td>0.93</td>
</tr>
<tr>
<td>40693715</td>
<td>124.04, 18.95</td>
<td>-0.14, 6.92</td>
<td>670.23</td>
<td>0.90</td>
</tr>
<tr>
<td>42936584</td>
<td>106.60, 16.22</td>
<td>6.40, 6.62</td>
<td>438.53</td>
<td>0.24</td>
</tr>
<tr>
<td>46184649</td>
<td>120.86, 11.74</td>
<td>1.86, 3.88</td>
<td>263.29</td>
<td>0.05</td>
</tr>
<tr>
<td>50857896</td>
<td>102.14, 9.93</td>
<td>6.14, 3.28</td>
<td>188.29</td>
<td>0.47</td>
</tr>
<tr>
<td>54141060</td>
<td>131.52, 10.92</td>
<td>10.06, 3.61</td>
<td>227.82</td>
<td>0.66</td>
</tr>
<tr>
<td>54436320</td>
<td>104.00, 15.67</td>
<td>6.00, 5.18</td>
<td>469.00</td>
<td>0.25</td>
</tr>
<tr>
<td>55835284</td>
<td>93.48, 9.97</td>
<td>9.34, 3.29</td>
<td>189.82</td>
<td>0.67</td>
</tr>
<tr>
<td>57004719</td>
<td>85.48, 8.09</td>
<td>5.34, 2.67</td>
<td>124.82</td>
<td>0.50</td>
</tr>
<tr>
<td>58498416</td>
<td>80.14, 15.59</td>
<td>6.74, 5.15</td>
<td>464.09</td>
<td>0.30</td>
</tr>
<tr>
<td>62037318</td>
<td>105.60, 20.16</td>
<td>9.90, 8.23</td>
<td>677.70</td>
<td>0.33</td>
</tr>
<tr>
<td>66004612</td>
<td>89.29, 13.16</td>
<td>4.89, 4.35</td>
<td>330.44</td>
<td>0.24</td>
</tr>
<tr>
<td>67209813</td>
<td>63.29, 14.02</td>
<td>9.29, 4.63</td>
<td>375.14</td>
<td>0.50</td>
</tr>
<tr>
<td>69824876</td>
<td>68.90, 19.31</td>
<td>6.77, 6.38</td>
<td>711.60</td>
<td>0.22</td>
</tr>
<tr>
<td>70191572</td>
<td>94.80, 13.07</td>
<td>2.40, 5.33</td>
<td>284.53</td>
<td>0.06</td>
</tr>
<tr>
<td>71368685</td>
<td>60.57, 7.91</td>
<td>6.57, 2.61</td>
<td>119.57</td>
<td>0.61</td>
</tr>
<tr>
<td>71622430</td>
<td>92.42, 16.39</td>
<td>8.33, 5.40</td>
<td>502.33</td>
<td>0.44</td>
</tr>
<tr>
<td>73528075</td>
<td>71.19, 12.19</td>
<td>6.86, 4.03</td>
<td>283.62</td>
<td>0.42</td>
</tr>
<tr>
<td>73540933</td>
<td>115.05, 12.35</td>
<td>-1.49, 4.08</td>
<td>291.18</td>
<td>0.03</td>
</tr>
<tr>
<td>75184998</td>
<td>83.10, 13.79</td>
<td>15.29, 4.66</td>
<td>320.76</td>
<td>0.84</td>
</tr>
<tr>
<td>77129176</td>
<td>105.00, 13.68</td>
<td>4.40, 4.52</td>
<td>357.30</td>
<td>0.19</td>
</tr>
<tr>
<td>77623950</td>
<td>70.14, 9.98</td>
<td>3.14, 3.30</td>
<td>190.29</td>
<td>0.19</td>
</tr>
<tr>
<td>78235405</td>
<td>87.33, 2.35</td>
<td>9.20, 0.78</td>
<td>10.53</td>
<td>0.97</td>
</tr>
<tr>
<td>81831287</td>
<td>104.81, 14.88</td>
<td>4.14, 4.91</td>
<td>422.62</td>
<td>0.15</td>
</tr>
<tr>
<td>83341631</td>
<td>114.43, 16.47</td>
<td>4.23, 5.44</td>
<td>517.77</td>
<td>0.13</td>
</tr>
<tr>
<td>84272178</td>
<td>77.81, 11.98</td>
<td>8.34, 3.96</td>
<td>273.82</td>
<td>0.53</td>
</tr>
<tr>
<td>84508420</td>
<td>102.62, 5.50</td>
<td>7.49, 1.82</td>
<td>57.68</td>
<td>0.81</td>
</tr>
<tr>
<td>85393271</td>
<td>97.48, 6.92</td>
<td>9.94, 2.28</td>
<td>91.32</td>
<td>0.83</td>
</tr>
<tr>
<td>86470057</td>
<td>87.90, 12.77</td>
<td>2.77, 4.22</td>
<td>311.10</td>
<td>0.10</td>
</tr>
<tr>
<td>92685750</td>
<td>93.50, 21.24</td>
<td>17.50, 16.45</td>
<td>541.50</td>
<td>0.53</td>
</tr>
<tr>
<td>96505784</td>
<td>109.14, 9.20</td>
<td>7.34, 3.04</td>
<td>161.49</td>
<td>0.59</td>
</tr>
<tr>
<td>97504261</td>
<td>113.24, 11.30</td>
<td>10.27, 3.44</td>
<td>174.97</td>
<td>0.75</td>
</tr>
<tr>
<td>98643146</td>
<td>56.90, 10.57</td>
<td>12.77, 3.49</td>
<td>213.10</td>
<td>0.77</td>
</tr>
</tbody>
</table>
Regression Models for Math Self-Efficacy

Figure 8. Regression models of math self-efficacy.

HLM has the flexibility to allow for both variability within persons as well as variability within persons due to group membership, allowing for variability for between persons. A beginning Hierarchical Linear model is one called an unconditional growth model, where the level-1 equation represents this within-person change over time.

The level-1 equation of a hierarchical linear model includes time as the only predictor. The model for the level-1 equation is

\[ Y_{ti} = \pi_{0i} + \pi_{1i} \text{Time}_{ti} + e_{ti} \]

Within-person effects are signified by the symbol \( \pi \). The \( Y_{ti} \) represents the math self-efficacy rating for individual \( i \) measured at time \( t \). \( \text{Time}_{ti} \) represents the variable measurement in terms of four-week intervals for this study, with 0 as the first time point. The intercept, \( \pi_{0i} \), is expected math self-efficacy rating of the individual at \( \text{Time}_{ti} = 0 \), or the initial math self-efficacy status. The slope of the equation is defined as \( \pi_{1i} \), which
represents the rate of change in individual \( i \). The last term in the equation, \( e_{ii} \), signifies the residuals for within-person predictions, in other words, the difference between observed and predicted values. This is often referred to as measurement error. The equation can be rewritten to illustrate these definitions as math self-efficacy for individual \( i \) = initial math self-efficacy for individual \( i \) + growth rate for individual \( i \times \) (time interval) + residuals. This model is often referred to as the unconditional growth model because no independent variables are used to explain the initial status or the rate of change other than time.

The level-2 equations for the hierarchical linear model for math self-efficacy allow for investigation of variations in the initial math self-efficacy and the rates of change for math self-efficacy between persons. Two equations are required, one for the initial math self-efficacy in the level-1 equation, \( \pi_{0i} \), and one for the rate of change in math self-efficacy, \( \pi_{1i} \). These between-person models are

\[
\pi_{0i} = \gamma_{00} + r_{0i} \\
\pi_{1i} = \gamma_{10} + r_{1i}.
\]

Between-person effects are designated by \( \gamma \). The initial math self-efficacy for individual \( i \), \( \pi_{0i} \), is the population average initial math self-efficacy, \( \gamma_{00} \). The coefficient \( \gamma_{01} \) represents the population average change in the initial math self-efficacy. These equations can be better understood if rewritten as

Initial math self-efficacy for individual \( i \) = population average initial math self-efficacy + random effect; and

Rate of change in math self-efficacy for individual \( i \) = population average rate of change in math self-efficacy + random effect.
The two random effects associated with the between-persons equations are $r_{0i}$ and $r_{1i}$.

The squares of these random effects yield variances $\tau_{00}$ and $\tau_{11}$, respectively, and a covariance $\tau_{0i}$. These variances and covariances can be used to determine the correlation between the initial status and the rate of change for linear growth curves.

Results of HLM analysis for the unconditional growth model are presented in Table 10 and can be used to answer research question one, “How does each student’s math self-efficacy score change over time?” The answer is that math self-efficacy scores Table 10.

HLM Results of Unconditional and Conditional Growth Models for Math Self-Efficacy.

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Growth Model</th>
<th>Conditional Growth Model (treatment group as predictor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>For Intercept ($\pi_{0i}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($\gamma_{00}$)</td>
<td>$93.90*** (p = 0.00)$</td>
<td>$90.43*** (p = 0.00)$</td>
</tr>
<tr>
<td>Treatment ($\gamma_{01}$)</td>
<td>______</td>
<td>$3.41 (p = .38)$</td>
</tr>
<tr>
<td>Variance ($\tau_{00}$)</td>
<td>$299.64*** (p = 0.00)$</td>
<td>$298.67*** (p = 0.00)$</td>
</tr>
<tr>
<td>For Slope ($\pi_{1i}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope ($\gamma_{10}$)</td>
<td>$6.86*** (p = 0.00)$</td>
<td>$6.82*** (p = 0.00)$</td>
</tr>
<tr>
<td>Treatment ($\gamma_{11}$)</td>
<td>______</td>
<td>$0.04 (p = 0.96)$</td>
</tr>
<tr>
<td>Variance ($\tau_{11}$)</td>
<td>$0.09 (p &gt; 0.50)$</td>
<td>$0.10 (p &gt; 0.50)$</td>
</tr>
<tr>
<td>Residual Variance $\sigma^2$</td>
<td>$297.94$</td>
<td>$299.37$</td>
</tr>
</tbody>
</table>

***$p < 0.001$

increase in a linear fashion, with significant intercept $(93.90, p = 0.00)$ and rate of change $(6.86, p = 0.00)$ for the within-person model. This can be interpreted as math self-
efficacy = 93.90 + 6.86(Time), where time is 0, 1, 2, 3, 4, or 5; measured in four-week intervals. For example, when time is equal to 2, math self-efficacy = 93.90 + 6.86(2) = 107.62. This is the predicted math self-efficacy for the second assessment given after the initial administration, or eight weeks after the initial assessment.

Further analysis of the results of the unconditional growth model includes examination of the variance components. Variance components describe the amount of outcome variability left after fitting the multilevel model (Singer & Willett, 2003). Analysis of the variance components distinguishes between the level-1 variation and the level-2 variation. The level-1 variance summarizes the average spread of an individual’s observed math self-efficacy scores around his or her own true rate of change. The unconditional growth model level-1 residual variance for math self-efficacy is 297.94. This number is difficult to evaluate since there is no comparison value but it can be compared to the variance components of other models for the data; a procedure that was conducted and explained after the next model.

Level-2 variance allows the distinction of level-1 variance into that portion due to within-person differences in the true initial math self-efficacy or the true rate of change in math self-efficacy. For the unconditional growth model, the variance for level-2 variation included 299.64 (p = 0.00) for the intercept and .09 (p > 0.50) for the rate of change. The variance due to the initial math self-efficacy, which is the intercept, is significant while the variance due to the rate of change is not. This indicates that predictors may exist that would explain the variance associated with the true initial math self-efficacy.
In addition to the variance components for levels one and two, there was a covariance associated with level-2 of the unconditional growth model. This evaluates the covariance between the true initial math self-efficacy and the true rate of change. The covariance for the unconditional growth model math self-efficacy was 1.79. Singer and Willett (2003) suggest that the covariance is more easily understood if it is transformed into a correlation coefficient. This is accomplished by dividing the covariance by the square root of the product of the variances for the level-2 equations. For this model,

\[
\frac{1.79}{\sqrt{(299.64)(0.09)}} = 0.34
\]

The correlation between true initial math self-efficacy and the true rate of change in math self-efficacy is positive but weak. The addition of predictor variables for the intercept and rate of change may help explain the variation in scores.

The second model is a conditional growth model because a predictor variable for the intercept and rate of change in the level-1 equation is introduced for the level-2 equations. The level-1 equation for the conditional growth model still includes time as the only predictor so that equation remains the same as the previous unconditional growth model:

Math self-efficacy for individual \( i \) = initial math self-efficacy for individual \( i \) + growth rate for individual \( i \) \* (time interval) + residuals.

The conditional growth model differs, however, from the unconditional growth model because an independent variable is introduced to explain the initial status or the rate of change beyond the effects of time. The level-2 equations for this HLM allow for investigation of variations in the initial math self-efficacy and the rates of change for math self-efficacy as a function of the feedback treatment group. This part of multilevel
modeling makes it possible to explore between-person differences in change in math self-efficacy due to an independent variable. For example, the rate of change in math self-efficacy may be different for members of one of the treatment groups than for members of another treatment group.

Two equations are required for level two, one for the initial math self-efficacy in the level-1 equation, \( \pi_{0i} \), and one for the rate of change in math self-efficacy, \( \pi_{1i} \). These between-person models are

\[
\pi_{0i} = \gamma_{00} + \gamma_{01} \text{(Treatment)} + r_{0i} \\
\pi_{1i} = \gamma_{10} + \gamma_{11} \text{(Treatment)} + r_{1i}.
\]

The initial math self-efficacy for individual \( i \), \( \pi_{0i} \), is the initial math self-efficacy for an individual in the control feedback treatment, \( \gamma_{00} \), where the control treatment is equal to 0. The coefficient \( \gamma_{01} \) represents how much change in the initial math self-efficacy is expected for the social-referenced feedback, treatment equal to 1, and self-referenced feedback, treatment equal to 2. The rate of change in math self-efficacy for individual \( i \), \( \pi_{1i} \), is equal to the rate of change for an individual in the control feedback treatment, signified by \( \gamma_{10} \). The coefficient \( \gamma_{11} \) represents how much increase or decrease is expected in the rate of change for the social-referenced feedback, treatment equal to 1, and self-referenced feedback, treatment equal to 2. The two random effects associated with the between-persons equations are the same as described for the unconditional growth model. These equations can be better understood if rewritten as

Initial math self-efficacy for individual \( i \) = expected initial math self-efficacy for the control group + a rate of change * (treatment group where 0 = control, 1 = social, 2 = self) + random effect; and
Rate of change in math self-efficacy for individual $i$ = expected rate of change in math self-efficacy for the control group $+$ a rate of change * (treatment group where 0 = control, 1 = social, 2 = self) $+$ random effect.

Results of HLM analysis for the conditional growth model are presented in Table 10 and were used to answer the second research question, “Do changes in students’ math self-efficacy ratings vary by feedback treatment?” The answer is no; the results of this study do not indicate that there was a difference in math self-efficacy, either in the initial math self-efficacy or the rate of change in math self-efficacy, with respect to the feedback treatment. Treatment was not a significant predictor for either the intercept or rate of change. This means that the true initial math self-efficacy was not affected by the treatment group nor was the rate of change in math self-efficacy affected by the treatment group. The conditional growth model can be examined further, however, with respect to the variance components and compared to the first model for verification.

Variance components for the second model are similar values to those obtained for the unconditional growth model; for the intercept, the true initial math self-efficacy, 298.67 ($p = 0.00$), and for the rate of change, 0.10 ($p > 0.50$). The covariance is slightly higher with the second model, 1.81 and calculation of the correlation coefficient in the same manner as for the unconditional growth model suggests a similar correlation.

$$\frac{1.81}{\sqrt{(298.67)(0.10)}} = 0.33$$

This correlation is expected since the feedback treatment was not a significant predictor, meaning no more variance was explained in this model than in the first model. There is still a positive weak relationship between true initial math self-efficacy and rate of change in math self-efficacy regardless of the feedback treatment. Although several
methods exist for comparing multilevel models, there is not a clear agreement in the literature at this time (Singer & Willett, 2003). One method is to compare the within-person residual variance to determine if any additional variation is explained by the additional variables used at level two. Since the additional variables were not significant for the second model, the comparison of the residual variances between the two models should be close 0%. The residual variance for the unconditional growth model was 299.94 and for the conditional growth model, it was 299.37. The decline is very minimal and calculation of the additional variation explained is negligible and in fact, less because the residual variance for the second model is greater than the residual variance for the first model. \( \frac{(297.94 - 299.37)}{297.94} = -0.0048 \). This verifies that the conditional growth model did not fit the data better than the unconditional growth model and the feedback treatment did not affect either the true initial math self-efficacy or the rate of change in math self-efficacy.

The same statistical analyses were conducted both for the grade self-efficacy ratings as well as the Diagnostic Exams to answer the remaining research questions. Both the unconditional growth model and conditional growth model using the feedback treatment group as a predictor were analyzed for each of the remaining two dependent variables. The level-1 and level-2 equations used were the same as those for the math self-efficacy ratings except for the outcome variable, which changed to grade self-efficacy rating and Diagnostic Exam in place of the math self-efficacy ratings. Preliminary graphs and regression models were also created, and the individual graphs for grade self-efficacy ratings are presented in Figure 9. These graphs indicated more
curvature in some of the graphs than what was seen in the individual graphs for the math

![Graphs showing self-efficacy ratings](image)

**Figure 9.** Line graphs for individual students for grade self-efficacy ratings.

self-efficacy, however, there appear to be more linear graphs in the grade self-efficacy than those seen in the individual graphs for math self-efficacy.

The regression models for grade self-efficacy ratings were also determined and are illustrated in Table 11. Variation in the initial status for the grade self-efficacy was very large, much like that seen with the math self-efficacy ratings. The possible scores on the grade self-efficacy rating ranged from 10 to 60 and the regression estimates based on the observed data for the intercept, or initial grade self-efficacy, ranged from 3.95 to
Table 11.

*Within-Person Ordinary Least Squares Regression Models for Grade Self-Efficacy.*

<table>
<thead>
<tr>
<th>ID</th>
<th>Initial Status</th>
<th>Rate of Change</th>
<th>Residual Variance</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>14519287</td>
<td>55.40</td>
<td>4.51</td>
<td>-4.58</td>
<td>1.49</td>
</tr>
<tr>
<td>21106350</td>
<td>55.90</td>
<td>2.68</td>
<td>-1.83</td>
<td>0.88</td>
</tr>
<tr>
<td>21266749</td>
<td>22.62</td>
<td>12.08</td>
<td>-4.16</td>
<td>0.45</td>
</tr>
<tr>
<td>22354166</td>
<td>37.81</td>
<td>2.89</td>
<td>-1.26</td>
<td>0.95</td>
</tr>
<tr>
<td>23104156</td>
<td>30.64</td>
<td>2.62</td>
<td>0.22</td>
<td>0.78</td>
</tr>
<tr>
<td>23439138</td>
<td>58.14</td>
<td>7.06</td>
<td>-1.26</td>
<td>2.33</td>
</tr>
<tr>
<td>23462792</td>
<td>49.43</td>
<td>3.51</td>
<td>-1.57</td>
<td>1.16</td>
</tr>
<tr>
<td>26111533</td>
<td>56.52</td>
<td>2.28</td>
<td>-2.03</td>
<td>1.06</td>
</tr>
<tr>
<td>29095097</td>
<td>31.95</td>
<td>12.01</td>
<td>-2.34</td>
<td>3.97</td>
</tr>
<tr>
<td>30879843</td>
<td>37.81</td>
<td>1.00</td>
<td>0.54</td>
<td>0.33</td>
</tr>
<tr>
<td>35063844</td>
<td>3.95</td>
<td>5.23</td>
<td>7.49</td>
<td>1.73</td>
</tr>
<tr>
<td>36056752</td>
<td>54.93</td>
<td>4.51</td>
<td>-5.71</td>
<td>1.65</td>
</tr>
<tr>
<td>42936584</td>
<td>54.80</td>
<td>3.57</td>
<td>-0.10</td>
<td>1.16</td>
</tr>
<tr>
<td>46184649</td>
<td>41.76</td>
<td>1.17</td>
<td>-0.37</td>
<td>0.39</td>
</tr>
<tr>
<td>50857896</td>
<td>39.95</td>
<td>7.69</td>
<td>4.89</td>
<td>2.54</td>
</tr>
<tr>
<td>54104060</td>
<td>55.57</td>
<td>1.09</td>
<td>0.77</td>
<td>0.36</td>
</tr>
<tr>
<td>54436320</td>
<td>51.70</td>
<td>1.39</td>
<td>-0.17</td>
<td>0.46</td>
</tr>
<tr>
<td>55835284</td>
<td>56.71</td>
<td>1.42</td>
<td>-0.09</td>
<td>0.47</td>
</tr>
<tr>
<td>57004719</td>
<td>47.86</td>
<td>4.58</td>
<td>-1.54</td>
<td>1.51</td>
</tr>
<tr>
<td>58498416</td>
<td>47.05</td>
<td>4.55</td>
<td>-3.29</td>
<td>1.50</td>
</tr>
<tr>
<td>62037318</td>
<td>56.40</td>
<td>1.06</td>
<td>0.20</td>
<td>0.43</td>
</tr>
<tr>
<td>66004612</td>
<td>29.90</td>
<td>3.81</td>
<td>1.17</td>
<td>1.26</td>
</tr>
<tr>
<td>67209813</td>
<td>47.48</td>
<td>2.36</td>
<td>1.14</td>
<td>0.78</td>
</tr>
<tr>
<td>69824876</td>
<td>31.95</td>
<td>12.01</td>
<td>-2.11</td>
<td>3.97</td>
</tr>
<tr>
<td>70191572</td>
<td>55.80</td>
<td>7.83</td>
<td>-6.40</td>
<td>3.20</td>
</tr>
<tr>
<td>71368685</td>
<td>33.24</td>
<td>3.21</td>
<td>-0.03</td>
<td>1.06</td>
</tr>
<tr>
<td>71622430</td>
<td>41.42</td>
<td>1.96</td>
<td>1.83</td>
<td>0.65</td>
</tr>
<tr>
<td>73528075</td>
<td>33.10</td>
<td>3.54</td>
<td>1.43</td>
<td>1.17</td>
</tr>
<tr>
<td>73540933</td>
<td>36.05</td>
<td>2.01</td>
<td>-2.49</td>
<td>0.66</td>
</tr>
<tr>
<td>75184998</td>
<td>49.54</td>
<td>4.47</td>
<td>1.20</td>
<td>1.51</td>
</tr>
<tr>
<td>77129176</td>
<td>54.67</td>
<td>2.61</td>
<td>-3.20</td>
<td>0.86</td>
</tr>
<tr>
<td>77623950</td>
<td>30.52</td>
<td>5.14</td>
<td>-0.54</td>
<td>1.70</td>
</tr>
<tr>
<td>78235405</td>
<td>34.71</td>
<td>5.74</td>
<td>4.31</td>
<td>1.90</td>
</tr>
<tr>
<td>81831287</td>
<td>53.14</td>
<td>1.79</td>
<td>-0.26</td>
<td>0.59</td>
</tr>
<tr>
<td>83341631</td>
<td>48.95</td>
<td>2.59</td>
<td>-2.31</td>
<td>0.85</td>
</tr>
<tr>
<td>84727217</td>
<td>42.29</td>
<td>2.39</td>
<td>-0.31</td>
<td>0.79</td>
</tr>
<tr>
<td>84508420</td>
<td>40.38</td>
<td>1.74</td>
<td>0.11</td>
<td>0.57</td>
</tr>
<tr>
<td>85393271</td>
<td>47.81</td>
<td>3.77</td>
<td>2.14</td>
<td>1.24</td>
</tr>
<tr>
<td>86470057</td>
<td>48.81</td>
<td>2.04</td>
<td>-4.26</td>
<td>0.67</td>
</tr>
<tr>
<td>92685750</td>
<td>50.83</td>
<td>4.10</td>
<td>-1.50</td>
<td>3.18</td>
</tr>
<tr>
<td>96505784</td>
<td>47.48</td>
<td>2.16</td>
<td>-0.46</td>
<td>0.71</td>
</tr>
<tr>
<td>97504261</td>
<td>34.05</td>
<td>8.14</td>
<td>3.34</td>
<td>2.48</td>
</tr>
<tr>
<td>98643146</td>
<td>30.33</td>
<td>2.62</td>
<td>2.40</td>
<td>0.87</td>
</tr>
</tbody>
</table>
The slope, or the rate of change in the grade self-efficacy over time, showed great diversity as well, with values ranging from -6.40 to 7.49. These values indicate that the rate of change in the grade self-efficacy over time varied from a decrease of about 6.40 points every four weeks to an increase of about 7.49 points every four weeks. A regression models graph containing all 46 of the regression models is presented in Figure 10. The overall trends in grade self-efficacy can be seen along with the intercepts and slopes. There are some regression models that increase and some that decrease. If students were more certain of achieving higher grades as the course progressed, the slope of the regression line would be positive and increasing. For students whose certainty of achieving higher grades decreased as the course progressed, the slope of the regression line would be negative and decreasing. \( R^2 \) values ranged from 0.00 to 0.91, which indicated that none of the variation between grade self-efficacy and time was explained.

Figure 10. Regression models of grade self-efficacy.
by the least squares regression of grade self-efficacy on time for some of the students. For other students, approximately 91% of the variation in grade self-efficacy was explained by the least squares regression of grade self-efficacy on time. There is a great deal of variability within these regression models and an overall trend that fits the entire class does not appear possible. Using HLM, however, may lead to a model, given its flexibility for fitting multilevel data.

Results of the HLM modeling for grade self-efficacy are presented in Table 12 for both the unconditional growth model and the conditional growth model. The unconditional growth model contained one predictor variable of time in the level-1 equation. Findings indicated that the answer to research question three, “How does each student’s grade self-efficacy rating change over time?” is that it remains fairly constant. The intercept for the within-person model was significant (44.31, \( p = 0.00 \)), and the slope, or rate of change, showed a slight decrease over time (-0.24, \( p = .548 \)) even though it was not significant. This means that grade self-efficacy = 44.31 – 0.24(Time), measured as with the math self-efficacy, 0 through 5, in four-week intervals. The slope has only a minor effect on the grade self-efficacy over time. An example of this slight effect eight weeks after the initial administration of the grade self-efficacy is grade self-efficacy = 44.31 - .24(2) = 43.83.

Additional analysis was conducted on the variance components for the unconditional growth model, also presented in Table 12. The unconditional growth model level-1 residual variance for grade self-efficacy is 37.27. As with the variance for math self-efficacy, this number is difficult to evaluate since there is no comparison value but it can be compared to the variance components of other models.
Table 12.

*HLM Results of Unconditional and Conditional Growth Models for Grade Self-Efficacy.*

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Growth Model</th>
<th>Conditional Growth Model (treatment group as predictor)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For Intercept</strong> ((\pi_{00}))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ((\gamma_{00}))</td>
<td>44.31***((p = 0.00))</td>
<td>42.51***((p = 0.00))</td>
</tr>
<tr>
<td>Treatment ((\gamma_{01}))</td>
<td></td>
<td>1.77((p = 0.40))</td>
</tr>
<tr>
<td>Variance ((\tau_{00}))</td>
<td>109.94***((p = 0.00))</td>
<td>110.59***((p = 0.00))</td>
</tr>
<tr>
<td><strong>For Slope</strong> ((\pi_{10}))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope ((\gamma_{10}))</td>
<td>-0.24((p = 0.55))</td>
<td>-0.06((p = 0.93))</td>
</tr>
<tr>
<td>Treatment ((\gamma_{11}))</td>
<td></td>
<td>-0.18((p = 0.71))</td>
</tr>
<tr>
<td>Variance ((\tau_{11}))</td>
<td>4.56***((p = 0.00))</td>
<td>4.70***((p = 0.00))</td>
</tr>
<tr>
<td>Residual Variance (\sigma^2)</td>
<td>37.27</td>
<td>37.28</td>
</tr>
</tbody>
</table>

\*\(*\ast\ast\ast p < 0.001\)

The level-2 variance for the unconditional growth model includes 109.94 \((p = 0.00)\) for the intercept, the initial grade self-efficacy, and 4.56 \((p = 0.00)\) for the rate of change. Both of these variance components were significant, a difference from the math self-efficacy. This indicated that there may be additional independent variables to explain the variation in both the true initial grade self-efficacy as well as the rate of change in the grade self-efficacy. This is examined in the second model of conditional growth. The covariance of the first model is analyzed before proceeding to the next model.

The covariance between the true initial grade self-efficacy and the true rate of change is -12.28. Again following the suggestion by Singer & Willett (2003), the correlation coefficient was calculated by dividing the covariance by the square root of the
product of the level-2 variances. Thus, the correlation coefficient for the first model was

\[ \frac{-12.28}{\sqrt{(109.94)(4.56)}} = -0.55. \]

This indicated a moderately strong negative correlation between the true initial grade self-efficacy and the true rate of change, meaning that as the true initial grade self-efficacy increased, the true rate of change decreased.

The independent predictor in the conditional growth model was the feedback treatment group, as with the math self-efficacy analysis. Results of this model are in Table 10 and provided the answer to the fourth research question, “Do changes in students’ grade self-efficacy ratings vary by feedback treatment?” The answer is no, the same as with math self-efficacy. The results of this study do not indicate that there was a difference in grade self-efficacy, either in the initial math self-efficacy or the rate of change in math self-efficacy, due to the type of feedback treatment. There was no significance in either of the coefficients for the effect of the feedback treatment on grade self-efficacy. Examination of the variance components for this model verified these findings and comparison of the residual variance to that of the unconditional growth model indicated that it was not a better fit for the data.

The variance for the true initial grade self-efficacy is 110.59 \((p = 0.00)\) and 4.70 \((p = 0.00)\) for the true rate of change. Both are very similar to the variances for the unconditional growth model previously discussed. The covariance is slightly higher, -12.44. The correlation coefficient, calculated as previously described,

\[ \frac{-12.44}{\sqrt{(110.59)(4.70)}} = -0.55, \]

was the same as for the unconditional growth model. Comparison of the residual variances from the two models verified that the conditional growth model did not explain any additional variation in the grade self-efficacy,
\[
\frac{(37.27 - 37.28)}{37.27} = -0.0003\]. The comparison of the residual variances between the two models for grade self-efficacy is very much the same as those for math self-efficacy.

As previously stated, the same statistical analyses were conducted on math achievement, measured by the Diagnostic Exams; to answer the remaining two research questions. The individual graphs of the Diagnostic Exams over time appear in Figure 11. As noted in the descriptive statistics section, the last measurement, Diagnostic Exam #6, exhibited a lower mean than the previous Diagnostic Exam. This is very apparent on the

Figure 11. Line graphs for individual students for grade self-efficacy ratings.
individual graphs, along with some other low values for measurements in the third or fourth months. The sharp increase for most of the graphs after only one month is interesting to note. The same topics and questions with different values were randomly ordered and presented at each measurement time on the Diagnostic Exams. The fact that so many students did well on the second diagnostic exam is curious, given that only four weeks of curriculum had been taught in that interval and only a few of the topics had been covered by that point in time.

Regression models were created for the Diagnostic Exams and the results appear in Table 13. The Diagnostic Exams were scaled in percent, using the number of correct responses. The intercept for the Diagnostic Exam regression models did not vary as much as was seen with math or grade self-efficacy ratings. Percents ranged from about 12% to 60.5%, a much narrower spread than the other outcome variables for the intercept. The slopes, or rates of change, for the regression models for the Diagnostic Exam varied from -20 to about 9. This indicated that some students scored about 20% lower on the Diagnostic Exam over time while some students increased about 9% over time.

The $R^2$ values varied about the same as the math and grade self-efficacies, from 0.00 to 0.97. This indicated that any where from 0% to 97% of the variation in the Diagnostic Exams was explained by the least squares regression of Diagnostic Exams on time, meaning that the amount left unexplained varied from 100% to 3%. This huge disparity in variation explained by the regression models fit the within-person variability present in the study. A graph of all of the regression models is illustrated in Figure 12.
Table 13.

*Within-Person Ordinary Least Squares Regression Models for Diagnostic Exams.*

<table>
<thead>
<tr>
<th>ID</th>
<th>Initial Status Estimate</th>
<th>Initial Status SE</th>
<th>Rate of Change Estimate</th>
<th>Rate of Change SE</th>
<th>Residual Variance</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>14519287</td>
<td>28.35</td>
<td>2.75</td>
<td>2.60</td>
<td>0.91</td>
<td>14.17</td>
<td>0.73</td>
</tr>
<tr>
<td>21106350</td>
<td>24.81</td>
<td>9.53</td>
<td>1.54</td>
<td>3.15</td>
<td>173.42</td>
<td>0.06</td>
</tr>
<tr>
<td>21266749</td>
<td>39.38</td>
<td>9.28</td>
<td>1.11</td>
<td>3.06</td>
<td>164.28</td>
<td>0.03</td>
</tr>
<tr>
<td>22354166</td>
<td>43.05</td>
<td>4.50</td>
<td>2.51</td>
<td>1.49</td>
<td>38.68</td>
<td>0.42</td>
</tr>
<tr>
<td>23104156</td>
<td>27.17</td>
<td>13.56</td>
<td>3.85</td>
<td>4.04</td>
<td>241.20</td>
<td>0.31</td>
</tr>
<tr>
<td>23439138</td>
<td>22.00</td>
<td>8.78</td>
<td>0.80</td>
<td>2.90</td>
<td>147.20</td>
<td>0.02</td>
</tr>
<tr>
<td>24162729</td>
<td>24.67</td>
<td>11.73</td>
<td>2.80</td>
<td>3.87</td>
<td>262.53</td>
<td>0.12</td>
</tr>
<tr>
<td>26091435</td>
<td>47.83</td>
<td>4.84</td>
<td>-6.50</td>
<td>3.75</td>
<td>28.17</td>
<td>0.75</td>
</tr>
<tr>
<td>26111533</td>
<td>32.29</td>
<td>12.44</td>
<td>0.49</td>
<td>4.11</td>
<td>295.34</td>
<td>0.00</td>
</tr>
<tr>
<td>29095097</td>
<td>37.67</td>
<td>9.88</td>
<td>-2.00</td>
<td>3.26</td>
<td>186.33</td>
<td>0.09</td>
</tr>
<tr>
<td>29339060</td>
<td>21.43</td>
<td>5.65</td>
<td>9.03</td>
<td>1.87</td>
<td>60.87</td>
<td>0.85</td>
</tr>
<tr>
<td>34808964</td>
<td>35.81</td>
<td>4.70</td>
<td>0.14</td>
<td>1.55</td>
<td>42.12</td>
<td>0.00</td>
</tr>
<tr>
<td>35063844</td>
<td>29.38</td>
<td>7.15</td>
<td>3.71</td>
<td>2.36</td>
<td>97.48</td>
<td>0.38</td>
</tr>
<tr>
<td>36505675</td>
<td>33.10</td>
<td>5.51</td>
<td>-2.57</td>
<td>1.82</td>
<td>57.90</td>
<td>0.33</td>
</tr>
<tr>
<td>40693715</td>
<td>23.50</td>
<td>10.31</td>
<td>1.00</td>
<td>3.77</td>
<td>198.50</td>
<td>0.03</td>
</tr>
<tr>
<td>42936584</td>
<td>37.40</td>
<td>9.20</td>
<td>-1.00</td>
<td>3.76</td>
<td>141.07</td>
<td>0.02</td>
</tr>
<tr>
<td>46184649</td>
<td>29.29</td>
<td>6.40</td>
<td>0.29</td>
<td>2.11</td>
<td>78.14</td>
<td>0.00</td>
</tr>
<tr>
<td>50857896</td>
<td>21.14</td>
<td>8.30</td>
<td>3.54</td>
<td>2.74</td>
<td>131.59</td>
<td>0.29</td>
</tr>
<tr>
<td>54141060</td>
<td>60.38</td>
<td>1.14</td>
<td>1.71</td>
<td>0.38</td>
<td>2.48</td>
<td>0.84</td>
</tr>
<tr>
<td>54436320</td>
<td>40.33</td>
<td>10.37</td>
<td>1.00</td>
<td>3.43</td>
<td>205.33</td>
<td>0.02</td>
</tr>
<tr>
<td>55835284</td>
<td>47.57</td>
<td>2.60</td>
<td>-1.83</td>
<td>0.86</td>
<td>12.87</td>
<td>0.53</td>
</tr>
<tr>
<td>57004719</td>
<td>31.62</td>
<td>2.92</td>
<td>2.89</td>
<td>0.96</td>
<td>16.28</td>
<td>0.69</td>
</tr>
<tr>
<td>58498416</td>
<td>23.67</td>
<td>6.06</td>
<td>2.60</td>
<td>2.00</td>
<td>70.13</td>
<td>0.30</td>
</tr>
<tr>
<td>62037318</td>
<td>35.60</td>
<td>9.39</td>
<td>2.00</td>
<td>3.83</td>
<td>147.07</td>
<td>0.08</td>
</tr>
<tr>
<td>66004612</td>
<td>27.00</td>
<td>6.14</td>
<td>2.00</td>
<td>2.03</td>
<td>72.00</td>
<td>0.20</td>
</tr>
<tr>
<td>67209813</td>
<td>36.76</td>
<td>4.21</td>
<td>-2.57</td>
<td>1.39</td>
<td>33.90</td>
<td>0.46</td>
</tr>
<tr>
<td>69824876</td>
<td>27.00</td>
<td>7.32</td>
<td>0.80</td>
<td>2.42</td>
<td>102.20</td>
<td>0.03</td>
</tr>
<tr>
<td>70191572</td>
<td>46.20</td>
<td>1.39</td>
<td>-5.40</td>
<td>0.57</td>
<td>3.20</td>
<td>0.97</td>
</tr>
<tr>
<td>71368685</td>
<td>31.76</td>
<td>7.38</td>
<td>0.43</td>
<td>2.44</td>
<td>103.90</td>
<td>0.01</td>
</tr>
<tr>
<td>71622430</td>
<td>36.07</td>
<td>3.27</td>
<td>0.72</td>
<td>1.08</td>
<td>19.95</td>
<td>0.13</td>
</tr>
<tr>
<td>73528075</td>
<td>30.19</td>
<td>4.43</td>
<td>0.46</td>
<td>1.46</td>
<td>37.42</td>
<td>0.02</td>
</tr>
<tr>
<td>73540933</td>
<td>35.10</td>
<td>7.92</td>
<td>0.03</td>
<td>2.62</td>
<td>119.70</td>
<td>0.00</td>
</tr>
<tr>
<td>75184998</td>
<td>26.19</td>
<td>7.85</td>
<td>-1.31</td>
<td>2.65</td>
<td>103.81</td>
<td>0.11</td>
</tr>
<tr>
<td>77129176</td>
<td>36.52</td>
<td>8.28</td>
<td>-0.54</td>
<td>2.74</td>
<td>130.92</td>
<td>0.01</td>
</tr>
<tr>
<td>77623950</td>
<td>36.48</td>
<td>4.27</td>
<td>2.34</td>
<td>1.41</td>
<td>34.82</td>
<td>0.41</td>
</tr>
<tr>
<td>78235405</td>
<td>27.33</td>
<td>2.82</td>
<td>4.60</td>
<td>0.93</td>
<td>15.13</td>
<td>0.86</td>
</tr>
<tr>
<td>81831287</td>
<td>23.10</td>
<td>11.32</td>
<td>4.83</td>
<td>3.74</td>
<td>244.70</td>
<td>0.29</td>
</tr>
<tr>
<td>83341631</td>
<td>29.48</td>
<td>6.79</td>
<td>2.74</td>
<td>2.24</td>
<td>87.92</td>
<td>0.27</td>
</tr>
<tr>
<td>84272178</td>
<td>33.29</td>
<td>2.22</td>
<td>1.89</td>
<td>0.73</td>
<td>9.44</td>
<td>0.62</td>
</tr>
<tr>
<td>84508420</td>
<td>39.24</td>
<td>6.22</td>
<td>2.57</td>
<td>2.06</td>
<td>73.90</td>
<td>0.28</td>
</tr>
<tr>
<td>85393271</td>
<td>36.52</td>
<td>9.77</td>
<td>7.86</td>
<td>3.23</td>
<td>182.12</td>
<td>0.60</td>
</tr>
<tr>
<td>86470057</td>
<td>31.19</td>
<td>9.81</td>
<td>1.86</td>
<td>3.24</td>
<td>183.62</td>
<td>0.08</td>
</tr>
<tr>
<td>92685750</td>
<td>50.67</td>
<td>12.67</td>
<td>-20.00</td>
<td>9.81</td>
<td>192.67</td>
<td>0.81</td>
</tr>
<tr>
<td>96505784</td>
<td>32.71</td>
<td>7.50</td>
<td>-0.69</td>
<td>2.48</td>
<td>107.44</td>
<td>0.02</td>
</tr>
<tr>
<td>97504261</td>
<td>12.03</td>
<td>7.09</td>
<td>5.92</td>
<td>2.16</td>
<td>68.90</td>
<td>0.71</td>
</tr>
<tr>
<td>98643146</td>
<td>23.43</td>
<td>4.79</td>
<td>0.03</td>
<td>1.58</td>
<td>43.87</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Figure 12. Regression models of diagnostic exam.

The majority of the slopes of the regression models are almost flat, indicating not much increase or decrease for a rate of change. A few unusual models, however, indicate positive, increasing lines while a dramatically negative line is most noticeable on the left of the graph. This participant was only in the study for three of the measurement points and scored very minimally on the last Diagnostic Exam administered, causing the huge negative slope.

Results of the HLM modeling for Diagnostic Exams are presented in Table 14 for both the unconditional growth model and the conditional growth model. The unconditional growth model contained one predictor variable of time. Findings indicated that the answer to research question five, “How does each student’s math achievement (measured by the Diagnostic Exam) change over time?” is that there is slight linear
growth over time. The intercept for the within-person model was significant (32.35, \( p = .00 \)) and the slope, or rate of change, was also significant (1.43, \( p = 0.00 \)). This means that the true initial Diagnostic Exam showed a slight increase over time. This means that the predicted Diagnostic Exam = 32.35 + 1.43(Time), where time is measured from 0 to 5, indicating the number of four-week intervals since the initial administration of the Diagnostic Exam. For instance, the predicted Diagnostic Exam for the third four-week interval is equal to 32.35 + 1.43(3) or 36.64. These results fit the pattern noted in the individual graphs of a slightly positive slope.

Table 14.

*HLM Results of Unconditional and Conditional Growth Models for Diagnostic Exam.*

<table>
<thead>
<tr>
<th></th>
<th>Unconditional Growth Model</th>
<th>Conditional Growth Model (treatment group as predictor)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For Intercept (( \pi_{0i} ))</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (( \gamma_{00} ))</td>
<td>32.35***((p = 0.00))</td>
<td>34.10***((p = 0.00))</td>
</tr>
<tr>
<td>Treatment (( \gamma_{01} ))</td>
<td>-</td>
<td>-1.74((p = 0.27))</td>
</tr>
<tr>
<td>Variance (( \tau_{00} ))</td>
<td>17.18((p = 0.07))</td>
<td>20.06((p = 0.06))</td>
</tr>
<tr>
<td><strong>For Slope (( \pi_{1i} ))</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope (( \gamma_{10} ))</td>
<td>1.43***((p = 0.00))</td>
<td>0.45((p = 0.48))</td>
</tr>
<tr>
<td>Treatment (( \gamma_{11} ))</td>
<td>-</td>
<td>0.98**((p = 0.05))</td>
</tr>
<tr>
<td>Variance (( \tau_{11} ))</td>
<td>1.28((p = 0.12))</td>
<td>1.05((p = 0.19))</td>
</tr>
<tr>
<td>Residual Variance ( \sigma^2 )</td>
<td>103.74</td>
<td>102.61</td>
</tr>
</tbody>
</table>

**\( **p < 0.05, ***p < 0.001**

Examination of the variance components for the unconditional growth model revealed no significant components for either the intercept (17.18, \( p = 0.07 \)) or the slope (1.28, \( p = 0.12 \)). This means that there is not likely to be significant predictors not
included in the model or, in other words, that the model explains most of the variance. These results are not the same as for the math and grade self-efficacies. The covariance for this model between the true initial Diagnostic Exam score and the true rate of change is 4.51. This is larger than the covariance for the math self-efficacy. The covariance for the grade self-efficacy was larger in magnitude (-12.28) but negative. The correlation coefficient was calculated for this model as well, $\frac{4.51}{\sqrt{(17.18)(1.28)}} = 0.96$. This is a much larger correlation coefficient than was seen for the math or grade self-efficacy models. This means that there is a strong positive correlation between the true initial Diagnostic Exam score and the true rate of change. As the initial score increases, the rate of change also increases. Squaring the correlation coefficient produced an $R^2$ of 0.92, which can be interpreted as 92% of the variation in the rate of change in Diagnostic Exam was explained by the rate of change in the Diagnostic Exam. This large amount of variation explained leaves a small amount unexplained by the model.

The second model is a conditional growth model with the independent predictor of feedback treatment group, the same as with the math and grade self-efficacies analyses. The results of this model answered the last research question, “Do changes in students’ math achievement (measured by the Diagnostic Exam) vary by feedback treatment?” The results of this study indicated that the answer is yes, there was a difference in Diagnostic Exam by feedback group (Table 14). The rate of change for the rate of change variable in the level-2 equation for Time was significant (.98, $p = 0.05$). This means that there was a greater rate of change for students in the social-referenced and self-referenced feedback groups. Considering all of the coefficients, the equation looks like
Predicted Diagnostic Exam = 34.10 - 1.75(Treatment) +[.45 + 0.98(Treatment)](Time),
where Treatment = 0 for the control group, 1 for the social-referenced feedback, and 2 for the self-referenced feedback. In other words, students in the self-referenced feedback group showed more rate of change than a student in the control or social-referenced feedback group. The equation for a student in the self-referenced group would look like this:

Predicted Diagnostic Exam = 34.10 - 1.75(2) +[0.45 + 0.98(2)](Time),
while the equation for a student in the control group would look like this:

Predicted Diagnostic Exam = 34.10 - 1.75(0) +[0.45 + 0.98(0)](Time).

The student in the self-referenced feedback treatment group has a rate of change of 2.41 (0.45 + 0.98(2) = 2.41) and the student in the control group has a rate of change of only 0.45 (0.45 + 0.98(0) = .45). Similar results are indicated for the social-referenced group but with a smaller rate of change than the self-referenced group.

Predicted Diagnostic Exam = 34.10 - 1.75(1) +[.45 + 0.98(1)](Time).

This student has rate of change of 1.43 (0.45 + 0.98(1) = 1.43) as compared to .45 for the control group and 2.41 for the self-referenced group. Even though these rates are small, the fact remains that the self-referenced group showed a rate that was more than 1.5 times that of the social-referenced group and more than five times that of the control group.

Further analysis on the variance components for the intercept (20.06, p = 0.06) and slope (1.05, p = 0.19) of the conditional growth model revealed that neither were significant. The covariance between the true initial Diagnostic Exam and the true rate of change was 4.47, slightly smaller than the covariance for the unconditional growth
model. Calculating the correlation coefficient as \[\frac{4.47}{\sqrt{(20.06)(1.05)}} = .97\], a much stronger positive correlation than what was exhibited by either the math or grade self-efficacy. This was a very strong positive correlation between the initial Diagnostic Exam score and the true rate of change, with the treatment group as an independent variable affecting the true rate of change. As the initial Diagnostic Exam score increased, the true rate of change also increased but increased at a faster rate for the self-referenced feedback group and the social-referenced feedback group. Squaring the correlation coefficient produced an \(R^2\) value of 0.94, meaning that about 94% of the variation in the rate of change with respect to the treatment group was explained by the least squares regression of the rate of change on the initial Diagnostic Exam scores, leaving only 6% unexplained by the model.

The residual variance for both models also appears in Table 14 and can be used to compare them, \[\frac{(103.74 - 102.61)}{103.74} = 0.0109\]. The difference in the residual variances does not indicate that the conditional model explains much more variation than the unconditional growth model, however, it is important to note that the component variance for the rate of change was slightly smaller in the conditional growth model, meaning that less variance remained unexplained by the second model.

The analysis for the research questions showed that there was growth in math self-efficacy, grade self-efficacy, and math achievement (as measured by the Diagnostic Exam). For the effects of the feedback treatment, however, only the Diagnostic Exam exhibited any difference among the three groups. The self-referenced feedback group experienced a rate more than five times that of the control group.
Summary of Findings

The research questions and results for this study were:

1. How does each student’s math self-efficacy rating change over time? The
result was that student math self-efficacy ratings increased in a linear pattern over time.

2. Do changes in students’ math self-efficacy ratings vary by feedback treatment?
There was not a significant difference in math self-efficacy ratings among the three

treatment groups.

3. How does each student’s grade self-efficacy rating change over time? The
result for the grade self-efficacy pattern was the same as for math self-efficacy, with the
total scores increasing in a linear pattern over time.

4. Do changes in students’ grade self-efficacy ratings vary by feedback
treatment? As with the math self-efficacy, there was also not a significant difference in
grade self-efficacy ratings among the three treatment groups.

5. How does each student’s math achievement (measured by the Diagnostic
Exam) change over time? The result was the same as for math and grade self-efficacies,
with the Diagnostic Exam scores increasing in a linear pattern over time.

6. Do changes in students’ math achievement (measured by the Diagnostic
Exam) vary by feedback treatment? There was a significant difference in the rate of
change for the Diagnostic Exam among the three treatment groups.

The major findings of the longitudinal analyses indicated that for all three of the
dependent variables, math self-efficacy, grade self-efficacy, and math achievement
(measured by Diagnostic Exam); there was a significant linear change over time. For the
effect of the treatment group, however, only the Diagnostic Exam showed any significant differences, and those differences related to rate of change in math achievement.

Analysis of the math self-efficacy indicated that the initial math self-efficacy rating was about 94 with an increase of approximately 7 points every four weeks. But the effect of the feedback treatment was not significant on the groups either for the initial status of math self-efficacy or the rate of change over time. Similar results were found for grade self-efficacy except the linear change was negative over time. The initial grade self-efficacy rating was around 44 with a very slight decrease over time, overall remaining constant throughout the study. There was not a significant difference between the rates of change or initial grade self-efficacies for the three feedback treatment groups, the same as seen with the math self-efficacy ratings. There were different rates of change between the three groups, but not significantly different.

The Diagnostic Exam not only produced significant results for linear change over time but also exhibited a significant difference in rates of change for the treatment groups. Even though the rates of change in Diagnostic Exam were smaller than expected, both the rate of change and initial Diagnostic Exam scores were significant, with an initial score of around 32% and an additional increase of 1.5% every four weeks. Furthermore, the most important finding was that the rates of change in Diagnostic Exam for the treatment groups were significantly different from each other. The rate of change for the self-referenced feedback group was approximately an additional 2.41% in the Diagnostic Exam score for each measurement point while the social-referenced group increased about 1.43% and the control group only increased .45%. Implications of these findings are discussed in the next chapter.
CHAPTER V

DISCUSSION

This chapter provides a brief summary of the purpose of the study and the major findings, along with implications of the findings and recommendations for future research.

This study aimed to serve three purposes. The first purpose was to extend previous research in a different setting by using high school participants. Much of prior research was conducted either in the elementary setting or with college age participants. The second purpose was to demonstrate that the use of specific types of feedback could increase math self-efficacy that would, in turn, increase math achievement. Thirdly, this study sought to examine trends in self-efficacy and its effects on math achievement over a longer time interval than what was often reported in the literature. Longitudinal research in the area of academic self-efficacy and academic achievement is minimal, especially at the high school level. Most studies measured the effect of feedback on self-efficacy or math achievement, either at one point, or at most two points, in the research process. Studies of the patterns in multiple measurements of self-efficacy or math achievement over time seem to be missing from the literature.

Results

Based on the findings from the literature, six research questions for this study were proposed and implemented. The questions and results are repeated here.
1. How does each student’s math self-efficacy rating change over time?
2. Do changes in students’ math self-efficacy ratings vary by feedback treatment?
3. How does each student’s grade self-efficacy rating change over time?
4. Do changes in students’ grade self-efficacy ratings vary by feedback treatment?
5. How does each student’s math achievement (measured by the Diagnostic Exam) change over time?
6. Do changes in students’ math achievement (measured by the Diagnostic Exam) vary by feedback treatment?

To answer these questions, a quasi-experimental study was designed, using two intact classes in one school, and randomly assigning the students to one of three feedback treatments. The three feedback treatments were self-referenced feedback, social-referenced feedback, and a control group where participants received innocuous comments. The research was conducted in a double-blind design. Further controls on the study included using the same teacher for both classes and randomly assigning the three treatments within each class. The math self-efficacy ratings, grade self-efficacy ratings, and the Diagnostic Exams were not reviewed until the conclusion of the study. This ensured no change on the part of the teacher in delivery of the instruction or feedback to the participants. The participants were unaware of the type of feedback being received; they were informed that they were participating in a study of teaching strategies. The written feedback appeared on the 5-minute checks completed in class in journals that were kept in the room. The participants completed all 5-minute checks anonymously and no grades were issued for completion of any of the instruments used in the study.
Results for Math Self-Efficacy

Results of the study indicated that the answer to research question one, “How does each student’s math self-efficacy rating change over time?”, is it grows linearly over time. The expected initial math self-efficacy was about 94 with an increase of approximately 7 points every four weeks. But the effect of the feedback treatment was not significant in its effect on the groups either for the initial status of math self-efficacy or the rate of change over time, indicating that the answer to research question two, “Do changes in students’ math self-efficacy ratings vary by feedback treatment?” is no, they do not vary by feedback treatment. Multilevel modeling for the longitudinal data revealed no differences in the rates of change over time for math self-efficacy ratings among the three treatment groups. In other words, the effects of feedback on math and grade self-efficacy were comparable across the six different measurement points for each of the three conditions of feedback.

One purpose of this study was to examine the effect of type of feedback on math self-efficacy. Math self-efficacy ratings did increase but there was no statistical significance in the rate of change over time among the three treatment groups. In fact, the math self-efficacy ratings fluctuated, decreasing between the second and third ratings and again between the fifth and final ratings. Although the results were nonsignificant, the mean math self-efficacy ratings of the social-referenced feedback group were less than the mean math self-efficacy ratings of the self-referenced feedback group, parallel to previous findings (Shih & Alexander, 2000).

One speculation for the lack of significant findings is possibly due to the participants selected for this study. Students with lower level math achievement were
chosen for this study. It is possible that if the same study were conducted with a group of students with higher levels of math achievement, such as in advanced classes, the mean math self-efficacy ratings of the social-referenced feedback may have been higher, in response to peer pressure and the competitive setting. In addition, the social-referenced feedback seemed to have a less sustained effect on math self-efficacy than the control treatment. In other words, the final math self-efficacy rating of the control group was higher than that of the social-referenced group. This also may not be true for other levels of students. Students of average ability may respond to social-referenced feedback more than what was exhibited in this study with lower level students.

Another explanation for the lack of significant results in the math self-efficacy across feedback treatment may be the fact that self-efficacy is difficult to increase. It takes many positive outcomes for a person to perceive a change in his or her capabilities to perform a certain task. Even though the study was five months long, it is possible that there was not enough of the type of feedback necessary to affect a significant change in math self-efficacy. Negative outcomes quickly lower self-efficacy but positive outcomes are very slow at raising self-efficacy.

Results for Grade Self-Efficacy.

Similar results were obtained for grade self-efficacy. The answer to third research question, “How does each student’s grade self-efficacy rating change over time?” also was a linear pattern but of a slight decrease rather than an increase. The possible grade self-efficacy ratings ranged from 10 to 60. The initial grade self-efficacy was around 44 with a slight decrease over time. Overall, grade self-efficacy remained fairly constant throughout the study. As with the math self-efficacy, there was not a significant
difference between the rates of change or initial grade self-efficacies for the three feedback treatment groups. This indicates that the answer to the fourth research question, “Do changes in students’ grade self-efficacy ratings vary by feedback treatment?” is also no. Even though not significant, however, the rates of change were different among the three feedback treatment groups.

The three feedback treatment groups also fluctuated in the mean grade self-efficacy ratings, even more erratically than what was seen in the mean math self-efficacy ratings. There was also a decrease between the second and third grade self-efficacy ratings for the self- and social-referenced feedback groups just as with the math self-efficacy ratings. The control group did not show a decrease in grade self-efficacy ratings until between the fifth and last ratings. The social-referenced feedback group also exhibited a decrease between the fifth and last ratings but the self-referenced feedback group showed an increase between these two ratings. Previous findings regarding grade self-efficacies have been more consistent than what was found in this study (Zimmerman, & Bandura, 1994). Zimmerman and Bandura (1994) found that most students in their study had a high certainty of earning at least a B, and the patterns of grade self-efficacy ratings were reported to be the same for all students.

An explanation for the decrease in grade self-efficacies may be the timing of grade reports at the school. Midterm reports for the first half of the course were issued prior to the third grade self-efficacy and it is possible that students’ perceptions were altered when presented with actual grades. Grade reports for the first half of the class were issued prior to the fourth grade self-efficacy rating and, overall, the grades for each class were perhaps lower than what the students expected to see. This was the first
semester in high school for the participants and it is possible that adjusting to the more
rigorous coursework and demands of high school schedules affected grades more than
students were expecting.

Results for Diagnostic Exam

Math achievement (measured by the Diagnostic Exam), the third dependent
variable, produced different results from the first two dependent variables. The fifth
research question “How does each student’s math achievement (measured by the
Diagnostic Exam) change over time?” showed significant results for both the initial
Diagnostic Exam as well as the rate of change over time. The initial Diagnostic Exam
grade was around 32% with an additional 1.5% increase each four weeks. While this was
not a very large initial score, it makes sense because the participants in this study were
beginning in the lowest level math course offered at this school. The 1.5% rate of change
was far less than expected. However, perhaps the fact that the Diagnostic Exam
contained questions covering all of the course topics may explain the low rate of change.

The effects of feedback on math achievement, measured by the Diagnostic Exam,
were not comparable across the six different measurement points for each of the three
conditions. The rate of change in the Diagnostic Exam score for the self-referenced
feedback treatment group was significantly different from the rates of change for the
other two treatment groups. This significant difference in the Diagnostic Exam by
feedback treatment group means the answer to the last research question, “Do changes in
students’ math achievement (measured by the Diagnostic Exam) vary by feedback
treatment?” is yes.
Findings that the Diagnostic Exam scores were different among the three feedback treatment groups along with significant differences in the rates of change across time are important. This means that the self-referenced feedback treatment group saw a higher rate of change in Diagnostic Exam score than the other groups. The rate of change for the self-referenced feedback group was more than five times that of the control group and almost double that of the social-referenced feedback group. Even though the initial Diagnostic Exam score was low and the rate of change small, the fact that the self-referenced group saw such a significant difference in the rate of change from the other two feedback groups very important. These results indicated the possibility that even though the type of feedback may not significantly impact the rates of change for math or grade self-efficacy, the effect of the type of feedback on the measure of math achievement was significant. These results of the effect of type of feedback are consistent with what has been reported in the literature (Shih & Alexander, 2000).

Even though the rates of change among the three feedback treatment groups were not consistent, they were all very small. The small rates of change for the Diagnostic Exam may be explained by the use of feedback in response to the 5-minute checks. Both the 5-minute checks and the Diagnostic Exams were related to the course curriculum, however, it is possible that the feedback was not specific enough to alter the math self-efficacy regarding the Diagnostic Exams. Meaning that perhaps the students did not make a connection between the written feedback to the 5-minute checks and their performances on the Diagnostic Exams.

Additionally, the students received no compensation for participation in the study and frequently asked if a grade was going to be issued for the Diagnostic Exams. Since
no grade was issued to the students for their performance on the Diagnostic Exams, one explanation for the low initial score on the Diagnostic Exam and the small rates of change may have been due to a lack of serious effort on the part of the participants. Some participants returned the Diagnostic Exams in a matter of a few minutes while others took much longer, working through all of the problems. The multiple-choice format of the Diagnostic Exam was designed to eliminate any subjective grading but it could have, in fact, reduced the accuracy of measuring math achievement since it allowed the students to simply choose answers.

Implications

Practical significance from this study for classroom teachers included a few points. First, the measurements obtained here indicated that one point in time reflected a lot of error or variation in performance. These widely varying results within each measurement point are an indication to high school classroom teachers that great fluctuation in assessment over time is to be expected, especially with students with lower levels of math self-efficacy and mathematics skills such as the participants in this study. High school classroom teachers implement a multitude of teaching strategies in an effort to increase academic achievement. A poor performance by the students could cause a teacher to discontinue a particular strategy when, in fact, the poor performance is merely a normal fluctuation among lower level high school students. The overall increase at the end of the mathematics course is more important and this study indicated that individual measurements during the course are prone to variation.

Secondly, while the results of the study were not significant by treatment group for math and grade self-efficacies, there was an indication that the self-referenced
feedback group did have a higher math self-efficacy than the social-referenced feedback or control groups. The impact of social-referenced feedback is more reflective of the atmosphere in advanced classes, where students are usually more competitive with one another, while in average or lower level classes the effects of social-referenced feedback might actually do more harm than good as it relates to math achievement. Typical high school classroom settings for average or lower level courses are not necessarily conducive to self-referenced feedback with thirty or more students, however, the results of this study indicated that self-referenced feedback might provide an additional strategy for improving math self-efficacy, grade self-efficacy, and math achievement among these groups of students. Self-referenced feedback in a high school classroom might pose a difficult shift for established classroom teachers but would be a simple and inexpensive strategy to implement. Comparative feedback regarding past performances on assessments or skills could provide necessary focus for lower level students who lack the self-regulatory mechanism to be successful.

As high school classroom teachers struggle to find teaching strategies to raise student achievement, they have the potential to do action research, for example, with self-referenced feedback, because in most cases they teach multiple classes of the same course. It would be relatively easy for a high school teacher to use one class as a control group and implement self-referenced feedback in another class and then compare the difference in performance between the two groups. This design would also allow the teacher to address other characteristics typically exhibited by individuals with high self-efficacy, such as persistence on task. Academic self-efficacy is difficult to improve and providing more opportunities within the classroom for students to experience self-
referred feedback have the potential to effect a positive change not only on self-efficacy but academic achievement as well.

Finally, much attention is given to students considered at-risk due to, among other variables, family economic situations and lower academic skills. More often than not, a large percentage of the lower level students are also considered at-risk. The results of this study indicated that the social-referenced feedback group often exhibited either similar means or lower means than the control group. A possible explanation for these results is that the class contained a large percentage of at-risk students who are quite aware of their academic standing among their peers and therefore the social-referenced feedback had no impact. At-risk students may also have a lack of family and peer support for academic achievement. Self-referenced feedback might provide the opportunity for these students to develop a higher self-efficacy, which, in turn, has been shown to predict higher academic achievement.

Limitations

This study was limited in some aspects that bear further scrutiny. For example, while the double-blind design increased the internal validity of the study, it also reduced its generalizability due to the many controls. Using the same instructor for both classes ensured that the instruction for the course was not confounded with the feedback treatment. If different instructors were used for the two classes, any significant results may have been due to different teaching styles rather than the effect of the feedback group. Keeping the participants in one level of math, algebra one, also increased the internal validity since the participants were more alike in skills, however, the results may not generalize to other diverse groups of students.
Controlling for the instructor and course impacted the sample size, which is another limitation of the study. The small sample size may have deterred the findings due to a lack of power, since power is directly related to sample size (Keppell & Zedeck, 1989; Stevens, 1996). Small sample sizes are often necessary when implementing quasi-experimental designs, however, working within the constraints of intact classes and school systems. Specifically for this study, increasing the sample size would have necessitated involving another instructor, possibly interfering with consistency of the manipulation of the feedback treatment. Designs of quasi-experiments often include trade-offs in that while certain features of the design strengthen internal validity, generalizability may be subsequently decreased.

The lack of a well-established math achievement assessment for ninth grade algebra students may have been a limitation. The use of the state Diagnostic Exam addressing the course content of algebra one did increase the specificity of the math self-efficacy as recommended in the literature (Bong, 2002; Bong & Hocevar, 2002; Kranzler & Pajares, 1997; Pajares, 1996a; Pajares & Miller, 1994, 1995; Pajares & Urdan, 2006; Smith & Fouad, 1999), but may not have been an accurate measurement. In addition, only a random sample of the items from the Diagnostic Exam were used and this may have affected the results, even though the same questions with different values were used on each of the Diagnostic exams. Use of the Diagnostic Exam, however, led to more internal validity than using a researcher-prepared assessment.

Possibly an additional limitation could have been that the researcher was the instructor. The study was tightly controlled, however, using a double-blind design and written feedback, to name two features of the study. Furthermore, none of the data was
examined until after the conclusion of the study, not allowing the instructor to alter
instruction. With the tightly controlled design, the researcher being the teacher for the
study was probably not a factor in the results.

Recommendations

Information gained from this longitudinal study provides valuable evidence that is
limited in the current literature. More study is needed to study the impact of self-
referred feedback on math self-efficacy in an effort to raise math achievement. This
study did not show significant differences in the rates of change in math self-efficacy due
to type of feedback even though there were significant differences in the rates of change
in the measure of math achievement between feedback groups. Designing studies with
larger sample sizes may provide more evidence of the relation but will require additional
attention to other potential variables that were controlled in this study. The teacher, for
instance, was the same for both classes, which controlled the implementation of the
treatment and ensured that the classroom environments were the same. As the sample
size increases, it will not be possible to have only one instructor. The concern with this
would be controlling the implementation of the treatment, especially if verbal feedback is
included, as well as environmental issues such as classroom expectations and rules. As
previously mentioned, the more controls on a study, the more internal validity it will have
but there is a trade-off. The more variables and diversity in the design, the more
generalizable it will be to other situations. This is an important point in education, where
the purpose for the research may be to apply the findings to other settings.

Another noteworthy endeavor might be to investigate the possibility that verbal
self-referenced feedback could make more of an impact on math self-efficacy than what
was observed in this study. Using verbal feedback would also allow the inclusion of some of the characteristics that were noted as being different between persons of different levels of self-efficacy (Bandura, 1982b). For instance, individuals with high self-efficacy were noted as being persistent and this could be used as a source of feedback with students in the classroom. The impact of the self-referenced feedback on math self-efficacy and math achievement might be greater if it were possible to provide verbal feedback to students, recognizing changes in their persistence on difficult problems, for example. Noting and giving feedback on other types of behaviors could prove to be effective as well, such as providing verbal feedback when students use different strategies to solve difficult problems, another characteristic exhibited by individuals with high self-efficacy.

The next step for these results would be to test its generalizability to other groups of students, perhaps starting with classes within the same school. If similar results to this study are obtained, then moving to other schools with different demographics that were not investigated in this study, such as SES or race. It would be important to know that math self-efficacy could be raised and subsequently math achievement, but this would be most important if math self-efficacy could be raised across a diverse student population. Public education has stringent expectations to meet the demands of both the government and society. Being able to boost math achievement by affecting change in math self-efficacy has great potential.
REFERENCES


Appendix A

Kentucky High School Diagnostic Mathematics Assessments [DRAFT]

Algebra 1, Version 1 Selected Items

Directions: Circle the best answer to each problem. ID Number ________________

1. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed.

<table>
<thead>
<tr>
<th>Number of pots of coffee</th>
<th>Number of customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>c</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>56</td>
</tr>
</tbody>
</table>

Which function best describes the relationship between the number of customers who were served coffee and the number of pots of coffee?

A. c = 4p + 8       B. c = 8p       C. p = c - 14       D. p = 8c

2. A rectangular garden is 6 feet longer than it is wide. If the perimeter of the garden is 52 feet, which equation can you use to find its width?

A. w + w + 6 = 52       B. 2w + 2(w + 6) = 52
C. 2w + 2(w-6) = 52     D. w + w – 6 = 52

3. Factor completely 2x³ - 5x² + 2x - 5

A. x²(2x - 5) + 2x - 5       B. (2x -5)x²       C. (2x - 5)(x² + 1)       D. (2x + 1)(2x - 5)

4. Using the table below, find the value of n if f(n) = 113

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>f(n)</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>14</td>
<td>…</td>
<td>113</td>
</tr>
</tbody>
</table>

A. 35       B. 36       C. 37       D. 38
5. For which inequality is the following the solution for?

\[ \frac{3}{4} \]

A. \( x - 1 < 4 \)  
B. \( x - 3 \leq 2 \)  
C. \( x + 2 > 3 \)  
D. \( x + 5 \geq 0 \)

6. Which of the following is an arithmetic sequence?

A. 1, 4, 9, 16, …  
B. 74, 71, 68, 65, …  
C. 2, 4, 8, 16, …  
D. 2, -3, 4, -5, 6, ...

7. Solve the equation \( V = \frac{1}{3} \)lwh for \( h \).

A. \( h = \frac{1}{3} \frac{V}{wl} \)  
B. \( h = V - \frac{1}{3}lw \)  
C. \( h = \frac{V}{3wl} \)  
D. \( h = \frac{3V}{wl} \)

8. What is the slope of a line that is perpendicular to the line \( y = 2x - 7 \)?

A. -2  
B. \( \frac{1}{2} \)  
C. \( -\frac{1}{2} \)  
D. 2

9. What is the slope of the line below?

A. \( -\frac{3}{2} \)  
B. \( \frac{3}{2} \)  
C. \( \frac{2}{3} \)  
D. \( -\frac{2}{3} \)
Appendix A, cont.

10. Solve the following equation for x:

\[ \frac{2}{3} x - 4 = 6 \]

A. 15  B. \( \frac{20}{3} \)  C. 11  D. 7

11. What is the domain of the function shown below?

A. \(-2 \leq x \leq 6\)  B. \(-4 \leq x \leq 6\)  C. \(-4 \leq y \leq 2\)  D. \(-2 \leq y \leq 6\)

12. Factor \(x^2 - 11x + 24\) completely.

A. \((x-2)(x+12)\)  B. \((x-3)(x-8)\)  C. \((x-6)(x-4)\)  D. \((x+3)(x-8)\)

13. Which equation below represents the equation of a line perpendicular to \(3x + 2y = 8\)?

A. \(2y + 3x = 12\)  B. \(-2x + 3y = -2\)  C. \(y = \frac{-3}{2} x + 6\)  D. \(y = -\frac{1}{3} x + 8\)

14. Simplify \(\frac{(r^2t)^3}{(rt^3)^2}\).

A. \(\frac{r^4}{t^5}\)  B. \(\frac{r}{t^2}\)  C. \(\frac{r^2}{t^2}\)  D. \(\frac{r^3}{t^2}\)
Appendix A, cont.

15. The ratio of the length \( l \) to the width \( w \) of a window is 5:3. If the window is 10 inches longer than it is wide, which proportion could be used to determine the dimensions of the window?

A. \( \frac{3}{5} = \frac{w + 10}{w} \)  
B. \( \frac{5}{3} = \frac{w + 10}{w} \)  
C. \( \frac{5}{3} = \frac{w}{w - 10} \)  
D. \( \frac{3}{5} = \frac{l}{w} \)

16. Given two lines are parallel, the solution set to the equations of the two lines would be:

A. infinitely many solutions  
B. no solution  
C. one solution  
D. two solutions

17. Which equation below represents the line that passes through the point (-6, 2) and is parallel to the line \( y = \frac{1}{3}x - 5 \)?

A. \( y = \frac{1}{3}x + 2 \)  
B. \( y = -3x - 16 \)  
C. \( y = \frac{1}{3}x + 4 \)  
D. \( y = 3x + 20 \)

18. Which graph best represents the equation \( 4y + 7x = 20 \)?

A.  
B.  
C.  
D.
19. The graph of the equation \( y = \frac{5}{3}x - 3 \) is given below. Graph \( y = -\frac{2}{3}x + 4 \) on the grid.

What is the solution to this system of equations?

A. (2, 1)  
B. (6.5, 8)  
C. (3, 2)  
D. (2, 3)

20. What must always be true of the solution set of a system of equations?

A. It is an infinite set of ordered pairs.  
B. It consists of all ordered pairs of numbers that satisfy both equations in the system.  
C. It consists of a single ordered pair of numbers.  
D. It consists of all ordered pairs of numbers that satisfy either equation in the system.
Appendix B

Kentucky High School Diagnostic Mathematics Assessments [DRAFT]

Algebra 1, Version 2 Selected Items

Directions: Circle the best answer to each problem. ID Number ________________

1. Which equation below represents the equation of a line perpendicular to $4x + y = -3$?
   A. $4x - y = -1$  
   B. $y = -x - 24$  
   C. $y = -4x + 1$  
   D. $x - 4y = -20$

2. Solve the equation $V = \frac{1}{3}lwh$ for $w$.
   A. $w = \frac{3V}{hl}$  
   B. $w = \frac{3V}{lh}$  
   C. $w = V - \frac{1}{3}lh$  
   D. $w = \frac{V}{3lh}$

3. Given two lines are the same, the solution set to the equations of the two lines would be:
   A. infinitely many solutions  
   B. no solution  
   C. one solution  
   D. two solutions

4. Factor $x^2 + 12x + 35$ completely.
   A. $(x - 7)(x+5)$  
   B. $(x - 7)(x - 5)$  
   C. $(x + 7)(x - 5)$  
   D. $(x + 7)(x + 5)$

5. Which equation below represents the line that passes through the point $(8, 5)$ and is parallel to the line $y = x + 3$.
   A. $y = x - 7$  
   B. $y = x + 12$  
   C. $y = x + 7$  
   D. $y = x - 12$

6. Which of the following is an arithmetic sequence?
   A. 0, 3, 8, 15, …  
   B. 3, 9, 27, …  
   C. 33, 35, 37, 39, …  
   D. 100, 50, 25, 12.5, …
Appendix B, cont.

7. From a recent United Parcel shipment of 750 packages, a sample of 30 packages were inspected and 5 were found to be broken. Estimate the number of broken packages from the entire shipment.

   A. 3750   B. 125   C. 180   D. 200

8. What is the slope of a line that is perpendicular to the line \( y = -\frac{4}{3} x + 1 \)?

   A. \( \frac{3}{4} \)   B. \( -\frac{4}{3} \)   C. \( \frac{4}{3} \)   D. \( \frac{3}{4} \)

9. A rectangular reflection pool is 8 feet longer than it is wide. If the perimeter of the pool is 60 feet, which equation can you use to find its width?

   A. \( 2w + 2(w + 8) = 60 \)   B. \( w + w + 16 = 60 \)

   C. \( 2w + 2(w - 8) = 60 \)   D. \( w + w - 8 = 60 \)

10. Which will never be true for the graphs of a system of linear equations?

   A. They are parallel.
   B. They are both the same line.
   C. They intersect in exactly one point.
   D. They intersect in exactly two points.

11. For which inequality is the following the solution for?

   ![Graph showing the solution set for the inequality.]

   A. \( x + 3 > 5 \)   B. \( x - 7 \leq -5 \)   C. \( x + 4 \geq 6 \)   D. \( x - 1 > -1 \)
12. Which graph best represents the equation $y - x = 5$?

A.  
B.  
C.  
D.  

13. Solve the following equation for $x$:

$$\frac{3}{5}x - 3 = 9$$

A. 10  
B. 15  
C. $\frac{18}{5}$  
D. 20

14. Using the table below, find the value of $n$ is $f(n) = 36$?

<table>
<thead>
<tr>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(n)$</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>12</td>
<td>...</td>
<td>36</td>
</tr>
</tbody>
</table>

A. 11  
B. 12  
C. 13  
D. 14
Appendix B, cont.

15. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed.

<table>
<thead>
<tr>
<th>Number of pots of coffee</th>
<th>Number of customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$c$</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>81</td>
</tr>
<tr>
<td>12</td>
<td>108</td>
</tr>
</tbody>
</table>

Which function best describes the relationship between the number of customers who were served coffee and the number of pots of coffee?

A. $p = 9c$  
B. $p = 6p + 9$  
C. $c = 9p$  
D. $c = p + 9$

16. What is the domain of the function shown below?

A. $2 < x < 9$  
B. $-2 \leq x \leq -9$  
C. $-2 \leq x \leq 9$  
D. $2 < x < -9$

17. Factor completely $3x^3 - 15x^2 + 2x - 10$.

A. $x^2(3x - 15) + 2(x - 5)$  
B. $3x^2(x - 5) + 2(x - 5)$  
C. $(3x^2 + 2)(x - 5)^2$  
D. $(3x^2 + 2)(x - 5)$
Appendix B, cont.

18. The graph of the equation \( y = \frac{-1}{4}x + 4 \) is given below. Graph \( y = \frac{5}{4}x - 2 \) on the grid.

![Graph of y = \( \frac{-1}{4}x + 4 \) and \( y = \frac{5}{4}x - 2 \)]

What is the solution to this system of equations?

A. (0, 4)  
B. (2, 3.5)  
C. (4, 3)  
D. No Solution

19. What is the slope of the line below?

![Graph of a line]

A. \( \frac{2}{3} \)  
B. \( -\frac{3}{2} \)  
C. \( -\frac{2}{3} \)  
D. \( \frac{3}{2} \)

20. Simplify \( \frac{-12x^5y^6}{-6x^4y^7} \)

A. \( \frac{-2x}{y} \)  
B. \( \frac{x}{2y} \)  
C. \( \frac{2x^9}{y^{13}} \)  
D. \( \frac{2x}{y} \)
Appendix C

Kentucky High School Diagnostic Mathematics Assessments [DRAFT]

Algebra 1, Version 3 Selected Items

Directions: Circle the best answer to each problem. ID Number ________________

1. Which equation below represents the line that passes through the point (-6, -2) and is parallel to the line \( y = x - 5 \)?
   A. \( y = \frac{1}{3}x \)    B. \( y = -3x - 16 \)    C. \( y = \frac{1}{3}x + 4 \)    D. \( y = 3x \)

2. What is the slope of a line that is perpendicular to the line \( y = 3x - 7 \)?
   A. 3    B. \(-\frac{1}{3}\)    C. \(\frac{1}{3}\)    D. -3

3. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of cups of coffee that were consumed.

<table>
<thead>
<tr>
<th>Number of cups of coffee</th>
<th>Number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>p</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>45</td>
</tr>
<tr>
<td>7</td>
<td>63</td>
</tr>
</tbody>
</table>

Which function best describes the relationship between the number of customers who were served coffee and the number of pots of coffee?

   A. \( c = 5p + 8 \)    B. \( p = 9c \)    C. \( p = c - 16 \)    D. \( c = 9p \)

4. Factor completely \( x^3 + 5x^2 + 3x + 15 \).
   A. \( 3x^2(x + 5) + 3(x + 5) \)    B. \( (x^2 + 3)(x + 5) \)    C. \( (x^2 + 3) + (x + 5) \)    D. \( (x^2 + 5)(x + 3) \)
Appendix C, cont.

5. Simplify \( \frac{(\frac{r^2}{t^3})^4}{(\frac{r^3}{t^3})^{13}} \).

A. \( \frac{r^5}{t^2} \)  
B. \( \frac{r^5}{t} \)  
C. \( \frac{r^{11}}{t^{13}} \)  
D. \( \frac{r^8}{t^9} \)

6. Which must always be true for the solution set of a system of linear equations?

A. It consists of all ordered pairs of numbers that satisfy both equations in the system.
B. It is an infinite set of ordered pairs.
C. It consists of all ordered pairs of numbers that satisfy either equation in the system.
D. It consists of a single ordered pair of numbers.

7. What is the domain of the function shown below?

A. \(-8 < x < 0\)  
B. \(-8 \leq x \leq 0\)  
C. \(8 < x \leq 0\)  
D. \(-8 < x \leq 0\)

8. The ratio of the length \( l \) to the width \( w \) of a window is 7:4. If the window is 8 inches longer than it is wide, which proportion could be used to determine the dimensions of the window?

A. \( \frac{4}{7} = \frac{w + 8}{w} \)  
B. \( \frac{7}{4} = \frac{w}{w + 8} \)  
C. \( \frac{7}{4} = \frac{w + 8}{w} \)  
D. \( \frac{4}{7} = \frac{l}{w} \)
Appendix C, cont.

9. What is the slope of the line below?

\[
\text{A. } -\frac{3}{2} \quad \text{B. } \frac{3}{2} \quad \text{C. } \frac{2}{3} \quad \text{D. } -\frac{2}{3}
\]

10. Which graph best represents the equation \(4y - 7x = 20\)?
Appendix C, cont.

11. Using the table below, find the value of \( n \) if \( f(n) = 48 \)?

<table>
<thead>
<tr>
<th>( n )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(n) )</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>16</td>
<td>…</td>
<td>48</td>
</tr>
</tbody>
</table>

A. 9  
B. 10  
C. 11  
D. 12

12. A rectangular garden is 8 feet longer than it is wide. If the perimeter of the garden is 32 feet, which equation can you use to find its width?

A. \( w + w + 8 = 32 \)  
B. \( 2w + 2(w + 8) = 32 \)  
C. \( 2w + 2(w - 8) = 32 \)  
D. \( w + w - 8 = 32 \)

13. The graph of the equation \( y = \frac{-2}{3} x - 3 \) is given below. Graph \( y = \frac{5}{3} x + 4 \) on the grid.

What is the solution to this system of equations?

A. (1, -3)  
B. (3, -1)  
C. (-3, -1)  
D. No Solution
14. Factor $x^2 - 7x + 12$ completely.
   A. $(x + 2)(x + 6)$  B. $(x - 4)(x - 3)$  C. $(x + 4)(x + 3)$  D. $(x + 12)(x - 1)$

15. Which of the following is an arithmetic sequence?
   A. 0, -3, -6, -15, …  B. 4, 16, 64, …  C. -27, -29, -31, -33, …  D. 100, 50, 25, 12.5, …

16. For which inequality is the following the solution?
   ![Number line](image)
   A. $x - 1 < 4$  B. $x - 3 \leq -1$  C. $x + 2 > 5$  D. $x + 3 \geq 0$

17. Solve the following equation for $x$:
   \[
   \frac{2}{3}x - 5 = 5
   \]
   A. 15  B. $x - 3 \leq -1$  C. 11  D. 6

18. Given two lines are parallel, the solution set to the equations of the two lines would be:
   A. no solution  B. two solutions  C. infinitely many solutions  D. one solution

19. Which equation below represents the equation of a line perpendicular to $2x - 3y = 8$?
   A. $2y - 3x = 12$  B. $y = \frac{3}{2}x - 6$  C. $y = -\frac{3}{2}x + 6$  D. $6x + 4y = 24$

20. Solve the equation $V = \frac{1}{3}lwh$ for $w$.
   A. $w = \frac{V}{\frac{1}{3}hl}$  B. $w = V - \frac{1}{3}lh$  C. $w = \frac{V}{3lh}$  D. $w = \frac{3V}{hl}$
Appendix D
Kentucky High School Diagnostic Mathematics Assessments [DRAFT]

Algebra 1, Version 4, Adapted Items

Directions: Circle the best answer to each problem. ID Number ________________

1. What is the slope of a line that is perpendicular to the line \( y = 7x - 2 \)?
   
   A. \(-7\)  
   B. \(\frac{1}{7}\)  
   C. \(-\frac{1}{7}\)  
   D. 7

2. The graph of the equation \( y = \frac{5}{3}x - 3 \) is given below. Graph \( y = -\frac{2}{3}x - 3 \) on the grid.

   ![Graph](image)

   What is the solution to this system of equations?
   
   A. (2, 1)  
   B. (6.5, 8)  
   C. (3, 2)  
   D. (0, -3)

3. Using the table below, find the value of \( n \) if \( f(n) = 113 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(n) )</td>
<td>11</td>
<td>14</td>
<td>17</td>
<td>20</td>
<td>…</td>
<td>113</td>
</tr>
</tbody>
</table>

   A. 35  
   B. 36  
   C. 37  
   D. 38
Appendix D, cont.

4. A rectangular garden is 4 feet longer than it is wide. If the perimeter of the garden is 52 feet, which equation can you use to find its width?

A. \( w + w + 4 = 5 \)  \quad B. \( 2w + 2(w + 4) = 52 \)  \quad C. \( 2w + 2(w-4) = 52 \)  \quad D. \( w + w - 4 = 52 \)

5. For which inequality is the following the solution for?

0 1 2 3 4 5 6 7

A. \( x - 1 < 4 \)  \quad B. \( x - 3 \leq 2 \)  \quad C. \( x + 2 > 3 \)  \quad D. \( x + 5 \geq 0 \)

6. Given two lines are perpendicular, the solution set to the equations of the two lines would be:

A. infinitely many solutions  \quad B. no solution  \quad C. one solution  \quad D. two solutions

7. What is the range of the function shown below?

A. \(-2 \leq x \leq 6\)  \quad B. \(-4 \leq x \leq 6\)  \quad C. \(-4 \leq y \leq 2\)  \quad D. \(-2 \leq y \leq 6\)

8. Solve the following equation for \( x \):

\[ \frac{3}{2} x - 4 = 6 \]

A. 15  \quad B. \( \frac{20}{3} \)  \quad C. 11  \quad D. 7
9. Which of the following is an arithmetic sequence?
   A. 1, 4, 9, 16, …   B. 74, -71, 68, -65,   C. 1, 2, 4, 8, …   D. -2, -3, -4, -5, -6, …

10. The ratio of the length \( l \) to the width \( w \) of a window is 7:3. If the window is 10 inches longer than it is wide, which proportion could be used to determine the dimensions of the window?
   A. \( \frac{3}{7} = \frac{w + 10}{w} \)   B. \( \frac{7}{3} = \frac{w + 10}{w} \)   C. \( \frac{7}{3} = \frac{w}{w - 10} \)   D. \( \frac{3}{7} = \frac{l}{w} \)

11. Simplify \( \frac{(r^3t^2)^2}{(rt^2)^3} \)
   A. \( \frac{r^4}{t^3} \)   B. \( \frac{r}{t^5} \)   C. \( \frac{r^2}{t^4} \)   D. \( \frac{r^3}{t^5} \)

12. Which graph best represents the equation \( 3y + 5x = -15 \)?
   A.   B.   C.   D.
Appendix D, cont.

13. Factor $x^2 - 10x + 24$ completely.
   A. $(x-2)(x+12)$  B. $(x-3)(x-8)$  C. $(x-6)(x-4)$  D. $(x+3)(x-8)$

14. Solve the equation $V = \frac{1}{4} lwh$ for $h$.
   A. $h = \frac{4V}{wl}$  B. $h = V - \frac{1}{4} lw$  C. $h = \frac{V}{4wl}$  D. $h = \frac{4V}{wl}$

15. What is the slope of the line below?
   A. $\frac{3}{1}$  B. $\frac{1}{3}$  C. $-\frac{1}{3}$  D. $-\frac{3}{1}$

16. Which of these does NOT describe the solution set of a system of two equations that are parallel?
   A. There is no solution.
   B. The two graphs of the equations have no points in common.
   C. The graphs of the equations do not intersect.
   D. There is one ordered pair that makes both equations true.
17. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed.

<table>
<thead>
<tr>
<th>Number of pots of coffee</th>
<th>Number of customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>c</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>7</td>
<td>49</td>
</tr>
</tbody>
</table>

Which function best describes the relationship between the number of customers who were served coffee and the number of pots of coffee?

A. \( c = 4p + 8 \)  B. \( c = 7p \)  C. \( p = c - 14 \)  D. \( p = 8c \)

18. Which equation below represents the equation of a line perpendicular to \( 3x - y = 8 \)?

A. \( 2y + 3x = 1 \)  B. \( -2x + 3y = -2 \)  C. \( y = -\frac{3}{2} x + 6 \)  D. \( y = -\frac{1}{3}x + 8 \)

19. Which equation below represents the line that passes through the point (6, -2) and is parallel to the line \( y = \frac{1}{3}x - 5 \).

A. \( y = \frac{1}{3}x + 2 \)  B. \( y = -3x - 16 \)  C. \( y = \frac{1}{3}x \)  D. \( y = 3x + 20 \)

20. Factor completely \( 2x^3 - 10x^2 + x - 5 \)

A. \( x^2(2x - 5) + 2x - 5 \)  B. \( (2x - 5)x^2 \)  C. \( (x - 5)(2x^2 + 1) \)  D. \( (2x + 1)(2x - 5) \)
Directions: Circle the best answer to each problem.  ID Number ________________.

1. Given two lines have the same slope but different y-intercepts, the solution set to the equations of the two lines would be:
   A. infinitely many solutions  B. no solution  C. one solution  D. 7

2. Factor \( x^2 - 14x + 24 \) completely.
   A. \((x-2)(x-12)\)  B. \((x-3)(x-8)\)  C. \((x-6)(x-4)\)  D. \((x+3)(x-8)\)

3. Which graph best represents the equation \( 3y - 5x = 15 \)?
   A. 
   B. 
   C. 
   D. 

Appendix E
Kentucky High School Diagnostic Mathematics Assessments [DRAFT]

Algebra 1, Version 5, Adapted Items
Appendix E, cont.

4. What must always be true of the solution set of a system of equations that describe the same line?

A. It is an infinite set of ordered pairs.
B. There is no solution.
C. It consists of a single ordered pair of numbers.
D. There are two solutions.

5. For which inequality is the following the solution for?

\[
\begin{array}{c}
\text{0} \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \\
\end{array}
\]

A. \(x - 1 < 4\)  
B. \(x - 3 \leq 2\)  
C. \(x + 2 > 3\)  
D. \(x - 2 \geq 0\)

6. Factor completely \(2x^3 + 2x^2 - 5x - 5\)

A. \(x^2(2x - 5) + 2x - 5\)  
B. \((2x - 5)x^2\)  
C. \((2x - 5)(x^2 + 1)\)  
D. \((x + 1)(2x^2 - 5)\)

7. Solve the following equation for \(x\):

\[
\frac{2}{3}x + 4 = 6
\]

A. 15  
B. \(\frac{20}{3}\)  
C. 3  
D. 7

8. The ratio of the length \(l\) to the width \(w\) of a window is 5:3. If the window is 7 inches longer than it is wide, which proportion could be used to determine the dimensions of the window?

A. \(\frac{3}{5} = \frac{w + 7}{w}\)  
B. \(\frac{5}{3} = \frac{w + 7}{w}\)  
C. \(\frac{5}{3} = \frac{w}{w - 7}\)  
D. \(\frac{3}{5} = \frac{l}{w}\)
9. Using the table below, find the value of n if \( f(n) = 113 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(n) )</td>
<td>2</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>…</td>
<td>113</td>
</tr>
</tbody>
</table>

A. 35  
B. 36  
C. 37  
D. 38

10. Solve the equation \( V = \frac{1}{4} lwh \) for \( l \).

A. \( l = \frac{4V}{wh} \)  
B. \( l = V - \frac{1}{4} wh \)  
C. \( l = \frac{V}{4wh} \)  
D. \( l = \frac{4V}{wh} \)

11. Which of the following is an arithmetic sequence?

A. 1, 4, 9, 16, …  
B. 74, 71, 77, 68, …  
C. 2, 4, 6, 8, …  
D. 2, -3, 4, -5, 6, …

12. What is the slope of the line below?

A. \( -\frac{3}{5} \)  
B. \( \frac{3}{5} \)  
C. \( \frac{5}{3} \)  
D. \( -\frac{5}{3} \)
Appendix E, cont.

13. Which equation below represents the equation of a line perpendicular to \(-2x + 3y = 8\)?

A. \(2y + 3x = 12\)  
B. \(-2x + 3y = -2\)  
C. \(y = -\frac{3}{2}x + 6\)  
D. \(y = -\frac{1}{3}x + 8\)

14. Which equation below represents the line that passes through the point \((9, 7)\) and is parallel to the line \(y = \frac{1}{3}x - 5\).

A. \(y = \frac{1}{3}x + 2\)  
B. \(y = -3x - 16\)  
C. \(y = \frac{1}{3}x + 4\)  
D. \(y = 3x + 20\)

15. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed.

<table>
<thead>
<tr>
<th>Number of pots of coffee</th>
<th>Number of customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>c</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>55</td>
</tr>
<tr>
<td>7</td>
<td>77</td>
</tr>
</tbody>
</table>

Which function best describes the relationship between the number of customers who were served coffee and the number of pots of coffee?

A. \(c = 4p + 8\)  
B. \(c = 11p\)  
C. \(p = c - 14\)  
D. \(p = 8c\)

16. What is the slope of a line that is perpendicular to the line \(y = -2x - 7\)?

A. \(-2\)  
B. \(\frac{1}{2}\)  
C. \(-\frac{1}{2}\)  
D. \(2\)

17. Simplify \(\frac{(rt^3)^2}{(r^2t)^2}\).

A. \(\frac{r^4}{t^3}\)  
B. \(\frac{t^4}{r^2}\)  
C. \(\frac{r^2}{t^4}\)  
D. \(\frac{r^3}{t^2}\)
Appendix E, cont.

18. The graph of the equation \( y = \frac{5}{3}x - 3 \) is given below. Graph \( y = -\frac{4}{3}x + 6 \) on the grid.

What is the solution to this system of equations?

A. (2, 1)  
B. (6.5, 8)  
C. (3, 2)  
D. (2, 3)

19. A rectangular garden is 6 feet longer than it is wide. If the perimeter of the garden is 52 feet, which equation can you use to find its length?

A. \( l + l + 6 = 52 \)  
B. \( 2l + 2(l + 6) = 52 \)  
C. \( 2l + 2(l-6) = 52 \)  
D. \( l + l - 6 = 52 \)

20. What is the range of the function shown below?

A. \( 1 \leq x \leq 6 \)  
B. \(-2 \leq x \leq 9 \)  
C. \(-2 \leq y \leq 9 \)  
D. \( 1 \leq y \leq 6 \)
Appendix F

Please indicate how much confidence you have that you can successfully solve each of the problems shown on the overhead projector by circling the number according to the following 6-point confidence scale. Each problem will be shown for a short period of time.

| ID Number | ________________ |
| CONFIDENCE SCALE |
| No Confidence at All | Very Little Confidence | Some Confidence | Much Confidence | Great Confidence | Complete Confidence |
| 1 | 2 | 3 | 4 | 5 | 6 |

1. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed. Which function best describes the relationship.....

2. A rectangular garden is 6 feet longer than it is wide. If the perimeter of the garden is 52 feet, which equation can you use....

3. Factor completely $2x^3 - 5x^2 + 2x - 5$

4. Using the table below, find the value of n if $f(n) = 113$.......

5. For which inequality is the following the solution for?

6. Which of the following is an arithmetic sequence?

7. Solve the equation $V = \frac{1}{3} lwh$ for $h$.

8. What is the slope of a line that is perpendicular to the line $y = 2x - 7$?
<table>
<thead>
<tr>
<th>No Confidence at All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. What is the slope of the line below?</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Solve the following equation for x: (\frac{2}{3}x - 4 = 6)</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. What is the domain of the function shown below?</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Factor (x^2 - 11x + 24) completely.</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Which equation below represents the equation of a line perpendicular to (3x + 2y = 8)?</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Simplify (\frac{(r^2t)^3}{(rt^3)^2}).</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. The ratio of the length (l) to the width (w) of a window is 5:3. If the window is 10 inches longer than it is wide, which proportion could be used to determine the dimensions?</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Given two lines are parallel, the solution set to the equations of the two lines would be...</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Which equation below represents the line that passes through the point (-6, 2) and is parallel to the line</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Which graph best represents the equation (4y + 7x = 20)?</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. The graph of the equation (y = \frac{5}{3}x - 3) is given below. Graph (y = -\frac{2}{3}x + 4) on the grid. What is the solution to this system of equations?</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. What must always be true of the solution set of a system of equations?</td>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix G
Math Self-Efficacy Scale

Please indicate how much confidence you have that you can successfully solve each of the problems shown on the overhead projector by circling the number according to the following 6-point confidence scale. Each problem will be shown for a short period of time.

ID Number ______________

CONFIDENCE SCALE

<table>
<thead>
<tr>
<th>No Confidence at All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

1. Which equation below represents the equation of a line perpendicular to $4x + y = -3$?

2. Solve the equation $V = \frac{1}{3}lwh$ for $w$.

3. Given two lines are the same, the solution set to the equations of the two lines would be...

4. Factor $x^2 + 12x + 35$ completely.

5. Which equation below represents the line that passes through the point $(8, 5)$ and is parallel to the line ...  

6. Which of the following is an arithmetic sequence?

7. From a recent united parcel shipment of 750 packages, a sample of 30 packages were inspected and 5 were found to be broken. Estimate the number of broken packages from the ...

8. What is the slope of a line that is perpendicular to the line $y = -\frac{4}{3}x + 1$?

9. A rectangular reflection pool is 8 feet longer than it is wide. If the perimeter of the pool is 60 feet, which equation can you use to find its width?

10. Which will never be true for the graphs of a system of linear equations?
11. For which inequality is the following the solution for?

\[ -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \]

12. Which graph best represents the equation \( y - x = 5 \)?

13. Solve the following equation for \( x \):

\[
\frac{3}{5} x - 3 = 9
\]

14. Using the table below, find the value of \( n \) is \( f(n) = 36 \).

<table>
<thead>
<tr>
<th>Confidence At All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

15. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed. Which function best describes the relationship?

16. What is the domain of the function shown below?

17. Factor completely \( 3x^3 - 15x^2 + 2x - 10 \).

18. The graph of the equation \( y = \frac{-1}{4} x + 4 \) is given below.

Graph \( y = \frac{5}{4} x - 2 \) on the grid. What is the solution to this system of equations?

19. What is the slope of the line below?

20. Simplify \( \frac{-12x^5y^6}{-6x^4y^3} \)
Appendix H

Math Self-Efficacy Scale

Please indicate how much confidence you have that you can successfully solve each of the problems shown on the overhead projector by circling the number according to the following 6-point confidence scale. Each problem will be shown for a short period of time.

ID Number ________________

CONFIDENCE SCALE

<table>
<thead>
<tr>
<th>No Confidence at All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

1. Which equation below represents the line that passes through the point (-6, -2) and is parallel to the line \( y = 3x - 7 \)?

2. What is the slope of a line that is perpendicular to the line \( y = 3x - 7 \)?

3. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of cups of coffee that were consumed. Which function best describes the relationship?

4. Factor completely \( x^2 + 5x^2 + 3x + 15 \).

5. Simplify \( \left( \frac{r^2}{t^2} \right)^2 \).

6. Which must always be true for the solution set of a system of linear equations?

7. What is the domain of the function shown below?

8. The ratio of the length \( l \) to the width \( w \) of a window is 7:4. If the window is 8 inches longer than it is wide, which proportion could be used to determine the dimensions?

9. What is the slope of the line below?

10. Which graph best represents the equation \( 4y - 7x = 20 \)?

11. Using the table below, find the value of \( n \) if \( f(n) = 48 \)...
<table>
<thead>
<tr>
<th>Confidence at All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

12. A rectangular garden is 8 feet longer than it is wide. If the perimeter of the garden is 32 feet, which equation can you use to find...

13. The graph of the equation \( y = \frac{-2}{3} x - 3 \) is given below.

Graph \( y = \frac{5}{3} x + 4 \) on the grid. What is the solution to this system of equations?

14. Factor \( x^2 - 7x + 12 \) completely.

15. Which of the following is an arithmetic sequence?

16. For which inequality is the following the solution?

![Number Line]

17. Solve the following equation for \( x \):

\[
\frac{2}{3} x - 5 = 5
\]

18. Given two lines are parallel, the solution set to the equations of the two lines would be...

19. Which equation below represents the equation of a line perpendicular to \( 2x - 3y = 8 \)?

20. Solve the equation \( V = \frac{1}{3} lwh \) for \( w \)
Appendix I

Math Self-Efficacy Scale

Please indicate how much confidence you have that you can successfully solve each of the problems shown on the overhead projector by circling the number according to the following 6-point confidence scale. Each problem will be shown for a short period of time.

**CONFIDENCE SCALE**

<table>
<thead>
<tr>
<th>No Confidence at All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

1. What is the slope of a line that is perpendicular to the line $y = 7x - 2$?

2. The graph of the equation $y = \frac{5}{3}x - 3$ is given below.
   - Graph $y = -\frac{2}{3}x - 3$ on the grid.
   - What is the solution to this system of equations?

3. Using the table below, find the value of $n$ if $f(n) = 113$...

4. A rectangular garden is 4 feet longer than it is wide. If the perimeter of the garden is 52 feet, which equation can you use to find ...

5. For which inequality is the following the solution for?

6. Given two lines are perpendicular, the solution set to the equations of the two lines would be...

7. What is the range of the function shown below?

8. Solve the following equation for $x$:
   $$\frac{3}{2}x - 4 = 6$$
<table>
<thead>
<tr>
<th>No Confidence at All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

**9.** Which of the following is an arithmetic sequence?  
1 2 3 4 5 6

**10.** The ratio of the length \( l \) to the width \( w \) of a window is 7:3. If the window is 10 inches longer than it is wide, which proportion could be used to determine the...  
1 2 3 4 5 6

**11.** Simplify \( \frac{(r^3 t)^2}{(r^2 t^3)^3} \)  
1 2 3 4 5 6

**12.** Which graph best represents the equation \( 3y + 5x = -15 \)?  
1 2 3 4 5 6

**13.** Factor \( x^2 - 10x + 24 \) completely.  
1 2 3 4 5 6

**14.** Solve the equation \( V = \frac{1}{4} lwh \) for \( h \).  
1 2 3 4 5 6

**15.** What is the slope of the line below?  
1 2 3 4 5 6

**16.** Which of these does NOT describe the solution set of a system of two equations that are parallel?  
1 2 3 4 5 6

**17.** The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed. Which function best describes the relationship between the number of customers who were served coffee and the number of pots of coffee?  
1 2 3 4 5 6

**18.** Which equation below represents the equation of a line perpendicular to \( 3x - y = 8 \)?  
1 2 3 4 5 6

**19.** Which equation below represents the line that passes through the point (6, -2) and is parallel to the line ....  
1 2 3 4 5 6

**20.** Factor completely \( 2x^3 - 10x^2 + x - 5 \)  
1 2 3 4 5 6
Appendix J

Math Self-Efficacy Scale

Please indicate how much confidence you have that you can successfully solve each of the problems shown on the overhead projector by circling the number according to the following 6-point confidence scale. Each problem will be shown for a short period of time.

**ID Number ________________**

**CONFIDENCE SCALE**

<table>
<thead>
<tr>
<th>No Confidence at All</th>
<th>Very Little Confidence</th>
<th>Some Confidence</th>
<th>Much Confidence</th>
<th>Great Confidence</th>
<th>Complete Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

1. Given two lines have the same slope but different y-intercepts, the solution set to the equations of the two lines would be...

2. Factor $x^2 - 14x + 24$ completely.

3. Which graph best represents the equation $3y - 5x = 15$?

4. What must always be true of the solution set of a system of equations that describe the same line?

5. For which inequality is the following the solution for?

6. Factor completely $2x^3 + 2x^2 - 5x - 5$

7. Solve the following equation for $x$: $\frac{2}{3}x + 4 = 6$

8. The ratio of the length $l$ to the width $w$ of a window is 5:3. If the window is 7 inches longer than it is wide, which proportion could be used to determine the ...

9. Using the table below, find the value of $n$ if $f(n) = 113$

10. Solve the equation $V = \frac{1}{4}lwh$ for $l$.

11. Which of the following is an arithmetic sequence?
12. What is the slope of the line below?  

13. Which equation below represents the equation of a line perpendicular to \(-2x + 3y = 8\)?  

14. Which equation below represents the line that passes through the point \((9, 7)\) and is parallel to the line ...  

15. The owner of a convenience store recorded the number of customers in the store from 6:00 A.M. to 11:00 A.M. who were served coffee and the number of pots of coffee that were consumed. Which function best describes the relationship between ...  

16. What is the slope of a line that is perpendicular to the line \(y = -2x - 7\)?  

17. Simplify \(\frac{(rt^3)^2}{(r^2t)^2}\).  

18. The graph of the equation \(y = \frac{5}{3}x - 3\) is given below. Graph \(y = -\frac{4}{3}x + 6\) on the grid. What is the solution to this system of equations?  

19. A rectangular garden is 6 feet longer than it is wide. If the perimeter of the garden is 52 feet, which equation can you use...  

20. What is the range of the function shown below?
Appendix K

Children’s Self-Efficacy Scale

Used by permission from Information Age Publishing, Inc.

This questionnaire is designed to help us get a better understanding of the kinds of things that are difficult for students. Please rate how certain you are that you can do each of the things described below by circling the appropriate number. Your answers will be kept strictly confidential and will not be identified by name. **ID Number __________________**

Rate your degree of confidence by circling a number from 1 to 7 using the scale given below:

<table>
<thead>
<tr>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Self-Efficacy in Enlisting Social Resources**

1. Get teachers to help me when I get stuck on schoolwork
   1  2  3  4  5  6  7

2. Get another student to help me when I get stuck on schoolwork?
   1  2  3  4  5  6  7

3. Get adults to help me when I have social problems
   1  2  3  4  5  6  7

4. Get a friend to help me when I have social problems
   1  2  3  4  5  6  7

**Self-Efficacy for Academic Achievement**

5. Learn general mathematics
   1  2  3  4  5  6  7

6. Learn algebra
   1  2  3  4  5  6  7

7. Learn science
   1  2  3  4  5  6  7

8. Learn biology
   1  2  3  4  5  6  7

9. Learn reading, writing, and language skills
   1  2  3  4  5  6  7
Appendix K, cont.

<table>
<thead>
<tr>
<th>Task</th>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. Learn to use computers</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>11. Learn a foreign language</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>12. Learn social studies</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13. Learn English grammar</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><strong>Self-Efficacy for Self-Regulated Learning</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Finish my homework assignments by deadlines</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>15. Get myself to study when there are other interesting things to do</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>16. Always concentrate on school subjects during class</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>17. Take good notes during class instruction</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>18. Use the library to get information for class assignments</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>19. Plan my school work for the day</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>20. Organize my school work</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>21. Remember well information presented in class and textbooks</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>22. Arrange a place to study without distractions</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>23. Get myself to do school work</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Appendix K, cont.

<table>
<thead>
<tr>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Efficacy for Leisure Time Skills and Extracurricular Activities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Learn sports skills well</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>25. Learn dance skills well</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>26. Learn music skills well</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>27. Do the kinds of things that are needed to work on the school newspaper</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>28. Do the kinds of things needed to serve school government</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>29. Do the kinds of things needed to take part in school plays</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>30. Do regular physical education activities</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>31. Learn the skills needed for team sports (for example, basketball, volleyball, swimming, football, soccer)</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
</tbody>
</table>

Self-Regulatory Efficacy

<table>
<thead>
<tr>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>32. Resist peer pressure to do things in school that can get me into trouble</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>33. Stop myself from skipping school when I feel bored or upset</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>34. Resist peer pressure to smoke cigarettes</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
<tr>
<td>35. Resist peer pressure to drink beer, wine, or liquor</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
</tr>
</tbody>
</table>
### Appendix K, cont.

<table>
<thead>
<tr>
<th>Task</th>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>36. Resist peer pressure to smoke marijuana</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>37. Resist peer pressure to use pills</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(uppers, downers)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. Resist peer pressure to have sexual intercourse</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>39. Control my temper</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**Self-Efficacy to Meet Others’ Expectations**

<table>
<thead>
<tr>
<th>Task</th>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>40. Live up to what my parents expect of me</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>41. Live up to what my teachers expect of me</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>42. Live up to what my peers expect of me</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>43. Live up to what I expect of myself</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**Social Self-Efficacy**

<table>
<thead>
<tr>
<th>Task</th>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>44. Make and keep friends of the opposite sex</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>45. Make and keep friends of the same sex</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>46. Carry on conversations with others</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>47. Work well in a group</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**Self-Assertive Efficacy**

<table>
<thead>
<tr>
<th>Task</th>
<th>Cannot do at all</th>
<th>Moderately can do</th>
<th>Highly Certain can do</th>
</tr>
</thead>
<tbody>
<tr>
<td>48. Express my opinions when other classmates disagree with me</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Cannot do at all</td>
<td>Moderately can do</td>
<td>Highly Certain can do</td>
</tr>
<tr>
<td>---</td>
<td>-----------------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>49. Stand up for myself when I feel I am being treated unfairly</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
<td></td>
</tr>
<tr>
<td>50. Get others to stop annoying me or hurting my feelings</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
<td></td>
</tr>
<tr>
<td>51. Stand firm to someone who is asking me to do something unreasonable or inconvenient</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
<td></td>
</tr>
<tr>
<td>52. Get my parent(s) to help me with a problem</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
<td></td>
</tr>
<tr>
<td>53. Get my brother(s) and sister(s) to help me with a problem</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
<td></td>
</tr>
<tr>
<td>54. Get my parents to take part in school activities</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
<td></td>
</tr>
<tr>
<td>55. Get people outside the school to take an interest in my school (for example, community groups, churches)</td>
<td>1 2 3 4</td>
<td>5 6 7</td>
<td></td>
</tr>
</tbody>
</table>
Appendix L

Bem Sex Role Inventory, Sample Items

Below, you will find listed a number of personality characteristics. We would like you to use those characteristics to describe yourself, that is, we would like you to indicate, on a scale from 1 to 7, how true of you each of these characteristics is. Please do not leave any characteristic unmarked.

<table>
<thead>
<tr>
<th>Never or almost never true</th>
<th>Usually not true</th>
<th>Sometimes but infrequently true</th>
<th>Occasionally true</th>
<th>Often true</th>
<th>Usually true</th>
<th>Always or almost always true</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

1. Defend my own beliefs

2. Affectionate

3. Conscientious

4. Independent

5. Sympathetic

ID Number ________________
Appendix M

List of Comments for Each Treatment Group and Control Group

<table>
<thead>
<tr>
<th>Self-Referenced Feedback</th>
<th>Social-Referenced Feedback</th>
<th>Control Group - No Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. This is the highest you’ve scored so far on this unit.</td>
<td>1. You scored higher on this than the majority of your classmates.</td>
<td>1. Review your notes before class.</td>
</tr>
<tr>
<td>2. Your scores are getting better.</td>
<td>2. Your scores are increasing more than your classmates.</td>
<td>2. Try using a different strategy to work the problem.</td>
</tr>
<tr>
<td>3. This section appears easier for you than the last topic.</td>
<td>5. This section appears easier for you than most of the other students.</td>
<td>3. Use your notes to help with the homework.</td>
</tr>
<tr>
<td>4. You are showing more of your work--good for you!</td>
<td>6. You show more of your work than other students.</td>
<td>4. Don’t give up on difficult problems; try something else!</td>
</tr>
<tr>
<td>5. You are making less careless mistakes than you were.</td>
<td>8. You make less careless mistakes than other algebra one students.</td>
<td>5. Be sure to take all of the notes in class.</td>
</tr>
<tr>
<td>6. You are more persistent on the challenging problems than you were.</td>
<td>9. You are more persistent than most of the other algebra one students.</td>
<td>6. Watch for careless errors.</td>
</tr>
<tr>
<td>7. You look more sure of yourself than when you first started algebra class.</td>
<td>7. You look more sure of yourself than most of your algebra one classmates.</td>
<td>7. Write down questions you have about the homework.</td>
</tr>
<tr>
<td>8. This is not your best work.</td>
<td>8. This is a lower score than the majority of the other students.’</td>
<td>8. Study your notes before the exam.</td>
</tr>
</tbody>
</table>
Appendix M, cont.

9. This is a lower score than what you have scored on this unit so far.

9. This is a lower score than what the other students have scored on this unit.

9. Be sure to complete all of the homework.

10. Try reading the textbook before class.

11. Be sure to ask questions if you have any.
Appendix N

Feedback Comments

Self-Referenced feedback is feedback about the individual compared to his/her previous performance or behavior.
Social-Referenced feedback is feedback about the individual compared to the performance or behavior of others who are similar to the individual.

Directions: Please read each statement and circle the appropriate letter if the statement reflects self-referenced feedback, social-referenced feedback, or neither type of feedback.

A=Self  B=Social  C=Neither

<table>
<thead>
<tr>
<th></th>
<th>Find a similar problem in the examples to use as a guide.</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>You seem to understand this material better than the last chapter.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>2</td>
<td>Your homework looks more accurate than it did for the last chapter.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>3</td>
<td>Your scores are increasing more than your classmates.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>Be sure to ask questions if you have any.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>You look more sure of yourself than when you first started algebra class.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>6</td>
<td>Try reading the textbook before class.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>7</td>
<td>You look more sure of yourself than most of your algebra one classmates.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>8</td>
<td>This section appears easier for you than the last topic.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>9</td>
<td>Your homework is more accurate than most of your classmates.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>10</td>
<td>This is the highest you’ve scored so far on this unit.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>11</td>
<td>Review your notes before class.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>12</td>
<td>Write down questions you have about the homework.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>13</td>
<td>Try using a different strategy to work the problem.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>14</td>
<td>Be sure to complete all of the homework.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>15</td>
<td>There has been an improvement in your grade this grading period.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>16</td>
<td>You show more of your work than other students.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>17</td>
<td>You have improved your grade more than most of the other algebra one students in this grading period.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>18</td>
<td>Use your notes to help with the homework.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>19</td>
<td>You worked harder in class yesterday than most of your classmates.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>20</td>
<td>You seem to understand this material better than most of the other algebra one students.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>21</td>
<td>You make less careless mistakes than other algebra one students.</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

215
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>You seem to be paying closer attention during class than you were.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>24</td>
<td>You seem to pay closer attention during class than the majority of your classmates.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>25</td>
<td>You stayed on task more than most of your classmates yesterday.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>26</td>
<td>You are making less careless mistakes than you were.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>27</td>
<td>You scored higher on this than the majority of your classmates.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>28</td>
<td>Use your time wisely.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>29</td>
<td>You are showing more of your work--good for you!</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>30</td>
<td>You stayed more on task during the class period yesterday than you have been.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>31</td>
<td>You are more persistent than most of the other algebra one students.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>32</td>
<td>You worked harder in class yesterday than you have been.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>33</td>
<td>This section appears easier for you than most of the other students.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>34</td>
<td>Your scores are getting better.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>35</td>
<td>Study your notes before the exam.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>36</td>
<td>Don’t give up on difficult problems; try something else!</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>37</td>
<td>Be sure to take all of the notes in class.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>38</td>
<td>Watch for careless errors.</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>39</td>
<td>You are more persistent on the challenging problems than you were.</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>
### Appendix O

**Student Weekly Graph of Goals and Actual Scores**

**ID Number _____________________**  
**Week of________________**

<table>
<thead>
<tr>
<th>Goal for</th>
<th>Actual score</th>
<th>Goal for</th>
<th>Actual score</th>
<th>Goal for</th>
<th>Actual score</th>
<th>Goal for</th>
<th>Actual score</th>
<th>Goal for</th>
<th>Actual score</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td></td>
<td>90%</td>
<td></td>
<td>80%</td>
<td></td>
<td>70%</td>
<td></td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td></td>
<td>40%</td>
<td></td>
<td>30%</td>
<td></td>
<td>20%</td>
<td></td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Weekly Graph of 5-minute Checks: Goals and Actual Scores*
Appendix P

July 17, 2006  
REVISED: April 2, 2007

Namok Choi, PhD  
(Debbie Thompson)  
College of Education  
University of Louisville

RE: 394.06 - The Effects of Evaluative Feedback on Math Self-Efficacy and Math Achievement of Ninth Grade Algebra Students: A Longitudinal Approach

Dear Doctor Choi:

This study has been reviewed by the chair of the Institutional Review Board (IRB) and approved through the Expedited Review Procedure, according to 45 CFR 46.110(b), since (7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

This study was approved for children under category 1 - 45 CFR 46.404 - Research not involving greater than minimal risk. No greater than minimal risk to children is presented, only if the IRB finds that adequate provisions are made for soliciting the assent of the children and the permission of their parents or guardians, as set forth in Sec. 46.408. This category requires the assent of child (if over age 7) and at least one parent signature.

The following items have been approved:

- Research Protocol, not dated
- Informed Consent, not dated
- Subject Assent, not dated

This study now has final IRB approval through 7/13/2007. You should complete and return the Progress Report/Continuation Request Form EIGHT weeks prior to this date in order to ensure that no lapse in approval occurs. The committee will be advised of this action at their next full board meeting.

Please note that the IRB follows the principles of the Belmont Report, is in compliance with Good Clinical Practice Guidelines as defined by the U.S. Food and Drug Administration and the Department of Health and Human Services under the Code of Federal Regulations (21 CFR Parts 50 and 56; 45 CFR 46) and International Conference on Harmonization (ICH) Guidelines (Section E6).
Site Approval

If this study will take place outside of the University of Louisville Campuses, site approval from the organization must be given before the research may begin.

Best wishes for a successful study. Please send all inquiries and electronic revised/requested items to our office email address at hsppofo@louisville.edu.

Sincerely,

Patricia K. Leitsch

Patricia K. Leitsch, Ph.D., Chair,
Social/Behavioral/Educational Institutional Review Board
PKL/elp
June 16, 2006

To Whom It May Concern:

Deborah Thompson has permission to do a research study and collect data from ninth grade algebra one students at Oldham County High School during the school year 2006-2007.

David L. Weedman
Principal
Oldham County High School
Appendix R

As of right now, how certain are you that you can earn each of the grades listed below for algebra one? Please indicate the strength of your belief that you can achieve each of the grades below by circling the number according to the 6-point scale.

No Certainty
Very Little Certainty
Some Certainty
Much Certainty
Great Certainty
High Certainty
1
2
3
4
5
6

For example, if you were highly certain that you could do 1 sit-up, you would mark a “6.” If you were not certain at all that you could do 1000 sit-ups, you would mark a “1.”

| Achieve an F in algebra one (0 to 72%). | 1 2 3 4 5 6 |
| Achieve a C- in algebra one (73 to 74%). | 1 2 3 4 5 6 |
| Achieve a C in algebra one (75 to 80%). | 1 2 3 4 5 6 |
| Achieve a C+ in algebra one (81 to 82%). | 1 2 3 4 5 6 |
| Achieve a B- in algebra one (83 to 84%). | 1 2 3 4 5 6 |
| Achieve a B in algebra one (85 to 90%). | 1 2 3 4 5 6 |
| Achieve a B+ in algebra one (91 to 92%). | 1 2 3 4 5 6 |
| Achieve an A- in algebra one (93 to 94%). | 1 2 3 4 5 6 |
| Achieve an A in algebra one (95 to 98%). | 1 2 3 4 5 6 |
| Achieve an A+ in algebra one (99 to 100%). | 1 2 3 4 5 6 |
Thank you so much for your valuable insight. I am honored that you took time to assist me with this question. I know that you are very busy but you are the best resource on this issue. I am grateful that you had time to provide direction for me and look forward to utilizing the information.

Thank you again,
Debbie Thompson

-----Original Message-----
From: "Prof. Frank Pajares" <mpajare@emory.edu>
To: "Thompson, Deborah Burnett" <dbthom01@netmail.louisville.edu>
Date: Mon, 3 Jul 2006 06:17:39 -0700 (PDT)
Subject: Re: math self-efficacy

I understand your dilemma. If you don't use an outcome measure with the same level of difficulty, it will be impossible to compare the influence of self-efficacy across administrations. Critics will argue that potential differences in level of difficulty will confound your findings.

I agree that it is unwise to use unit assessments. Not only would this be problematic as regards differing levels of difficulty, you face the possibility of obtaining elevated levels of both self-efficacy and achievement.

One way around this would be to find an assessment that has multiple questions testing the same concept, or, alternatively, an appropriate test that has several equivalent forms.

With a test that has a broad bank of questions testing the same concepts, you could split the test into several subtests and administer each test once per month. You would be testing the same concepts each time, but with a different set of questions. A test with alternate equivalent forms is equally useful.

What is essential, I think, is to ensure that all of the tests administered (and the self-efficacy assessments obtained) are equivalent forms of the final test. Only in this way can you successfully measure growth.

This would give you a measure of self-efficacy for performance. You could compare that with self-efficacy for learning by including a self-efficacy assessment each time of their confidence to obtained specific grades in mathematics (I've done that several times and it works very
well). Each provides different but quite useful information. You'll no doubt find that, early on, the students' self-efficacy for performance will be quite low when faced with math problems sure to be unfamiliar, but their grade self-efficacy will not be similarly affected (this gets at the idea of "I know I can't solve these unfamiliar problems right now, but I believe I will succeed in learning how to solve them").

I believe that the most interesting wrinkle of a longitudinal study would be to compare the evolution of the two types of self-efficacy.

"Thompson, Deborah Burnett" <dbthom01@louisville.edu> wrote:

Hi Dr. Pajares;
Thank you so very much for your response to my inquiry. I totally understand if you do not have time to assist me, but just in case, I have provided a description of my proposed design. The question is the specific nature of the monthly math achievement assessment. I had planned to use an overall algebra one diagnostic exam but am concerned that it is too broad to use for the entire semester. On the other hand, if I want to analyze a trend in math self-efficacy and math achievement over the semester, then it does not seem appropriate to use unit assessments.

Again, thank you so very much for responding to my request. If you do not have time to advise me, I understand.

I am grateful for your response to my initial inquiry.
Sincerely,
Debbie Thompson

Design
This research study is a quasi-experimental using repeated measures with a control group and two treatment groups. Two intact classes of ninth grade algebra I students will be used for the research. Most of the ninth grade students recommended for algebra one in the ninth grade have taken prealgebra in the eighth grade while a small number of ninth grade have taken algebra one in the eighth grade and are told to repeat algebra one in the ninth grade. The school uses a block schedule so the algebra one course will be completed by the middle of the school year (January, 2007). The algebra one-all year course is designed for students who are lacking in basic skills and also for students with special needs. The purpose of this study is to increase math self-efficacy and it is noted in the literature that this will not override a lack of basic skills. Therefore, the algebra one course was chosen since this is the group of students who are more likely to have lower levels of self-efficacy but adequate basic skills. The freshman geometry and advanced freshman geometry classes are designed for the upper level and gifted students. Typically, students taking classes at that level have higher self-efficacy than those students taking the lower level courses such as algebra one.

One third of each of the algebra one classes will be used as a control and the remaining two thirds of each class as the treatment groups. There will be a constant treatment of goal setting for all students and one independent variable of type of feedback. The independent variable will have 3 conditions, self-referenced feedback, social-referenced...
feedback, and no feedback. One third of the students in each of the classes will be randomly assigned to receive self-referenced feedback and another third in each of the classes randomly assigned to receive social-referenced feedback. The remaining one third of each class will receive no feedback. The approximate sample sizes are 16, depending upon actual enrollment once school begins. Math self-efficacy and math achievement will be measured monthly throughout the semester.
Hierarchical Linear Modeling (HLM) is comprised of a system of equations, one for each level of analysis. A basic HLM system involves a linear equation at level one with one dependent variable, its intercept and slope with one independent variable. Such as, $Y = \beta_0 + \beta_1 X_1$, where $Y$ is the dependent variable, $\beta_0$ is the intercept, $\beta_1$ is the slope, and $X_1$ is the independent variable. This level one equation is analogous to Ordinary Least Squares regression. In the level two equation for HLM, however, the intercept and the slope both become dependent variables with corresponding intercepts and slopes. So in a level two equation for this example, $\beta_0 = \gamma_{00} + \gamma_{01} W_1$, where $\beta_0$ is the intercept as a dependent variable, with $\gamma_{00}$ as the intercept, $\gamma_{01}$ as the slope, and $W_1$ as the independent variable. In addition to the intercept $\beta_0$ becoming a dependent variable, the slope, $\beta_1$, also becomes a dependent variable in a separate equation, such as $\beta_1 = \gamma_{10} + \gamma_{11} W_1$. The slope, $\beta_1$, now has an intercept, $\gamma_{10}$, a slope, $\gamma_{11}$, and an independent variable, $W_1$.

Combining the system of equations into one produces a mixed model equation like:

$$Y = \gamma_{00} + \gamma_{01} W_1 + \gamma_{10} X_1 + \gamma_{11} W_1 X_1.$$ 

Using HLM also allows for partitioning of the variance, or error, associated with each level of the model. The example just used is then expanded into $Y = \beta_0 + \beta_1 X_1 + r$, where $r$ is the random effect or error associated with level one. Analogous to the random effects in level one, each of the equations in level two have unique effects, or error. So the two previous equations then become $\beta_0 = \gamma_{00} + \gamma_{01} W_1 + u_0$ and $\beta_1 = \gamma_{10} + \gamma_{11} W_1 + u_1$, where $u_0$ and $u_1$ are unique effects, or error, associated with the intercept and slope.
Appendix T, cont.

respectively. The mixed model now includes each of the random and unique effects, or error terms: \( Y = \gamma_{00} + \gamma_{01}W_1 + \gamma_{10}X_1 + \gamma_{11}W_1 X_1 + u_0 + u_1X_1 + r. \)

Different models can be built by adding more independent variables to each of the equations. Multilevel modeling analysis is typically accomplished by starting with the intercept model and building more complex models. More complex models can be structured to examine the effects of added independent variables to equations in the model or using the slopes in the model as dependent variables predicted by independent variables. The fit of models is measured by the deviance, that is, the measure of the lack of fit between the data and the model. This deviance is compared from one model to another to determine the best fit. A chi-square test can be conducted on the difference between two deviances to determine if one model is significantly better than another. The difference in deviances is compared to a chi-square distribution using the difference in parameters of the models as the degrees of freedom (Singer & Willett, 2003).

One advantage of using HLM, in addition to the ability to analyze complex designs and partitioning the variance according to levels, is the ability to analyze data when the sample sizes are unequal. For example, when considering longitudinal designs, it is often the case that all measurements are not obtained for each participant. In other statistical methods of analysis, the participants with incomplete measurements would have to be either omitted or data imputed. With HLM, unequal numbers of measurements are allowed. Multilevel modeling focuses on the influence of the group membership of the individual characteristics. The more alike group members are, the smaller the sample size can be; and the less alike group members are, the larger the
sample needs to be. Therefore, HLM does not require all measurements on each individual or for missing values to be replaced by statistical methods, such as using a mean value (Luke, 2004).
Subject Informed Consent for
Effects of Teaching and Learning Methods on Math Achievement of
Ninth Grade Algebra One Students

Investigators:
Debbie Thompson
Oldham County High School
P.O. Box 187
Buckner, KY 40014

Dr. Namok Choi
College of Education & Human Development, Rm. 338
University of Louisville
Louisville, KY 40292

Site where study is to be conducted:
Oldham County High School

Phone number for subjects to call for questions: 502-222-9461 ext 236

Your child is invited to participate in a research study. The study is being conducted by
Dr. Namok Choi and Debbie Thompson, PhD candidate. The study is sponsored by the
University of Louisville, Department of Education and Human Development. The study
will take place at Oldham County High School. Approximately 60 ninth grade algebra
one students will be invited to participate. Debbie Thompson will be the instructor for
the algebra one class but the informed consent and assent letters will be collected by a
third party and will be held from the instructor until grades have been submitted at the
end of the semester.

The purpose of this study is to study the long-term effects of specific teaching and
learning methods. The purpose is to identify methods that will increase math
achievement of ninth grade algebra one students.

In this study, your children will be asked to complete personal inventories about
personality characteristics, confidence levels, and complete a short algebra test every few
weeks. The students will be asked to keep track of daily progress goals and graph their
results. Previous math scores will be obtained from school records. Students will be
randomly assigned to one of three learning methods. Randomly assigned is a process like
drawing a number from a hat. The study will last the entire semester. The time required
for the progress algebra tests will require about one class period every month. Students
will be given a complete course in algebra one. They may decline to answer any question
that makes them uncomfortable.

There are no foreseeable risks associated with this study of ninth grade algebra one
students, although there may be unforeseeable risks. Grades of students will not be
included for any performance on the monthly algebra progress tests.

Page 1 of 3
Appendix U, cont.

Effects of Teaching and Learning Methods on Math Achievement of Ninth Grade Algebra One Students

The possible benefits of this study include an increase in math achievement or other academic skills. The information collected may not benefit your child directly. The information learned in this study may be helpful to others.

Total privacy cannot be guaranteed. Your privacy will be protected to the extent permitted by law. If the results from this study are published, your name will not be made public. While unlikely, the following may look at the study records:

The University of Louisville Institutional Review Board, Human Subjects Protection Program Office.

The information gathered from the surveys and algebra tests and previous math scores will be identified by an assigned identification number instead of a name. Only the researcher will know the identification numbers. Any data or information gathered will be kept in a locked place throughout the semester or in a password protected computer file in the co-principal investigator’s classroom.

Taking part in this study is voluntary. You may choose for your child not to take part at all. If you decide for your child to be in this study you may have your child stop taking part at any time. If you decide for your child not to be in this study or if you have your child stop taking part at any time, your child will not lose any benefits for which they may qualify. You will be told about any changes that may affect your decision to continue in the study.

If you have any concerns or complaints about the study or the study staff, you have three options.

1. You may contact the principal investigator at 502-852-4014.

2. If you have any questions about your rights as a study subject, questions, concerns or complaints, you may call the Human Subjects Protection Program Office (HSPPO) (502) 852-5188. You may discuss any questions about your rights as a subject, in secret, with a member of the Institutional Review Board (IRB) or the HSPPO staff. The IRB is an independent committee composed of members of the University community, staff of the institutions, as well as lay members of the community not connected with these institutions. The IRB has reviewed this study.

3. If you want to speak to a person outside the University, you may call 1-877-852-1167. You will be given the chance to talk about any questions, concerns or complaints in secret. This is a 24 hour hot line answered by people who do not work at the University of Louisville.
Appendix U, cont.

Effects of Teaching and Learning Methods on Math Achievement of Ninth Grade Algebra One Students

This paper tells you what will happen during the study if you choose for your child to take part. Your signature means that this study has been discussed with you, that your questions have been answered, and that your child will take part in the study. This informed consent document is not a contract. You are not giving up any legal rights by signing this informed consent document. You will be given a signed copy of this paper to keep for your records.

Printed Name of Subject/Legal Representative

Signature of Subject/Legal Representative Date Signed

Signature of Person Explaining the Consent Form (if other than the Investigator) Date Signed

Signature of Investigator Date Signed

LIST OF INVESTIGATORS
Dr. Namok Choi 502-852-4014
Debbie Thompson 502-222-9461 ext 236

UNIVERSITY OF LOUISVILLE
INSTITUTIONAL REVIEW BOARD
Date Appro'd 7/14/04 Valid Thru 7/13/07

230
Appendix V

ASSENT FORM
Effects of Teaching and Learning Methods on Math Achievement of Ninth Grade Algebra One Students

I am invited to be in a research study being done by Mrs. Debbie Thompson. When a person is in a research study, they are called a "subject". I am invited because I am a student in Algebra One class this semester.

This means that I will complete two short surveys on the first day of class and then complete a survey and take an algebra test every few weeks to assess my progress in the class. In addition, my previous math grades will be accessed. I will also record my daily goals. There are no known risks for my participation in this research study. Completing the items will not be included in my grades.

This study will last the entire semester. Participating in this study may increase my math achievement or it may have no direct benefit to me. The results of the study may provide information that assists in the preparation of more effective math lessons.

My family, Mrs. Debbie Thompson, administrative staff of Oldham County High School, and classmates will know that I’m in the study. If anyone else is given information about me, they will not know my name. A number will be used instead of my name.

I have been told about this study and know why it is being done and what I have to do. My parent(s) have agreed to let me be in the study. If I have any questions I can ask Mrs. Debbie Thompson. She will answer my questions. If I do not want to be in this study or I want to quit after I am already in this study, I can tell Mrs. Debbie Thompson and she will discuss this with my parents.

Printed Name of Subject                        Signature of Subject                        Date Signed

Printed Name of Parent/Guardian                Signature of Parent/Guardian                Date Signed

Printed Name of Investigator                   Signature of Investigator                   Date Signed

UNIVERSITY OF LOUISVILLE
INSTITUTIONAL REVIEW BOARD
Date Appr's 7/14/00, Valid Thru 7/13/07
For IRB Approval Stamp

Assent version date __________________________

UNIVERSITY OF LOUISVILLE
INSTITUTIONAL REVIEW BOARD
Date Appr's 7/14/00, Valid Thru 7/13/07
For IRB Approval Stamp
Provided you properly cite IAP this can be used on a one time basis and not again without written permission.

Good luck

G

*************************************************
George F. Johnson
IAP - Information Age Publishing Inc
203-661-7602
www.infoagepub.com

-----Original Message-----
From: Thompson, Deborah Burnett [mailto:dbthom01@louisville.edu]
Sent: Sunday, July 02, 2006 2:38 PM
To: infoage@infoagepub.com
Subject: inquiry

Hello;
I am a doctoral student and would like permission to use the Children's Self-Efficacy Scale on pages 326 and 327 of Self-Efficacy Beliefs of Adolescents (ISBN 1-59311-366-8) as an instrument in my dissertation research. Thank you.

Debbie Thompson
Doctoral Student,
University of Louisville
Date: ___March 15, 2007___

To whom it may concern,

This letter is to grant permission for: __Deborah J. Thompson___
to use the following copyright material;

Instrument:  

_Bem Sex Role Inventory_

Author:  

_Sandra Lipsitz Bem_

Copyright:  

_1978 Consulting Psychologists Press. All Rights Reserved._

for her/his thesis research.

In addition, five (5) sample items from the instrument may be reproduced for inclusion in a proposal, thesis or dissertation.

The entire measure may not at any time be included or reproduced in other published material.

Sincerely,

[Signature]

Sandra Darrow
Customer Service
CURRICULUM VITAE

Deborah Thompson

WORK ADDRESS
Oldham County High School
P.O. Box 187
Buckner, KY 40010
502-222-9461 ext 236
fax: 502-222-0558
debbie.thompson@oldham.kyschools.us

CERTIFICATION HELD
Commonwealth of Kentucky
Rank I, Effective 7-01-2001

- KMA Professional Certificate for Teaching Mathematics, Grades 8 through 12;
  Exp. Date: 6-30-2011
- PHB Provisional High School Certificate, Grades 7 through 12
  Exp. Date: Life
- A92 Area Specialization: Home Economics Vocational
- B71 Teaching Major: Mathematics

EDUCATION
- Expected date of graduation: May 2007, University of Louisville, Louisville, KY
  Doctorate of Philosophy in Educational Leadership and Organizational Development,
  with an emphasis in Evaluation
- June 2001, University of Louisville, Louisville, KY
  Rank I Planned Program of 30 graduate hours, at least 15 hours are 600 level or
  Higher with an emphasis in Mathematics and Educational Statistics
- August 1988, University of Louisville, Louisville, KY
  Post-Baccalaureate undergraduate study for Mathematics concentration, 24 hours
- August 1983, University of Louisville, Louisville, KY
  Master of Education, 30 graduate hours, Secondary Education with an emphasis in
  Consumer Education
- June 1978, University of Kentucky, Lexington, KY
  Bachelor of Science, Vocational Home Economics Education
PROFESSIONAL EXPERIENCE

- September 1993 to present, Mathematics Instructor, Oldham County High School, Buckner, KY.
  o Teach full-time in the mathematics department
  o Assist in the creation of the School Improvement Plan as a member of the committee that oversees the two-year plan based on a needs assessment of the school

- January 1991 to October 1993, Homebound Instructor, Oldham County School System, Buckner, KY
  o Instructed students unable to attend school for medical reasons
  o Visited homebound students, Kindergarten through Grade 12 in their homes, twice a week for the duration of their illness
  o Number of students varied from one to ten per week
  o Retrieved assignments from each teacher of each homebound student
  o Responsible for grading and reporting of grades back to the home school

- August 1986 to January 1994, Part-time Instructor, Kentucky Community & Technical College System-Jefferson State (formerly known as Jefferson State Vocational School), Louisville, KY
  o Instructed adults in basic reading and mathematics skills in the learning center of the school
  o Individualized learning plans for each student
  o Administered the placement test as needed

- December 1980 to September 1984, Full-time Instructor, Kentucky Community & Technical College System-Jefferson State (formerly known as Jefferson State Vocational School), Louisville, KY
  o Instructed adults in Consumer and Family Life Skills
  o Trained adult students in job skills including applications and resumes
  o Individualized learning plans to meet the needs of faster learners

PROFESSIONAL AFFILIATIONS

Current memberships held in:
- Golden Key
- AERA (American Educational Research Association)
- KAPE (Kentucky Association of Professional Educators)
- NCTM (National Council of Teachers of Mathematics)
- KCTM (Kentucky Council of Teachers of Mathematics)
- Kentucky Chapter of the ASA (American Statistical Association)

AWARDS AND GRANTS

- Teacher of the Year, 2000-2001, Oldham County High School
- Pyramid Award, $1000 grant, from the Oldham County Education Foundation (OCEF), 2002/2003, "The Nature of Mathematical Modeling in a Technological World," grant to allow technology purchase for connection between mathematical functions and real-world quantities
• Pyramid Award, $500 grant, from the Oldham County Education Foundation (OCEF), 1999/2000, “Doing the Sine Wave,” grant to allow technology purchase for music lesson in Algebra 2 class
• Pyramid Award, $500 grant, from the Oldham County Education Foundation (OCEF), 1998/1999, “The Seeds of Technology,” grant to allow students in need to purchase graphing calculators on an installment plan
• Pyramid Award, $500 grant, from the Oldham County Education Foundation (OCEF), 1995/1996, grant to purchase data gathering equipment
• Texize National Award for Excellence in Consumer Education, 1982, $1000 cash prize, student-developed consumer pamphlets

PRESENTATIONS
• Presented the INFINITY Project to Tech Alliance in 2004. Tech Alliance is a professional organization of mathematics and science teachers in the Louisville area that meet twice a year to learn about new technology for the classroom.
• Selected as an INFINITY Project instructor for a 40 hour training session for new instructors, held at SMU in Dallas, Texas in the summers of 2002 and 2003.
• Selected for an educational video published by VideoJournal in 2002. I was filmed teaching a mathematics lesson using the multiple intelligences.
• Presented Multiple Intelligence activities as part of a team of 5 teachers to two different Kentucky high schools in the summer of 2001.
• Co-presented with another high school teacher and a university professor at a session of KCTM in October of 2000 on “Incorporating the History of Mathematics.”

DISSERTATION
Effects Of Evaluative Feedback On Math Self-Efficacy, Grade Self-Efficacy, And Math Achievement Of Ninth Grade Algebra Students: A Longitudinal Approach

SERVICE
• I currently Coordinate the Calculator Installment Loan Program-coordinate application process, procurement and dissemination of calculators to students in need, and secure payments from students.
• Served as the Math Portfolio Cluster Leader for nine years-responsible for all math portfolios for the entire student body at Oldham County High School. I trained the math teachers and tracked the portfolios for graduation requirements.
• Served as assistant Cheerleading Coach for 7 years-responsible for the budget and fund-raising activities.
• In past years, I have served on the following committees at Oldham County High School:
  o The New American High School Application Committee-one of 7 teachers and administrators that prepared a 55 page document. I was responsible for the ten pages of statistical information and also for compiling the binders for the site visit.
- **Title IX Committee** - I served as one of six teachers and administrators. We met twice a year to discuss athletic concerns and issue.

- **Rewards Committee** - I was selected both times our school qualified for rewards money to serve on this committee. The committee decides on shares and portions of the reward amount for each position.

- **Assessment Committee** - I was one of six teachers asked to serve on this committee to look at ways we could improve our test scores. As a result of this work, I was asked to write the mathematics review activities for the school.