

University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

12-2012

Predictive relationships of teacher efficacy, geometry knowledge for teaching, and the cognitive levels of teacher practice on student achievement.

Paul Klein
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>

Recommended Citation

Klein, Paul, "Predictive relationships of teacher efficacy, geometry knowledge for teaching, and the cognitive levels of teacher practice on student achievement." (2012). *Electronic Theses and Dissertations*. Paper 763.
<https://doi.org/10.18297/etd/763>

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.

PREDICTIVE RELATIONSHIPS OF TEACHER EFFICACY,
GEOMETRY KNOWLEDGE FOR TEACHING,
AND THE COGNITIVE LEVELS OF TEACHER PRACTICE
ON
STUDENT ACHIEVEMENT

By

Paul Klein
University of Louisville

A Dissertation
Submitted to the Faculty of the
College of Education and Human Development
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

December 2012

Copyright 2012 by Paul Joseph Klein

All rights reserved

PREDICTIVE RELATIONSHIPS OF TEACHER EFFICACY,
GEOMETRY KNOWLEDGE FOR TEACHING,
AND THE COGNITIVE LEVELS OF TEACHER PRACTICE
ON
STUDENT ACHIEVEMENT

By

Paul Joseph Klein
M.Ed., Xavier University, 2002
M. Eng., University of Louisville, 1990

A Dissertation Approved on

November 26, 2012

By the following Dissertation Committee:

William Bush, Dissertation Chair

Robert Ronau

Susan Peters

Chad Buckendahl

Molly Sullivan

ACKNOWLEDGEMENTS

This journey was only made possible through the help of others. From the beginning of my time as a research assistant, I have felt lucky to have Dr. William Bush as my chair and mentor. He is not only a respected and talented professional in the mathematics education arena, but is a man of great generosity, personal integrity, and kindness. I was so lucky to have somehow landed under his guidance and I am grateful. Dr. Robert Ronau has also been an invaluable source of wisdom and encouragement during this journey. His expertise in research design and attention to detail continues to show me what it means to be a high quality researcher, and his friendship has been a nice support along the way. Dr. Susan Peters is a real powerhouse. Her analytical ability in finding the heart of an argument or the true meaning of a sentence or phrase never ceases to impress me. Yet she still remains so human and approachable. I don't mind repeating what I have said many times before, that she is a true asset to this university, and I have learned much from her during these last months of writing. Between his frequent flights to and from Las Vegas, Dr. Chad Buckendahl of Alpine Testing Solutions added his psychometric expertise always in a constructive and professional way. Thanks Chad for giving me your time, professional advice, and friendship. Finally, Dr. Molly Sullivan entered the scene when I needed an extra help. It was nice to have comments coming from the real-world, administration side of things, especially for me, as I tend to live in the clouds sometimes. Thanks Molly for stepping in when needed and giving me support throughout. To close, I want to thank my two friends Dr. Jane Jones and Patty Grimes. It

was to them that I would take my personal doubts, fears, and worries and always found words of encouragement and kindness, making it somehow easier to take that next step. Many thanks to all of you for getting me to the end.

ABSTRACT

PREDICTIVE RELATIONSHIPS OF TEACHER EFFICACY, GEOMETRY KNOWLEDGE FOR TEACHING, AND THE COGNITIVE LEVELS OF TEACHER PRACTICE ON STUDENT ACHIEVEMENT

Paul Joseph Klein

November 26, 2012

This study explored the predictive relationships of teacher efficacy, teacher knowledge, and teacher practices with student achievement. More specifically, secondary mathematics teachers' efficacy beliefs, geometry knowledge for teaching, and the cognitive complexity of the teachers' classroom practices were examined for 72 teachers in both urban and rural districts across Kentucky, along with the student achievement data of their students. Teacher and student data were obtained from the NSF-funded Geometry Assessment for Secondary Teachers (GAST) project, which administered geometry teacher knowledge assessments at the beginning and end of the school year, and collected cognitive complexity data from lessons through three classroom observations. Student achievement was measured using a modified geometry end-of-course assessment with a geometry readiness test as a covariate. Teacher efficacy data was obtained from the same teachers through an online survey at the end of the GAST project. Correlation, multiple regression, and hierarchical linear modeling techniques were used to analyze the data. Results revealed that the cognitive level of

teacher practices significantly predicted student achievement. This finding provides support for increasing teacher awareness of the importance of high cognitive instruction by helping them recognize the essential features of classroom activities that provide this instruction and assisting them to plan and implement high cognitive tasks in their classrooms.

LIST OF TABLES

Table 1. Study Design and Research Questions	61
Table 2. Descriptive Statistics of Teacher Variable Measurements	81
Table 3. Descriptive Statistics of the EGT and EOCA Variables	86
Table 4. Teacher Level Bivariate Correlations.....	88
Table 5. Ordinary Least Squares Regression Results.....	89
Table 6. Multiple Regression Model Summaries	93
Table 7. Multiple Regression Coefficient Statistics	93
Table 8. Results from One-Way ANOVA, Unconditional Model, HLM	95
Table 9. HLM Model Statistics with Teacher Practice as an Intercept Variable.....	97
Table 10. HLM Model Statistics with Teacher Practice as EGT Slope Variable.....	99
Table 11. Final Model, TP and Yrs_Exp as Intercept Variables	101
Table 12. HLM Model Variances.....	102

LIST OF FIGURES

Figure 1. Conceptual Framework	9
Figure 2. Conceptual Framework with Variables.....	10
Figure 3. Triadic Reciprocal Causation	17
Figure 4. Webb Efficacy Scale Item.....	27
Figure 5. Domains of Mathematical Knowledge for Teaching	40
Figure 6. Mathematical Tasks of Teaching.	43
Figure 7. Teacher Experience Distribution.....	63
Figure 8. GAST Assessment Distributions for Form A and Form B	82
Figure 9. Teacher Efficacy Distribution	83
Figure 10. Distribution of Teacher Practice Scores.....	85
Figure 11. Distribution of the Entering Geometry Test.....	86
Figure 12. Distribution of Student End of Course Assessment Scores	87
Figure 13. Histogram of Standardized Residuals & PP Plot.	91
Figure 14. Scatterplot of Residuals vs. Predicted Errors.	92

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iii
ABSTRACT	v
LIST OF TABLES.....	vii
LIST OF FIGURES	viii
CHAPTER 1	1
Current Trends in Mathematics Education	1
Student Achievement and Teacher Effectiveness	2
Teacher Efficacy	4
Mathematical Knowledge for Teaching.....	5
Teacher Practice.....	8
Conceptual Framework	8
CHAPTER 2	12
The Theoretical Development of Teacher-Efficacy.....	13
Self-Efficacy	13
Teacher Efficacy Framework.....	19
Teacher Efficacy Measures.....	25
Teacher Knowledge Framework.....	35
Teacher Practice and Cognitive Complexity	43
Student Achievement	48
CHAPTER 3	60
Research Design.....	60
Population and Sample.....	62
Sampling Procedures.....	63
Instrumentation	64
Limitations	74
CHAPTER 4	81
Descriptive Statistics.....	81
Correlation Analysis	87

Regression Analysis.....	88
HLM Analysis.....	93
CHAPTER 5	104
Discussion	104
Teacher Efficacy and Student Achievement.....	105
Geometry Knowledge for Teaching and Student Achievement	108
Teacher Practice and Student Achievement.....	110
Conclusion	114
REFERENCES	119
APPENDICES	130
CURRICULUM VITAE.....	158

CHAPTER 1

STUDY RATIONALE AND RESEARCH QUESTIONS

Current Trends in Mathematics Education

For several years students in the United States have consistently scored lower on mathematics achievement tests when compared to students from other industrialized countries. The latest assessment, PISA 2009 (Program for International Student Assessment), further confirms this poor showing. Of the sixty-five countries participating in the assessment, the US ranked thirty-second, and fewer than a third of US students scored at the proficient level in mathematics. The results of the 2007 TIMSS study (Trends in Mathematics and Science Study) also showed poor results when compared to Asian countries (very few European countries participated in the TIMSS study). Fourth-grade students from Hong Kong, Singapore, China, Japan, Russia, and England far surpassed US fourth graders in mathematics, and similar results were true for eighth-grade students.

Consequences of this trend could be tragic, both for the students and for the nation. More states require four years of mathematics or particular mathematics courses as required subjects, and universities are setting higher and stricter admission standards for mathematics. Because of this, many college hopefuls will lose opportunities to attend a postsecondary institution simply because they lack the mathematical skills and knowledge necessary to advance to the next level of education. The effects on the financial future of the US are also potentially devastating. A recent report predicted that

the US could lose \$75 trillion over the next 80 years if the current trend in mathematics proficiency continues (Peterson, Woessmann, Hanushek, & Lastra-Anadon, 2011).

These findings further suggest that what most Americans have realized since the publication of *A Nation at Risk* (Gardner, 1983) is true—American education is in dire need of reform, particularly in mathematics. Clearly a problem exists; the multi-million dollar question is “What is the solution?”

Student Achievement and Teacher Effectiveness

With the publication of *Equality of Educational Opportunity*, Coleman and colleagues (1966) found that student achievement was determined more by socioeconomic background than by school characteristics. Much subsequent research substantiated these findings (Sirin, 2005; White, 1982). Coleman’s study led the way to the development of busing and student assignment plans designed to help integrate schools and reduce the achievement gap for low-income and minority students. However, the Coleman Report also revealed that, among various school resources (which districts and policy makers can control more directly), teacher quality, as measured by teacher’s educational background, verbal abilities, level of parent education, and other teacher variables had significant correlations to student achievement. Except for peer effects determined by the composition of the student body, teacher quality explained more student achievement variance than any other school characteristic. Sanders (2004) also found that teacher effects explained substantial amounts of student achievement variance. In a longitudinal analysis spanning twenty-two years, Sanders partitioned variability in student achievement into three categories: among districts, among schools within districts, and among teachers within schools within districts. His findings revealed

that districts accounted for 5% of the variance; schools within districts accounted for about 30% of the variance; and teachers within schools within districts accounted for about 65% of the variance. Sanders concluded, "Differences in teacher effectiveness is the dominant factor affecting student academic progress. This is true in all subjects but is pronounced in Math" (p. 2). Repeatedly, the competence of the teacher has been found to be a major factor in boosting student achievement. Darling-Hammond (2007) noted:

A long line of studies has established that the single most important school influence on student learning is the quality of the teacher. Students lucky enough to have teachers who know their content and how to teach it well achieve substantially more. And the effects of a very good (or very poor) teacher last beyond a single year, influencing their students' learning for years to come. Indeed, expert teachers are the most fundamental resource for improving education. (p. 67)

Nevertheless, throughout the history of academic research the important characteristics that define exactly what makes a good teacher remain hard to measure (Goldhaber, 2002; Goldhaber, Brewer, & Anderson, 1999). Researchers found that easily measured teacher variables such as education level, experience, and certification status explained only 3% of the variance in student achievement, if the effect was detectable at all (Boyd, Goldhaber, Lankford, & Wyckoff, 2007; Goldhaber, 2002). Especially today, when federal and state budgets are affected by a slumping economy, administrators need to effectively and accurately identify quality teachers from those who are not. Following the demands of high stakes accountability, exceptional teachers should be rewarded appropriately for teaching excellence and helping others improve their teaching.

Researchers must strive to further understand how to measure the salient qualities of effective instruction and identify those aspects of teacher education programs that are essential for developing competent teachers.

Teacher Efficacy

In the late seventies, the application of Julian Rotter's locus of control theory (1966) and Albert Bandura's self-efficacy theory (1977) to educational research led to the identification of teacher efficacy, an important teacher attribute that appeared to have a positive relationship to student achievement. The first major studies on teacher efficacy were conducted by the RAND Corporation (Berman, McLaughlin, Bass, Pauly, & Zellman, 1977). The results of the extensive study showed a strong positive relationship between teacher efficacy and all dependent variables in the analysis, and in fact, the effects of teacher efficacy on outcomes were "among the strongest relationships identified in the analysis," making it "the most powerful teacher attribute" (McLaughlin & Marsh, 1978, p. 85). Dependent variables included student performance, percentage of goals achieved, amount of teacher change, and continued use of project methods and materials after the project ended. Other teacher-level variables such as years of experience and verbal ability also showed positive effects, although their association with the dependent variables was not as large. With such positive results, the RAND studies stimulated further research on the relationship between teacher efficacy and student achievement (Allinder, 1995; Ashton & Webb, 1986; Maguire, 2011; Moore & Esselman, 1992; Ross, 1992, 1994; S. Watson, 1991). Not only were the RAND conclusions supported, but teacher efficacy also was found to be correlated with many other areas such as: student motivation (Midgley, Feldlaufer, & Eccles, 1989); the effort

teachers expend in a teaching situation (Gibson & Dembo, 1984); teachers' willingness to place low SES students in regular classrooms (Meijer & Foster, 1988; Podell & Soodak, 1993; Soodak & Podell, 1993); teachers' level of professional commitment (Coladarci, 1992); teacher career retention (Glickman & Tamashiro, 1982) teacher willingness to try innovation (Allinder, 1994; Guskey & Passaro, 1994; Smylie, 1988) teacher valuing innovations (Cousins & Walker, 2000); classroom management skills (Ashton, Webb, & Doda, 1982; Woolfolk & Hoy, 1990); and low teacher stress (Greenwood, 1990; Parkay, Greenwood, Olejnik, & Proller, 1988). Clearly, teacher efficacy has been found to be an important teacher characteristic that is highly associated with successful student performance and improved classroom environments.

Unfortunately, recent studies directly relating teacher efficacy to student achievement have been scarce. In a review of teacher efficacy research from 1998 to 2009, Klassen and colleagues (2011) suggested, "Establishing a stronger research base that provides evidence for links between teachers' self-efficacy and student outcomes are needed, especially at the classroom level where the influence of teacher characteristics plays a critical role in influencing student achievement" (p.40). Palardy and Rumberg (2008) also agreed, "as a class of teacher effects, teacher attitudes and beliefs about their ability to teach and students' ability to learn have been under-examined in the literature" (p. 114). The current study proposes to answer this call for additional research by analyzing teacher efficacy and student achievement relationships.

Mathematical Knowledge for Teaching

Though teacher efficacy appears to have a strong association with teacher effectiveness, it cannot fully account for a teacher's influence on student achievement on

its own. Raudenbush, Rowan, and Cheong (1992) suggested that teacher efficacy is only part of the solution. They wrote,

“...positive feelings of [teacher] self-efficacy are necessary (but not sufficient) for effective teaching. Following Bandura we argue that positive feelings of self-efficacy produce a “generative capability” that enables teachers to construct new teaching strategies and increase their levels of effort in the face of difficult and uncertain teaching circumstances. From this perspective, feelings of positive self-efficacy cannot guarantee effective teaching, since teachers with high levels of perceived self-efficacy may lack the requisite knowledge or skills to be effective.” (p. 151)

Here Raudenbush and colleagues implied that requisite knowledge is also needed for effective teaching. The question then follows: “What knowledge is requisite?” This question has received much attention in mathematics education research. For centuries, mathematics knowledge has been categorized into distinct disciplines and content areas, but the notion of *mathematical knowledge for teaching* has only recently been proposed (Ball, Lubienski, & Mewborn, 2001). The applied mathematics that engineers use differs from the more theoretical and abstract mathematics studied by pure mathematicians. But what mathematics content should a mathematics *teacher* know? Mathematical knowledge for teaching reaches beyond what has been traditionally taught in mathematics and education programs, though the scope of that specific knowledge is difficult to define. Shulman’s work (1986) on pedagogical content knowledge introduced a major advancement in classifying teacher knowledge, but the most significant advances

regarding a teacher's knowledge of mathematics has come from a team of researchers at the University of Michigan. Deborah Ball and her colleagues applied Shulman's ideas of teacher knowledge to *mathematical knowledge for teaching* (Ball et al., 2001; Ball et al., 2008; Hill, Rowan, & Ball, 2005). The framework they introduced has become fundamental in classifying characteristics of mathematical knowledge for teaching and has provided mathematics education programs guidelines for exploring subsequent curriculum changes. Though research measuring mathematical knowledge for teaching and its effects on student achievement is just beginning (see Hill et al., 2005), studies have shown that mathematical teacher knowledge contributes to student success (Baumert et al., 2010; Hill et al., 2005). More research that measures teacher knowledge directly is warranted to further explore these relationships.

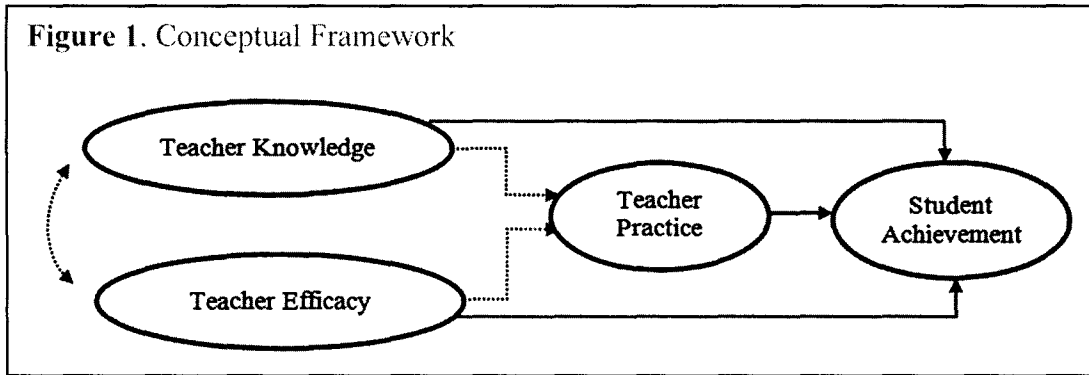
In addition, following Raudenbush's idea that teacher efficacy is a necessary but not sufficient condition for effective teaching, exploring the relationship between a teacher's mathematical knowledge for teaching and teacher efficacy as they relate to student achievement is a logical next step. Previous researchers have explored the extent to which teacher knowledge is related to efficacy (Benz, Bradley, Alderman, & Flowers, 1992; Campbell, 1996; Hoy & Woolfolk, 1993), but these investigations were embedded in larger questions and did not focus on the efficacy-knowledge-achievement relationship. Moreover, none of them looked specifically at mathematical knowledge for teaching when analyzing teacher knowledge.

Teacher Practice

Though teacher efficacy and teacher knowledge seem necessary prerequisites for good teaching, neither of them would likely have much of an impact on student achievement if students were not exposed to high-cognitive learning possibilities. The QUASAR project (Silver & Stein, 1996) revealed that the cognitive demand of classroom tasks was a key element of a teacher's instruction related to student learning. Students who were in classrooms where tasks with high-cognitive demand were set up and implemented showed the highest achievement on the mathematics assessment (Stein & Lane, 1996; Stein, Lane, & Silver, 1996). However, although a teacher may plan high-cognitive lessons, what unfolds during actual instructional time does not always reflect a high-cognitive level (Stein, Grover, & Henningsen, 1996). This research revealed that teachers' actions in the classroom can trump their pedagogical and content knowledge. Teacher practices, consisting of high-cognitive instruction, demonstrations, and tasks give students opportunities to learn to think at higher levels and are key aspects of effective teaching.

Conceptual Framework

This study added to existing research by exploring the extent to which teacher efficacy, teacher knowledge, and the cognitive level of a teacher's instruction (labeled teacher practice) predicted student mathematics achievement. The conceptual framework for the study is exemplified by Figure 1 below.

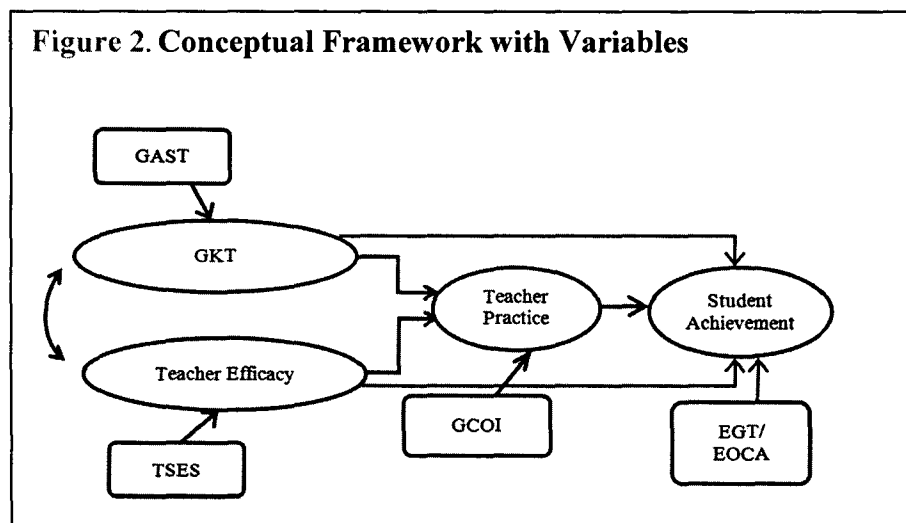


Secondary geometry teachers and their classes from both urban and rural Kentucky high schools participated in the study during the two academic years 2010-2012. The researcher measured the constructs of teacher efficacy, geometry knowledge for teaching (GKT), and the cognitive complexity of teacher practice for each teacher in the study using instruments that will be described below. Then using correlation analysis, ordinary least squares regression, multiple regression, and Hierarchical Linear Modeling the researcher explored the predictive nature of these constructs on the student achievement scores. The relationships between the teacher-level variables were also analyzed. The researcher hypothesized that the three variables would predict student performance and be positively correlated with each other.

Data for much of the study was obtained from the National Science Foundation funded Geometry Assessment for Secondary Teachers (GAST) project (grant no. GB080891). For the GAST study, both public and private high school mathematics teachers participated during a one-year high school geometry course. The teacher knowledge construct was measured using the Geometry Assessment for Secondary Teachers (GAST), administered to teachers at the beginning and end of the school year. GAST is a teacher assessment that was specifically designed to capture the geometry knowledge necessary for effective teaching. Although this instrument was not strictly

designed using Ball and colleagues' mathematical knowledge for teaching framework, many of the items were representative of this framework, and both content knowledge and pedagogical content knowledge categories were assessed. To measure student geometry achievement, the researcher used an adapted version of a Kentucky End-of-Course Geometry Assessment (EOCA). The test was designed to assess secondary geometry knowledge at a deeper level than simple recall and procedure repetition, and was reduced in content to match the specific content areas of the GAST test. To compare students across classes, an Entering Geometry Test (EGT) was administered at the beginning of the year and used as a covariate of student achievement. This assessment was product of researchers at the University of Chicago and was designed to measure geometry knowledge of incoming geometry students (Usiskin, 1982). Teacher efficacy, the only variable not obtained from GAST project data, was measured using the Teacher Sense of Efficacy Scale (TSES) developed by Tschannen-Moran and Hoy (2001). This instrument was administered on-line to participating GAST teachers.

The final conceptual framework including measurement variables is shown below.



Applying correlation analysis, regression analyses, and Hierarchical Linear

Modeling with these data, the study sought to answer the following questions:

1. Are teacher efficacy, geometry knowledge for teaching, and the cognitive levels of teacher practice correlated?
2. Does teacher efficacy or geometry knowledge for teaching predict the cognitive levels of teacher practice?
3. Does teacher efficacy, geometry knowledge for teaching, or the cognitive levels of teacher practice predict student achievement?

CHAPTER 2

RESEARCH ON TEACHER EFFICACY, MATHEMATICAL KNOWLEDGE FOR TEACHING, AND COGNITIVE LEVELS OF TEACHER PRACTICES

Introduction

This chapter describes the concepts *teacher efficacy*, *mathematical knowledge for teaching*, *teacher practices* and *student achievement* along with the research exploring these constructs. The first part of the chapter focuses on teacher efficacy, which has roots in two theories of motivation grounded in Social Cognitive Psychology: Julian Rotter's *locus of control* (1954, 1966) and Albert Bandura's *self-efficacy* (1977, 1986, 1997). Rotter's theory centered on an individual's belief that desired outcomes were either due to his or her own actions or to external factors. Bandura's self-efficacy focused on an individual's beliefs about his or her own *capabilities* as a major influence in human action. Both of these belief systems proposed that the person was the author of his or her own actions, rather than a by-product of the environment. The concept of teacher efficacy will be presented as an outgrowth of these two advances in modern psychology, followed by a brief exposition of the development of teacher efficacy measurements. The chapter then traces the development of mathematical knowledge for teaching, beginning first with the knowledge frameworks developed by Bloom, followed by Shulman's focus on the knowledge specific to teaching, and concluding with Ball and her colleagues' mathematical knowledge for teaching (MKT) framework. Finally, the teacher practice construct is presented from the perspective of teacher moves and tasks, and using Webb's

depth of knowledge framework as a basis to measure the cognitive complexity of classroom instruction. The chapter concludes with a brief exposition of the historical roots of student achievement research.

The Theoretical Development of Teacher-Efficacy

The primary basis for the theory of teacher efficacy is self-efficacy.

Self-Efficacy

The development of the self-efficacy construct is closely tied to the progression of human behavior and motivation theories of the 20th century (Pajares, 2002). At the core of self-efficacy theory is the assumption that the beliefs people hold to be true about themselves form the basis of their actions and vitally influence success or failure. This theory is based on the study of the self, which was a central point in William James' "The Principles of Psychology" (1890). However, from 1920 through 1940, the focus of psychology was behaviorist oriented (following the work of Watson, Pavlov, and Skinner). No longer was the internal self a principle of action; rather, actions were studied as observable responses to external stimuli. Watson (1930) wrote:

"Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select -- doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors" (p. 82).

Early studies on the inner life of an individual were criticized as "unscientific." Educational research followed this trend until Maslow (1954) introduced his theories of

motivation based on the human desire to fulfill certain needs, and a resurgence of the interest in the self occurred. Constructs such as self-esteem and self-concept became popular, and research on these constructs flourished during the 1960s and 1970s.

However, some efforts were misguided and the push to promote “high self-esteem” led to poor results in education. In 1982 Hansford and Hattie published a meta-analysis of 128 studies and found that correlations between self-concept and academic achievement did not yield consistent results. Still, the shift to cognitive processes of the individual led to the development of self-efficacy research. Though the theoretical framework of teacher efficacy is based heavily on Bandura’s self-efficacy theory, teacher efficacy research found its theoretical and historical roots in Julian Rotter’s theory of locus of control (1966).

Rotter’s Theory of Reinforcement and Locus of Control

During the 1950s, Julian Rotter challenged behaviorists’ theories by claiming that human behavior depended not so much upon external forces, but on judgments made about the outcomes of behavior. Reinforcement, reward, or gratification was generally accepted as a key element in the acquisition and performance of skills and knowledge. Thus, reinforcement served to strengthen the expectancy that a certain behavior would be followed by that reinforcement if the behavior was repeated. Rotter hypothesized that if the subject believed the reinforcement was due to, or contingent upon, his/her own actions or skills, the expectation that the reinforcement would follow the same behavior in the future would be stronger (1966). Rotter labeled this a belief in *internal control*. If on the other hand, the subject believed that the reinforcement was brought about by external forces (fate, luck, chance, or powerful others), the expectation that the

reinforcement would reoccur in the future was not as strong, representing a belief in *external control*. The subject's perception of locus of control then had a direct effect on motivation for future behavior. Rotter wrote:

“It is hypothesized that this variable [locus of control] is of major significance in understanding the nature of learning processes in different kinds of learning situations and also that consistent individual differences exist among individuals in the degree to which they are likely to attribute personal control to reward in the same situation” (p.1).

Rotter's research validated his theory and educational researchers began to take notice. This construct seemed to fit the K–12 academic environment, where a multitude of variables contributes to student learning, in particular the teacher. How much do teachers believe they have more influence on student outcomes than other factors, and does this belief affect how well students perform? Research has shown that teachers make a difference. Is this characteristic consistent and can it help explain effective teaching? Researchers began to study the relationships between teacher's internal/external locus of control and student motivation and achievement. Thus Rotter's social learning theory was used as a framework to study teacher belief systems and provided the foundation for the first studies of teacher efficacy.

Bandura's Self Efficacy

The other major theory underpinning teacher efficacy was Albert Bandura's concept of self-efficacy. During the second half of the century, Bandura developed and refined a social learning theory centered on cognitive processes which eventually included aspects of self-beliefs. His theory not only focused on the self, but more

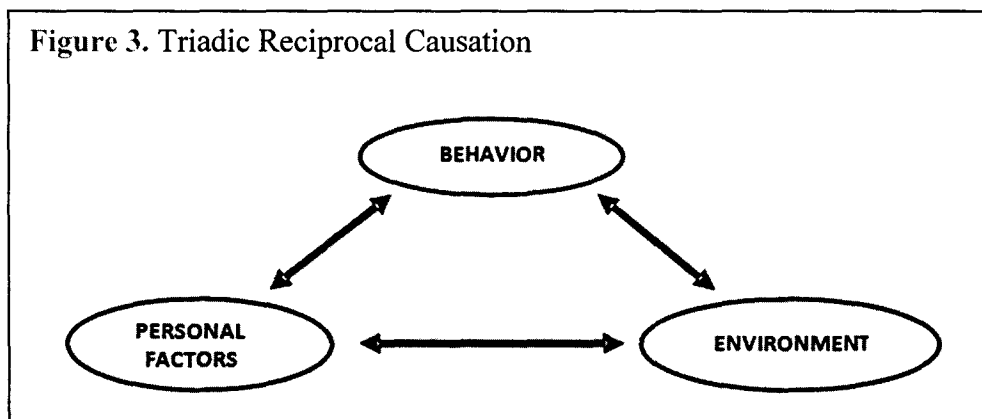
specifically on *beliefs of capabilities* (1977). Central to his theory was the proposition that human behavior can often be better predicted by the beliefs people held about their own capabilities than by what they were actually capable of accomplishing.

Self-efficacy can be defined as “the belief in one’s capabilities to organize and execute the courses of action required to manage prospective situations” (1995, p. 2). Individuals possess self-beliefs that enable them to exercise some control over their thoughts, feelings, and actions. In this sense, individuals are not simply passive objects upon which the environment acts and shapes; but through reflection, they are also authors of their own behavior and outcomes. Self-efficacy provides the foundation for human motivation, well-being, and personal accomplishment.

What specific effects does self-efficacy have on behavior? According to Bandura, self-efficacy influences the choices a person makes and the amount of effort one exerts to complete a task. It shapes how long one perseveres when obstacles present themselves, how one feels about the process, and how much resistance one puts forth against the fear of failure. Bandura’s theory states that self-efficacy is based, not on what is objectively true about capabilities, but on what is believed to be true. As such, the theory explains why some people behave confidently, though in reality they have very little experience to support this disposition, and at the same time, why very talented people are sometimes hindered by self-doubt (Pajares, 2002).

Bandura further developed his theory with a framework for human agency he called *triadic reciprocal causation* (see Figure 3). Within this framework, behavior, personal factors (cognitive, affective, and biological events), and environmental factors all interact bidirectionally, though not necessarily at the same strength. For example,

people's thoughts and feelings are affected by the environment in which they find themselves, but this environment is at the same time influenced by the characteristics of each person. (A classroom environment, for example, can be affected by the personal characteristics of a boisterous and aggressive child.) Also the structures and norms of a society influence human behavior, while at the same time human actions create and modify these same social structures. Finally, a person's thoughts, beliefs and expectations help determine behavior, yet the nature and external effects of this behavior influence the person's thoughts and emotions. This exchange of influence is not instantaneous, but involves a time lag between events. Self-efficacy resides in the personal factors that help determine future behavior.



Sources of Self-Efficacy

Four distinct sources contribute to a person's self-efficacy beliefs: mastery experience, vicarious experience, verbal persuasion, and physiological states (Bandura, 1997). Mastery experience represents the most influential source. Nothing builds a person's beliefs about his/her own capabilities more than successfully completing the task at hand. While successes build a firm belief in personal efficacy, failures are

extremely detrimental, especially if they occur before high self-efficacy has been established. This is why, for example, the first experiences of pre-service teachers are particularly important. An early experience of failure can be devastating.

Vicarious experience also enhances or diminishes self-efficacy beliefs. Observing others succeed after sustained effort can be especially influential if persons identify similarities between themselves and the acting subject. Once this connection is established, the observer forms the conviction “I can do it too.” If, however, the persons cannot identify with the acting subject, self-efficacy beliefs can remain generally unaffected.

A third source of efficacy beliefs is verbal persuasion. People who are verbally persuaded that they have the capabilities to perform a task exert a more sustained effort than those who have self doubt. But verbal persuasion alone is often not sufficient to build efficacy beliefs. Unrealistic expectancies can quickly be negated by the results of a poorly performed task, and placing persons prematurely in a role for which they are not prepared for can also have adverse effects. Positive verbal persuasion must be followed by structured opportunities for persons to succeed and reinforce their beliefs.

Finally, physiological states such as stress, anxiety, and mood swings can strengthen or weaken efficacy beliefs. The physiological state itself does not affect the change in beliefs; rather, the persons’ interpretations of that state as being supportive or detrimental cause the change. One person might use high energy and nervousness as positive facilitators of action, while another may see them as debilitating. In Bandura’s self-efficacy theory, the role of cognition and interpretation of these four sources is central in determining behavior.

Several characteristics help distinguish self-efficacy from self-concept/esteem (Gist & Mitchell, 1992). Self-efficacy is context specific; it is an assessment of competence to perform a specific task in a particular domain. In contrast, self-concept/esteem is an integrated view of the self, stemming from judgments of the self across various dimensions. Gist and Mitchell give the example of a rocket scientist who has a very low self-efficacy pertaining to dancing. Upon reflection, the scientist decides that this is acceptable, and considering other personal characteristics, he does not diminish his self-evaluation. Self-efficacy is also action oriented. It answers the question “Can I do this?” Self-concept/esteem constructs, on the other hand, are more being/feeling-oriented and answer the questions “Who am I?” or “How do I feel?” These two distinctions are central to any self-efficacy analysis. Contrasting elements of self-concept and self-efficacy are included in Appendix A.

Teacher Efficacy Framework

The following section describes the development of the self-efficacy construct as applied to educational contexts, specifically with regard to how these ideas appear in research on teachers. Teacher efficacy was first conceptualized and tested in Rotter’s locus of control framework. Bandura’s self-efficacy theories were then applied to further develop the theoretical framework. This progression will be described below, followed by an investigation of the prominent teacher efficacy instruments.

Development of Teacher Efficacy Research

The construct of teacher efficacy was first introduced, analyzed, and advanced in educational research within Rotter’s social learning theory framework. Berman, McLaughlin, Bass, Pauly, and Zelman (1977) defined the construct as “a belief that the

teacher can help even the most difficult or unmotivated students” and also “the extent to which the teacher believes he or she has the capacity to affect student performance” (pp. 136-137). In their study, student motivation and achievement were reinforcements for teacher instruction and behavior. As members of the RAND research team, the researchers measured the extent to which teachers believed they could influence these outcomes. Teachers who were confident their behavior would positively affect student motivation and achievement exhibited an internal locus of control (high “sense of efficacy”) whereas teachers who believed that environmental factors such as school and district characteristics, student family structures, or socio-economic status overpowered their own efforts exhibited external locus of control (low “sense of efficacy”).

Essentially, researchers asked “How does a teacher’s locus of control belief structure affect student outcomes?” To answer this question, the RAND team used the responses of over 1000 teachers to two items on a self-report survey:

1. When it comes right down to it, a teacher really can’t do much [because] most of a student’s motivation and performance depends on his or her home environment.
2. If I really try hard, I can get through to even the most difficult or unmotivated students.

Responses to these two items were combined to form a single measure labeled “teacher sense of efficacy.” The results of the study showed a strong positive relationship among teacher efficacy and all of the outcome variables in the analysis. Berman and colleagues wrote, “teacher sense of efficacy is positively related to the percent of project goals achieved, the amount of teacher change, improved student performance, and continuation of both project methods and materials. Teacher’s attitudes about their own professional competence, in short, appeared to have major effects on what happens to projects and

how effective they are” (p. 137). Thus teacher efficacy was found to be strongly associated with improved student achievement and other important teacher-level variables.

The findings of the RAND study led to additional studies of teacher efficacy. Ashton, Doda, and Webb (1982) incorporated Bandura’s self-efficacy concept to give theoretical support to the construct. These researchers further specified the distinction between the two RAND items. The first RAND item referred to the *general relationship* between student learning and environmental factors. A teacher may be convinced of his or her own capabilities, but because of external influences such as the skills of the particular student, the school where the teacher works, SES, race, gender, special needs of the class, etc., the teacher’s sense of efficacy is lessened. For example, if a teacher believes that high student achievement is due solely to student ability, the teacher will have a low sense of teaching efficacy. If, on the other hand, the teacher is convinced of Bloom’s theory (1978) that all students have the same learning ability when given the appropriate learning conditions, the teacher will have a high sense of teaching efficacy. Teacher efficacy as interpreted from this perspective has become known as *general teaching efficacy* (GTE) (Tschannen-Moran, Hoy, & Hoy, 1998, p. 204).

The second RAND item “If I really try hard, I can get through to even the most difficult or unmotivated students” related directly to a teacher’s *personal teaching efficacy* (PTE), or “the teacher’s general sense of effectiveness as a teacher” (Ashton et al., 1982, p. 92). Teachers with high PTE generally had confidence in their abilities as teachers and were satisfied with their training and knowledge of the subject. This helped them overcome obstacles to effective teaching and learning. The distinction between

GTE and PTE and the incorporation of Bandura's framework to teacher efficacy led to further studies with efforts to discover meaningful relationships of both dimensions of the construct to many different school related variables.

Teacher Efficacy, Student Achievement, and other With-in Teacher Effects

Subsequent research continued to support teacher efficacy as a significant contributor to many aspects of teaching and learning. After the RAND study, teacher efficacy was linked to improved student achievement (Anderson, Greene, & Loewen, 1988; Ashton & Webb, 1986; Moore & Esselman, 1992; Ross, 1992; S. Watson, 1991) and increased student motivation and sense of efficacy (Anderson et al., 1988; Henson, 2002; Midgley et al., 1989). Teacher efficacy was also shown to have positive correlations with teacher-level variables. Teachers with high efficacy were more willing to implement instructional change (Ashton & Webb, 1986; Guskey, 1988; Haney, Wang, Keil, & Zoffel, 2007), work longer with struggling students (Gibson & Dembo, 1984), and were less inclined to refer difficult students to special education (Meijer & Foster, 1988; Podell & Soodak, 1993; Soodak & Podell, 1993). Teachers were also more likely to stay longer in the teaching profession (Burley, Hall, Villeme, & Brockmeier, 1991; Glickman & Tamashiro, 1982) and use classroom management techniques that stimulate student autonomy (Woolfolk, Rosoff, & Hoy, 1990). Furthermore, Muijs and Reynolds (2002) showed that teacher efficacy affected specific teacher behaviors which in turn led to increased student achievement. Thus teacher efficacy ultimately was linked to student achievement measured directly or through within-teacher variables, all of which enhance student performance. This finding was true especially during the first 20 years of teacher efficacy research, 1978-1998.

However, relatively few studies in the last several years have attempted to reinforce the research findings between teacher efficacy and student achievement. In a review of 218 studies from the twelve year period 1998-2009, Klassen, Tze, Betts, and Gordon (2011) reported that only two studies (0.09%) examined links between teacher efficacy and student outcomes, and only four studies concentrated on teachers' collective efficacy and student outcomes. In the first of the two, Ross, Hogaboam-Gray, and Hannay (2001) found that teacher efficacy variables explained 7% - 9% of the variance of student outcomes (in this case, a student's computer skills and computer self-efficacy), though correlations were relatively small. In the second study, Caprara, Barbranelli, Steca, and Malone (2006) investigated over 2000 Italian teachers and found a significant but small effect of teacher efficacy on student outcomes one year later. A reciprocal effect was also discovered: student achievement tended to boost teacher efficacy beliefs.

More recently Bruce, Esmonde, Ross, Dookie, and Beatty (2010) examined the link between teacher efficacy and professional learning and development experiences. They found that teacher efficacy had little effect on student outcomes unless connected to teacher professional learning opportunities, which together led to improved student achievement. However, if teachers had inflated teacher efficacy beliefs (that did not reflect actual skills), they would not seek additional strategies to improve student performance, and no higher scores resulted. Another recent study by Dale, Phillips, and Sianjina (2011) sought to show the mediating effects of teacher efficacy on instructional and transformational leadership when analyzed with mathematics achievement. The sample included 177 elementary teachers from six Maryland counties. Although teacher

efficacy did not have a mediating effect on the two variables, the study suggested a significant direct impact on student mathematics achievement.

School Context Effects and Teacher-Efficacy

According to Bandura, personal factors such as self-efficacy can be influenced by environmental factors. For a teacher, these environmental factors include the school organizational structure and climate, principal leadership, and even student subjects. Moore and Esselman (1992) found that teachers who perceived a positive school atmosphere and a strong academic press in the school staff scored higher on Gibson and Dembo's Teacher Efficacy Scale (1984). Principal leadership has also been linked to strong teacher efficacy. Schools with a high academic emphasis and a principal who has strong influence and supports the teaching staff were found to be significantly related to teacher efficacy, though effect sizes were relatively small (Hoy & Woolfolk, 1993). The sense of community in the school also had influence. Using High School and Beyond data, Lee, Dedrick and Smith (1991) found that the school's sense of community was a significant predictor of teacher efficacy. (Sense of community was measured by a composite of items that capture how much teachers can count on the staff for help, felt like they were in a big family, and shared the beliefs, values, and mission of the school). Though these results may not be surprising, many teachers still do not receive adequate support, nor do they feel closely connected to the school's community life.

In addition to these contextual factors, students and classes also may affect teacher efficacy levels. Raudenbush and colleagues (1992) used a single efficacy item, similar to the RAND 2 item, to study teacher efficacy variations across classes during a single day. Results showed that efficacy can vary within teachers depending on the

classes they were teaching and the particular group of students with whom they were working. Teachers tended to feel most efficacious when teaching high-track classes.

Teacher Efficacy Measures

After the discovery of strong associations between teacher efficacy and student achievement, as demonstrated by the two items on the RAND survey, researchers designed more comprehensive measures of teacher efficacy (Tschannen-Moran et al., 1998). Initially, researchers used Rotter's Internal-External Locus of Control (I-E) Scale to measure teacher locus of control relationships with student achievement (Murray & Staebler, 1974). Rose and Medway, however, argued that the I-E scale was never intended to measure the specific dynamics of classroom teaching or student-teacher interactions, and may lead to lower correlations. In response, they developed the 28-item Teacher Locus of Control (TLC) instrument (1981) specifically designed to measure elementary school teachers' perceptions of control in the classroom. Half of the items described positive or successful situations, and half described negative or failure situations. Teachers were asked to attribute the result to an explanation representing either an internal or external locus of control.

Results of the TLC scale were more predictive than the Rotter I-E scale of teachers' willingness to implement new instructional techniques and was overall a better predictor of behavior (Rose & Medway, 1981). For example, teachers with an internal locus of control in low SES schools gave fewer disciplinary commands to students ($r = -.68, p < .02$). However, TLC scores were weakly correlated to the RAND items individually as well as the cumulative RAND TE score (Coladarci, 1992). In addition, Greenwood, Olejnick, and Parkay (1990) studied the relationship between locus of

control as measured by the TLC and RAND items. The four possible combinations for the two RAND items were grouped into four categories (for example, one category represented a positive response pattern to both questions “I can...; teachers can...”). Teachers who responded with Pattern 3 (I can; teachers can) were significantly more internally-oriented in their beliefs about student success and failure than teachers of Pattern 1 (I can’t; teachers can’t). Pattern 3 teachers also displayed less stress than Pattern 1 teachers. Thus, when teachers had more confidence in their abilities, as well as in the abilities of other teachers, they tended to show less stress carrying out their classroom responsibilities.

As Rose and Medway developed the TLC scale, Guskey (1981) and Ashton, Doda, and Webb (1982) developed other instruments measuring teacher efficacy. The first was Guskey’s Responsibility for Student Achievement (RSA) scale. This 30-item survey presented a prompt describing a positive or negative student achievement experience, followed by two choices for the teacher. One of the choices represented an event caused by the teacher, the other an event caused by something outside the teacher’s control. However, because teachers often viewed classroom events as stemming from more than a single cause, the “either-or” choice format was found lacking (p. 44-45). Instead, Guskey asked teachers to assign a weight to each choice, with the total weight between them totaling 100 points or 100 percent.

The RSA was designed to measure how much teachers assumed responsibility for student outcomes. Two subscales represented assumed responsibility for student success (R+) and for student failure (R-). After administering the survey to 215 elementary and secondary school teachers and performing factor analyses, Guskey found the factors R+

and R- explained 60.9% of the variance in item responses. He then compared scores from the RSA to the RAND items (1988, pp. 8-9) and found significant, positive correlations between R+ and R- and the RAND items ($r = 0.42$, and $r = 0.43$) and no significant correlations between the factors themselves. These findings indicated that teachers' beliefs about affecting positive results were not on the same continuum as their beliefs about preventing negative ones. This result was also found by Guskey in a previous study which investigated contextual influences on teacher efficacy (1987, p. 45).

The other teacher efficacy instrument developed during this time was the Webb Efficacy Scale. Ashton, Doda, and Webb (1982) sought to build upon the RAND instrument by enhancing theoretical support and adding more items to increase reliability. They also wanted to reduce social desirability bias caused by the natural tendency for an "ego-defensive" response to survey questions. In other words, teachers may rate themselves according to how their response supports their image as a teacher instead of honestly answering the question. To avoid this problem, the researchers designed a forced-choice format with items matched for social desirability. An example item is shown below in Figure 4.

Figure 4. Webb Efficacy Scale Item

A. A teacher should not be expected to reach every child; some students are not going to make academic progress.

B. Every child is reachable. It is a teacher's obligation to see to it that every child makes academic progress.

Circle one:

1. I agree most strongly with A
2. I agree most strongly with B

Results revealed that two of the Webb items representing Personal Teaching Efficacy significantly correlated with the second RAND item, but no large correlations with RAND item 1 were found. Also, few inter-correlations among the items themselves were found, revealing weak reliability. The authors admitted they failed to find an internally consistent measure of teacher efficacy (p. 96).

Using another approach, a group of researchers (Ashton, Buhr, & Crocker, 1984; 1982) introduced a series of situational vignettes to capture teacher efficacy. After reading a scenario of a typical classroom situation, teachers were asked to rate their own effectiveness in dealing with the situation. In this way, the researchers attempted to add specific contexts to the teachers' beliefs about their capabilities. Teachers in the 1982 study ranked themselves on a self-effectiveness scale from 1 (extremely ineffective) to 7 (extremely effective). The 1984 study included a "norm-referenced" scale by adding a comparison to other teachers, e.g. "much less effective than other teachers." Responses to the norm-referenced items yielded significant correlations to the RAND items, but responses to the self-effectiveness items did not. Both sets of data produced high internal consistencies ($\alpha = .95$ for the self-effectiveness scale and $\alpha = .94$ for the norm-referenced scale). Benze, Bradley, Alderman and Flowers (1992) used the Ashton vignettes and found that preservice teachers and college faculty were more optimistic about their effectiveness in dealing with student socialization and motivation than were in-service K - 12 teachers.

A few years after the Aston vignettes were developed, Gibson and Dembo (1984) created a more extensive instrument, the Teacher Efficacy Scale (TES), which eventually became the primary tool in teacher efficacy research. In a recent review of teacher

efficacy studies, Klassen, Tze, Betts, and Gordon (2011, p. 36) noted that almost one third of the 218 empirical studies examined in their review used the TES or a variation of the instrument. In developing the TES, Gibson and Dembo began with the RAND items and applied Bandura's social cognitive ideas as a theoretical foundation. The first item aligned with outcome expectancy, which reflected "the degree to which teachers believed that the environment could be controlled, that is, the extent to which students can be taught, given such factors as family background, IQ, and school conditions" (Gibson & Dembo, 1984, p. 570). The second item corresponded to Bandura's self-efficacy construct. The authors stated, "One would predict that teachers who believe student learning can be influenced by effective teaching, and who also have confidence in their own teaching abilities, should persist longer, provide a greater academic focus in the classroom, and exhibit different types of feedback than teachers who have lower expectations concerning their ability to influence student learning" (p.570).

To develop the TES, Gibson and Dembo reviewed relevant studies and interviewed classroom teachers to construct a preliminary pool of 53 sample items. These items were then administered to 90 classroom teachers as a pilot. Based on the results of a principal factor analysis, the item count was reduced to 30, forming the Teacher Efficacy Scale. To further pilot the instrument, Gibson and Dembo gave the TES to 208 elementary school teachers, along with 55 teachers enrolled in graduate education courses. Factor analysis revealed two substantial factors. Factor 1 (accounting for 18.2 % of the variance) represented a teacher's sense of Personal Teaching Efficacy and aligned with RAND item 2 (Cronbach's $\alpha = .78$). Factor 2, the General Teaching Efficacy factor (accounting for 10.6% of the variance, $\alpha = .75$), represented the belief that

any teacher's ability to bring about change is significantly limited by factors external to the teacher. Gibson and Dembo claimed this factor clearly aligned with Bandura's outcome expectancy dimension, and corresponded to RAND item 1. Because only 16 of the items loaded heavily on these two factors, Gibson and Dembo recommended further research on the scale, limiting the items to these 16. Soodak and Powell (1993) and Hoy and Woolfolk (1990, 1993) followed up on this suggestion and found that one of the GTE items loaded unexpectedly on the PTE factor and another item seemed not to belong to either factor. Hoy and Woolfolk (1993) further reduced the scale to 10 items, which was found just as reliable as the original version in their study, although these findings have not always been consistent. Continuing research using the TES has confirmed the existence of the two factors found by Gibson and Dembo, which in general have had weak inter-correlations and appear not to be related (Anderson et al., 1988; Burley et al., 1991; Hoy & Woolfolk, 1993; Moore & Esselman, 1992; Soodak & Podell, 1993).

Eclectic Measures

Other researchers used items across different instruments to produce measures of teacher efficacy. Midgley, Feldlaufer, and Eccles (1989) created a personal teaching efficacy instrument that revealed large significant differences in personal efficacy between middle school and high school mathematics teachers. This instrument was built using the RAND personal efficacy item, two items from the academic futility measurement (Brookover et al., 1978), one item from the Webb Scale, and one original item (Cronbach's alpha, $\alpha = .65$). Also, two groups of researchers (Lee et al., 1991; Newmann, Rutter, & Smith, 1989) selected two items from the High School and Beyond data to measure teacher efficacy: "To what extent do you feel successful in providing the

kind of education you would like to provide for most of your students?" and "I sometimes feel it is a waste of time to try to do my best as a teacher." These items were combined with two satisfaction items: "I usually look forward to each working day at this school" and "How much of the time do you feel satisfied with your job in this school?" to measure teacher efficacy. Tschannen-Moran et al. (1998, p. 218) criticized this combination. Though the measurements were highly correlated, the constructs were clearly different conceptually and should not have been confounded by combining the items representing them. Finally, Raudenbush and colleagues (1992) used a single item measured on a 4-point Likert scale to represent teacher efficacy, "To what extent do you feel successful in providing the kind of education you would like to provide for this class?" The researchers found that high school teachers showed higher self-efficacy when they taught high-track, honors classes.

Critiques of Teacher Efficacy Measures

Development of valid and reliable instruments to measure the teacher efficacy since the RAND items has not been without challenges. According to Henson (2002), two major issues have emerged. The first issue is embedded in the theoretical nature of the construct as defined by Bandura and is a perennial one. Since the nature of self-efficacy belief is necessarily linked to the specific context in which an action is performed, a true measure of the construct should also be tied to the specific context and task. This creates a problem when generalizing the results. If the measurement is too context specific, the conclusions drawn from the study will be severely limited to the specific task or context incorporated in the research. Pajares wrote, "specificity and precision are often purchased at the expense of external validity and practical

relevance''(1996, p. 561). On the other hand, if the measurement instrument is created to apply to a broader range of circumstances, the validity of the instrument in measuring the teacher's self-efficacy beliefs might be weakened. Again Pajares noted:

...self-efficacy beliefs should be assessed at the optimal level of specificity that corresponds to the criterial task being assessed and the domain of functioning being analyzed. This caution [made by Bandura (1986)] has often gone unheeded in educational research, which has resulted in self-efficacy assessments that reflect global or generalized attitudes about capabilities bearing slight or no resemblance to the criteria task with which they are compared...Omnibus tests that aim to assess general self-efficacy provide global scores that decontextualize the self-efficacy-behavior correspondence and transform self-efficacy beliefs into a generalized personality trait rather than the context-specific judgments Bandura suggests they are (p. 547).

The results of these measurements, according to Bandura, lack predictive relevance because it is difficult to judge what exactly is being assessed. The agent (which is the teacher in this case) must make a judgment about an imagined task or set of tasks not specifically referenced by the instrument.

Riggs and Enochs (1989) attempted to create a more domain specific instrument, the Science Teaching Efficacy Belief Instrument (STEBI) based on the TES (Riggs & Enochs, 1990). For this instrument, all items from the TES were modified to reflect an elementary science classroom setting. The STEBI incorporated two subscales: the Personal Science Teaching Efficacy Belief Scale and the Science Teaching Outcome

Expectancy Scale. Additional items were also added to control for acquiescence responding, which occurs when survey respondents simply answer agreeably to the statement, regardless of the question content (Mueller, 1986). Enochs, Huinker, and Smith (2000) further modified the STEBI to measure mathematics specific teacher efficacy. Items in the Mathematics Teaching Efficacy Belief Instrument (MTEBI) followed the phrasing of the STEBI except questions were oriented toward mathematics. “Even if I try very hard, I will not teach *mathematics* as well as I will most subjects (emphasis added).” Here the word science was simply replaced by mathematics. The MTBEI also had subscales for personal teaching efficacy (thirteen items) and outcome expectancy (eight items). These instruments were designed for elementary teachers only. It would make no sense to have this type of domain specific instrument for secondary teachers, as all of their teaching is domain specific, and rewording would not add specificity.

Henson’s second measurement concern involved the construct validity of the primary instruments developed during the first phase of teacher efficacy research. As stated previously, the roots of teacher efficacy studies stem from the two RAND items, which have theoretical underpinnings in Rotter’s locus of control theory. But Bandura’s self-efficacy theories published just before teacher efficacy research blossomed, became the primary basis for teacher efficacy constructs, especially with its adoption by Gibson and Dembo in the development of the TES (1984). Both theories then are intermingled in the development of the teacher efficacy construct. Upon closer examination of the TES items, Guskey and Pasaro found an “interesting anomaly” (1994, p. 630). The PTE items in the instrument were all positively worded (“I can”), and reflected an internal locus of

control, whereas the GTE items were all negatively worded ("Teachers can't") and reflected an external locus of control. Was the instrument measuring PTE and GTE or internal/external locus of control? Guskey and Pasaro reworded the items by making half of the items positive and the other half negative for each construct. Interestingly, results did not reflect the dimensions of PTE and GTE but instead showed signs of a dichotomous internal/external locus of control construct. Principal component factor analysis results showed two factors; yet, the loading of the PTE and GTE items on each factor appeared random. By revising the positive/negative structure of the items, the loading on the factors was essentially changed even though the content of the items remained the same. Reworded items revealed factors corresponding to the internal/external locus of control.

“The teachers we surveyed, both prospective and experienced, did not distinguish between their personal ability to affect students and the potential influence of teachers in general. Rather, the distinctions they drew related to beliefs about the influence they and all teachers have, or do not have, on the learning of students, even those who may be considered difficult or unmotivated. Whether the item referent was "my influence" or "teachers' influence" made no difference” (page. 637).

However, the internal/external locus of control construct is typically thought of as bipolar, with each dimension on opposite sides of the continuum. Guskey and Pasaro did not find this to be the case because the data revealed two independent and distinct factors. The authors believed the strength associated with each factor depended on the teachers'

beliefs of the strength of their personal influence on the internal factor versus the strength of external influences that lie outside the classroom.

At this point, all survey measurements on this topic have flaws with regard to measuring teacher efficacy. However a relatively recent instrument has been well received. After an extensive review of teacher efficacy research, Klassen and others (2011, p. 40) recommended the Teacher Sense of Efficacy Scale (TSES) created by Tschannen-Moran in collaboration with Barr and Woolfolk Hoy (Tschannen-Moran & Barr, 2004; Tschannen-Moran & Hoy, 2001) which “shows considerably more congruence with self-efficacy theories than many of the other measures” among the hundreds of studies they reviewed. The development team of the TSES had extensive teaching experience and methodically created items to reflect essential aspects of teaching that surface in a teacher’s daily life. Items such as “How much can you do to adjust your lessons to the proper level for individual student?” and “How well can you keep a few problem students from ruining an entire lesson?” exemplify the specific focus of the instrument toward important teacher capabilities. The instrument has three subscales: Efficacy in Student Engagement, Efficacy in Instructional Practices, and Efficacy in Classroom Management. Two forms of the scale are available, a full 24-item form and a 12-item short form. The 24-item long form tends to have a slightly stronger factor structure with varying populations and will be incorporated in this study. TSES items are listed in Appendix E.

Teacher Knowledge Framework

Most university programs are designed to equip graduates with the knowledge necessary for their future career. How ironic is it that, only in the last 25 years, scholars

have begun to analyze the specific knowledge that is needed for teachers? This section describes the landscape of current mathematical knowledge for teaching structures proposed by leading researchers in the field. A brief introduction to content knowledge frameworks by Bloom will first be discussed, followed by a more detailed explanation of knowledge specific to teaching as proposed by Lee S. Shulman. Finally the structures of mathematical knowledge for teaching by Deborah Ball and her colleagues will be shared. These frameworks provide the context for using the Geometry Assessment for Secondary Teachers (GAST) scores as meaningful measures of mathematics teacher knowledge.

Classifying knowledge is not a new enterprise. Ancient and medieval philosophers such as Aristotle and Roger Bacon wrote monumental works on the classification of knowledge. But without question, the classification of knowledge related to educational research builds upon the fundamental work by Benjamin S. Bloom entitled, "Taxonomy of Educational Objectives; the Classification of Educational Goals (1956). In the 1950s, Bloom organized a team of researchers to define learning objectives for students to help educators design and develop assessments and curricula. The team identified three domains of educational activities: cognitive, affective, and psychomotor. Under the cognitive domain, six levels of knowledge were delineated according to the difficulty that students would encounter either answering questions orally or on written assessments. These levels, listed from least difficult to most difficult, include: 1. *Knowledge* -recall previously learned information; 2. *Comprehension* – understand the meaning of material; 3. *Application* – use a concept in a new situation, 4. *Analysis* – separate concepts into component parts; 5. *Synthesis* – build a structure or pattern from diverse elements; 6. *Evaluation* – make judgments about the value of ideas.

Using this framework, the level of difficulty of a task or item can be classified and tests/curricula can be developed according to the needs of the student and intent of the educator. Since its inception, Bloom's taxonomy has undergone revisions, albeit only minor ones (the fifth and sixth levels were switched, for example), but the work has been foundational. Subsequently, mathematics educators have developed mathematical content knowledge frameworks, finding levels of difficulty and understanding in mathematics much like Bloom did with knowledge in general (Hiebert, 1989; Hiebert & Carpenter, 1992; Krauss et al., 2008; Skemp, 1976, 1987). For example, Skemp (1976) classified the understanding of mathematics into relational understanding and instrumental understanding. Relational understanding is a deep, conceptual understanding of mathematics: "knowing both what to do and why" (p. 89), whereas instrumental understanding is hardly to be considered understanding at all. Skemp describes instrumental understanding as "rules without reasons" and gives examples such as "borrowing" and "take it over to the other side and change the sign." He argues that a "mathematical mismatch" occurs when students aim to understand mathematics instrumentally while the teacher's goal is for them to have relational understanding (or vice-versa). The difference between these two types of understanding is so great that Skemp claimed, "there are two effectively different subjects being taught under the same name 'mathematics' " (p. 91). Similarly, Hiebert and Carpenter divided mathematics into conceptual knowledge (understanding concepts and relationships of principles) and procedural knowledge (knowledge of rules, algorithms, and procedures) (Hiebert, 1986; Hiebert & Carpenter, 1992). These two categories are not entirely unrelated; both must be mastered by the student in order to obtain a thorough understanding of mathematics.

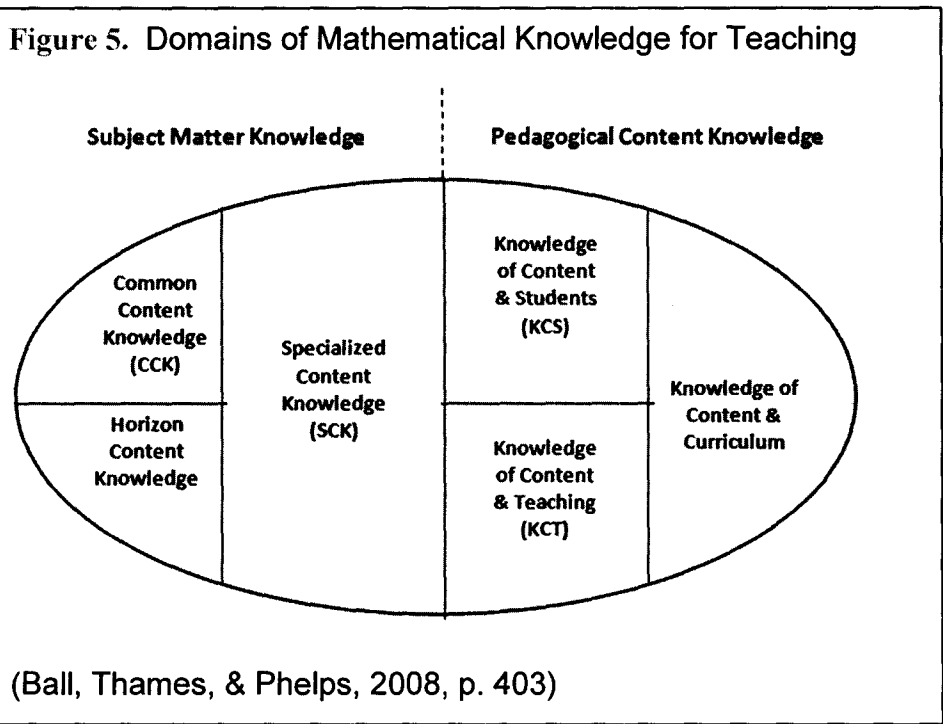
In this way knowledge classification was continued by content area experts to develop classifications systems in their own field.

Shulman's Pedagogical Content Knowledge

A second major contribution regarding *teacher* knowledge classification was introduced in 1985 by Lee Shulman during his Presidential Address at the AERA meeting in Chicago. In his talk, Shulman specifically addressed teacher knowledge as a content area in its own right. He defined *pedagogical content knowledge* (PCK) as “the ways of representing and formulating the subject that makes it comprehensible to others” (1986, p. 9). Note that PCK focuses on applying pedagogical principles to a specific subject domain. Teachers should not only master content knowledge but also know a variety of representations and alternate teaching methods accumulated either from research or from their own experience. This knowledge includes a deep understanding of the complexities that make the learning of specific lessons easy or difficult. Teachers should understand what learners bring to a lesson that facilitate or hinder the understanding of a concept, along with what strategies assist or deter student learning of specific content. In addition to pedagogical knowledge, Shulman's analysis included *content knowledge* (the concepts, facts and structures of the subject matter), as well as *curricular knowledge*, which encompasses all extraordinary methods and tools to teach the subject at hand. Curricular knowledge also includes a broader knowledge of content outside the teacher's subject area, the other subjects that students learn at the same time (lateral curriculum knowledge), and how this subject matter fits in the bigger picture of the complete education (1986, p. 10).

Mathematical Knowledge for Teaching

Since Shulman's insights were published, research on pedagogical content knowledge and teacher knowledge in general has been popular in educational research. In mathematics education, efforts concentrated on defining specific categories of mathematics teacher knowledge to revise curriculum and expectations for effective pre-service and inservice teacher programs. Researchers classified *mathematical knowledge for teaching* as the mathematics that teachers need to know that is different from what is required for other professions (Ball & Bass, 2000; Cuoco, 2001; Hill et al., 2005; Ma, 1999; Usiskin, 2001). Specifically, it is "the mathematical knowledge needed to carry out the work of teaching mathematics" (Ball et al., 2008, p. 395). Building on Shulman's content knowledge and PCK categories, Ball, Hill, and colleagues created a mathematical knowledge for teaching framework shown in Figure 5. This framework divides teacher knowledge into two major categories, *Subject Matter Knowledge* and *Pedagogical Content Knowledge*. Each of these categories is further divided into three distinct knowledge types described below:



Subject Matter Knowledge – corresponds to Shulmans’s content knowledge; the knowledge of the subject and its organizing structures. This knowledge is classified into:

a. *Common content knowledge (CCK)* – mathematical knowledge and skill used in settings other than teaching. Teachers must know this content as they must set-up and solve the problems they are assigning.

b. *Horizon content knowledge* – knowledge of the mathematics beyond the mathematics that one teaches, as well as the knowledge of the connection of these levels as contained in the mathematics curriculum. For example, first-grade teachers should understand how first-grade mathematics content relates to third-grade mathematics topics.

c. *Specialized Content Knowledge (SCK)* - mathematical knowledge and skill unique to teaching (and not typically useful in other

ways). This knowledge includes looking for patterns in student errors and evaluating whether a non-standard approach to a problem works.

Pedagogical Content Knowledge – the representations of specific content ideas, as well as an understanding of what makes the learning of a specific topic difficult or easy for students.

a. *Knowledge of content and students* (KCS) – the conceptions and misconceptions that students of different ages and backgrounds bring with them when learning the most frequently taught topics and lessons; combines knowledge about students and knowledge about mathematics.

b. *Knowledge of content and teaching* – the ways of representing and formulating the subject that make it comprehensible to others; combines knowledge about teaching and knowledge about mathematics.

c. *Knowledge of content and curriculum* – includes knowledge outside the teacher’s subject area and the connections with this knowledge.

Two sub-domains of Pedagogical Content Knowledge, *Knowledge of Content and Students* (KCS) and *Knowledge of Content and Teaching* (KCT), were empirically identified in the 2008 study. Furthermore, the addition of the specialized content knowledge is a particularly interesting category. Notice that this category is on the content side of the diagram--which indicates it does not relate to student background, learning, cognition, or motivation. In addition, although this knowledge is content related, it is specific to the teaching field, with little purpose or worth outside of teaching. This combination is particularly intriguing. Ball writes,

Perhaps most interesting to us has been evidence that teaching may require a specialized form of pure subject matter knowledge—“pure” because it is not mixed with knowledge of students or pedagogy and is thus distinct from the pedagogical content knowledge identified by Shulman and his colleagues and “specialized” because it is not needed or used in settings other than mathematics teaching. This uniqueness is what makes this content knowledge special (2008, p. 396).

Ball gave an example of this type of knowledge with the subtraction problem below:

$$\begin{array}{r} 307 \\ -168 \\ \hline 261 \end{array}$$

Here a person with minimal mathematics knowledge can detect the answer is incorrect.

However, to help the student learn from the error, a teacher must recognize that the student has simply subtracted the smaller number from the larger, no matter the order.

Using this knowledge, the teacher can give appropriate feedback. This type of knowledge reflects mathematical content, and yet, it would never be needed (or desirable) outside the realm of teaching. Figure 6. includes Ball’s list of teaching tasks which often demand SCK. Ball used empirical evidence obtained from analyzing an extensive set of classroom videotapes to show that this category of mathematics knowledge surfaces in the classroom and is an essential part of teacher knowledge.

Figure 6.

Mathematical Tasks of Teaching

- Presenting mathematical ideas**
 - Responding to students' "why" questions**
 - Finding an example to make a specific mathematical point**
 - Recognizing what is involved in using a particular representation**
 - Linking representations to underlying ideas and to other representations**
 - Connecting a topic being taught to topics from prior or future years**
 - Explaining mathematical goals and purposes to parents**
 - Appraising and adapting the mathematical content of textbooks**
 - Modifying tasks to be either easier or harder**
 - Evaluating the plausibility of students' claims (often quickly)**
 - Giving or evaluating mathematical explanations**
 - Choosing and developing useable definitions**
 - Using mathematical notation and language and critiquing its use**
 - Asking productive mathematical questions**
 - Selecting representations for particular purposes**
 - Inspecting equivalences**
-

Ball et al., 2008, p. 400

Teacher Practice and Cognitive Complexity

From the above analysis, teacher knowledge and teacher beliefs appear to be key teacher qualities that promote student learning. Ultimately, however, what actually happens in the classroom determines how much learning takes place. Teachers might have a deep understanding of mathematics and mathematics knowledge for teaching and believe they have the knowledge to teach effectively, but if they do not give students relevant and frequent opportunities to learn in the classroom, these constructs will have

little influence on student achievement. Repeating what is proposed in NCTM's *Professional Standards for Teaching Mathematics* (NCTM 1991) Stein, Smith, Henningsen, and Silver (2000) wrote "opportunities for student learning are not created simply by putting students into groups, by placing manipulatives in front of them, or by handing them a calculator. Rather, it is the level and kind of thinking in which students engage that determines what they will learn" (p. 11). If a teacher uses strategies to encourage higher mathematical thinking in the classroom, it seems reasonable that students will later exhibit this type of thinking on their own when solving problems or when completing a challenging project or assessment. On the other hand, if higher cognitive mathematical thinking is not encouraged and demanded during a lesson, it is not realistic to expect students to develop high levels of thinking, either in the classroom or later in the workplace. This relationship is precisely the one studied in the GAST project.

During a typical lesson, a teacher's presentation of mathematics is only part of the picture. On a practical level, mathematical tasks influence what students learn because students spend most of their time in the classroom doing these tasks in the classroom (Boston & Smith, 2009). Here again, not all mathematical tasks promote learning. The cognitive complexity of tasks affects student achievement. The QUASAR project (Silver & Stein, 1996) showed that students in classrooms in which tasks with high cognitive demand were set up and implemented showed the highest achievement on the mathematics assessment (Stein & Lane, 1996; Stein, Lane, et al., 1996). Hiebert and Wearne (1993) found that both instructional tasks and the nature of classroom discourse likely "influence learning by affecting the kinds of cognitive processes students engage

(p.422)” Students who were given opportunities to engage in reflective thought and self-expression performed better, and these thought processes were linked to the nature of instructional tasks and classroom discourse presented in the classroom.

The connection between high-level cognitive tasks and learning has been widely supported in educational research. Referring to recent reviews and meta-analyses summarizing research on learning and instruction, Baumert and colleagues (2010) wrote “although terminology differs, three components of instruction have emerged consistently as being crucial for initiating and sustaining insightful learning processes in mathematics lessons. These three components are as follows: cognitively challenging and well-structured learning opportunities; learning support through monitoring of the learning process, individual feedback, and adaptive instruction; and efficient classroom and time management” (p.146). Clearly, the cognitive levels of the activities presented to students in the classroom either support or decrease the effectiveness of the learning opportunities in which students engage.

Stein and colleagues (2000) developed a systematic way to differentiate the levels of cognitive demand for mathematical tasks in written form. Their guide can be used by teachers to select tasks based on the kinds of thinking that the tasks demand of the students rather than the superficial features often associated with cognitive complexity (such as the format of the task – e.g. “a word problem”, or that the problem was presented in a real-life context). The guide distinguishes between low level tasks, e.g. those focusing on *memorization*, and *procedures without connections*, and higher level tasks, e.g., those focusing on *procedures with connections* and *doing mathematics*. Memorization tasks involve reproducing formulas, rules, or definitions, while making no

reference to the meaning of the connections or concepts that underlie the fact being memorized. *Procedures without connections* are similar in that the underlying reasoning for the procedure is not explored; rather, the procedure is simply algorithmic. Students are taught how to perform the procedure without any mathematical understanding of why the procedure is appropriate or how or why it produces the correct result. For example, a student may be asked to convert the fraction $\frac{3}{8}$ to a decimal by performing the long division algorithm, with no reference to the concepts of a fraction or its decimal equivalent. On the other hand, asking a student to use a 10 x 10 grid to explain why $\frac{3}{5} = 0.60$ and allowing them to make connections to the procedure and give meaning to their work is an example of using *procedures with connections*. Finally, tasks classified as *doing mathematics* require complex and non-algorithmic thinking. Students explore mathematical problems and discover relationships without having a clear path to the solution. In the previous example, students would be given the task of exploring fractional and decimal representations of $\frac{3}{5}$ using a grids of different sizes; however, they would not be given the conversion procedure (at least not initially). Through their own reasoning, students would represent the shaded region as a percent, decimal, and a fraction. They must draw upon their knowledge of these concepts in ways they have not been previously exposed. However, the guide developed by Stein and her colleagues (1996) only classifies the complexity of written tasks, not tasks as they are delivered by the teacher. In fact, they found that only about one-third of the tasks that started out at a high level remained on this level as the lesson unfolded in the classroom. My proposed study, while differentiating cognitive complexity levels in ways similar to Stein, attempts to capture the cognitive complexity of lessons using Webb's Depth of Knowledge levels.

Cognitive Complexity and Webb's Depth of Knowledge Levels

In the 1990s, Norman Webb was charged with determining the extent to which different states' standards and assessments aligned, an essential attribute for an effective standards-based system (Webb, 1997, 1999, 2002). Webb developed a systematic process and the analytical tools for judging the alignment among standards and assessments. To do this, he chose four criteria for analysis: categorical concurrence, depth-of-knowledge consistency, range of knowledge correspondence, and balance of representation. Webb wrote, "Standards and assessments can be aligned not only on the category of content covered by each, but also on the basis of the complexity of knowledge required by each. *Depth-of-knowledge consistency between standards and assessment indicates alignment if what is elicited from students on the assessment is as demanding cognitively as what students are expected to know and do as stated in the standards [emphasis added]*" (Webb, 2002, p. 4). To interpret and assign depth-of-knowledge measures, Webb devised the following levels, briefly summarized below:

Level 1. Recall. Includes recall of a fact, information, or procedure.

Level 2. Skill/Concept. Includes use of concepts and mental processing beyond habitual response.

Level 3. Strategic Thinking. Requires abstraction, reasoning, and developing a strategic plan involving some complexity. Often times there may be more than one answer.

Level 4. Extended Thinking. Requires investigation and extended thinking to process multiple conditions of the problem and solution.

Using these levels, educators can determine the depth of knowledge involved in teaching, classroom and homework assignments, student tasks, textbook problems, and tests. For example, a test which includes mostly DOK 3 tasks has higher cognitive complexity than a test which includes mostly DOK 1 tasks. Applying Webb's Depth of Knowledge framework to the moves and tasks used during classroom instruction can give a quantitative value of the cognitive complexity of the lesson as a whole, and ultimately of the level of teaching that the teacher plans or executes.

Student Achievement

Effective methods to improve student achievement have long been sought after, yet, after decades of research and experimentation, questions and controversies remain about the factors that actually lead to achievement gains. The main impetus of this student achievement research began in 1966 with the publication of "Equality of Educational Opportunity," better known as the "Coleman Report." Commissioned by the US Department of Education, Coleman and his team reported to the President and Congress the reasons for the "lack of availability of equal educational opportunities for individuals by reason of race, color, religion, or national origin in public educational institutions" (Section 402 of the Civil Rights Act of 1964). Coleman measured educational opportunity using student scores on achievement tests. He wrote, "One way of assessing the educational opportunity offered by schools is to measure how well they perform this task [teaching reading, writing, calculating, and problem solving]. Standard achievement tests are available to measure these skills...what they measure are the skills which are among the most important in our society for getting a good job and moving up to a better one, and for full participation in an increasingly technical world" (p. 20).

Coleman found that student background and socioeconomic status related more strongly to student achievement than did school resources. The report also indicated that achievement of minority students was particularly affected by differences in school characteristics, such as school resources, libraries, characteristics of teachers, or composition of the student body. One of the most significant contributions of Coleman's work was a new focus on student achievement rather than years spent in school, as had been commonly used. Since this report was released, educational researchers have increased their focus on understanding the analysis of the report and giving further evidence for the various factors that improve student achievement.

What has the research found about resources that improve student achievement? Unfortunately research on the effects of school resources on student achievement, though extensive, is conflicting. For example, Greenwald, Hedges, and Laine (1996) performed a meta-analysis of 60 primary research studies on the effects of school resources on student achievement and found that resources such as per-pupil expenditure, school size, teacher education and salary, and teacher/pupil ratio were positively related to student outcomes. The researchers added, "while many would have hoped that increasing resources would be positively related to achievement, we did not expect that the synthesis of data from a wide variety of studies over a three-decade period would yield conclusions so uniform in direction and comparable in magnitude" (p. 385). At the same time, Hanushek (1997) performed a meta-analysis on close to 400 studies exploring the effects of school resources on student achievement, and his conclusions were decidedly different. His findings "have a simple interpretation: There is no strong or consistent relationship between school resources and student performance" (p. 148). Hanushek did

not argue that schools do not make a difference, only that the measurable resources commonly used in studies (as those referred to in Greenwald, Hedges, and Laine) are not related to student outcomes.

The Coleman report indicated that among school-related influences, teacher quality was significantly correlated to student achievement, supporting Hanushek's finding that teachers have a great influence on student success. One of Hanushek's studies found that teachers account for at least 7.5 percent of the total variation in student achievement, with strong support that the true contribution is even higher (Rivkin, Hanushek, & Kain, 2005). He also found that one year with a good teacher can move an average student (based on student achievement) from the 50th percentile to the 58th percentile, which in economic terms can add more than \$20,000 to the student's lifetime earnings (Hanushek, 2011). Other researchers agree that teachers matter. William Sanders, a leading expert in statistical analysis, also found evidence regarding the impact of teacher effects on student achievement. His 2004 study revealed that teachers accounted for 65% of the variance in student achievement, and teacher effects were the dominant factor influencing student academic progress. Darling-Hammond (2007), summing up many decades of research, stated "the single most important school influence on student learning is the quality of the teacher" (p. 67). Though controversy exists among researchers as to whether or not typical, measureable school resources matter, most believe that the quality and effectiveness of the teacher is a major factor influencing student academic achievement. The problem lies with identifying which teacher characteristics make a difference.

MKT and Student Achievement

Only recently have mathematical knowledge for teaching frameworks been clearly defined, and the work showing the effects of teacher knowledge on student achievement is just beginning. Since the early 1960s, two approaches of exploring teacher relationships to student achievement emerged (Hill et al., 2005). The first, labeled “process-product” studies, analyzed the effects of *teaching practices* on student achievement. These studies found positive relationships among certain teaching practices and student achievement gains, and the results have been well documented (see reviews by Brophy & Good (1986), and Gage (1978)). The research showed, for example, that student achievement improved when teachers emphasized educational objectives in establishing expectations and allocating time. Achievement also improved when teachers led students quickly through the curriculum while using small steps that allowed for high success rates, or adapted instructional materials based on their knowledge of the students’ characteristics (Brophy, 1986). But many studies relied excessively on correlation and little attention was given to the subject matter taught (Hill et al., 2005). The teaching practices that worked well for reading did not necessarily work well for mathematics.

The second approach, labeled “educational production function” research, linked student achievement gains to the resources possessed by students, teachers, and schools. Teacher preparation and experience, such as education level, certification status, and years of classroom experience, were used as predictors of student achievement. Analyses of these production function results have shown positive relationships between teacher education and student achievement. In the study mentioned above, Greenwald,

Hedges, and Laine (1996) concluded “resource variables that attempt to describe the quality of teachers (teacher ability, teacher education, and teacher experience) show very strong relations with student achievement” (p. 384). However, these claims have been disputed on methodological grounds with regard to sampling procedures and meta-analysis techniques (Hanushek, 1996). Other studies that measured teacher knowledge more directly through certification examinations or subject matter tests have also shown higher student achievement correlations (Boardman, Davis, & Sanday, 1977; Strauss & Sawyer, 1986; Tatto, Nielsen, Cummings, Kularatna, & Dharmadasa, 1993), but this research did not compare specific teacher subject matter knowledge with student gains in that specific subject matter. Finally, teacher preparation in science and mathematics has been shown to have a positive effect on student gains in these specific areas (Monk, 1994). This study used variables such as the number of undergraduate and graduate mathematics and mathematics education courses, earned mathematics major, and earned Master’s degree as measurements of mathematics and science preparation.

Studies that analyzed the relationship of teachers’ mathematics knowledge on students’ mathematical achievement gains have been few, and for the most part, are not generalizable. Mullens, Murnane, and Willet (1996) found positive effects of subject matter knowledge and teacher effectiveness, but the study was done in Belize where the teacher preparation and ability is assumed extremely variable. The same is true for the Harbison and Hanushek study (1992), which analyzed teachers in Brazil. In addition, the teacher assessments for these two studies, did not measure mathematical knowledge for teaching as described previously. The Mullens (1992) study used scores on a primary-school-leaving examination (given to all students) as the teacher knowledge measure, and

Harbison and others (1996) used the same fourth-grade mathematics assessment used to measure the students as the teacher measure.

Based more heavily on Shulman's conceptual framework, a study by Rowan, Chiang and Miller (1997) used teachers' scores on the mathematics quiz included in NELS:88 as a measure of teacher content knowledge. This measurement was essentially a multiple-choice response on only one item. Without further studies on the validity of this measurement, generalization of the results should be made with caution. Metzler and Woessmann (2010) found teachers' subject matter knowledge had a significant positive effect on student outcomes in mathematics and reading, but the study was also conducted in a third world country (Peru), again limiting generalizability.

A study by Hill, Rowan, and Ball (2005) used a measure specifically focused on the specialized mathematical knowledge and skills used in teaching mathematics. To capture the teacher knowledge construct, researchers designed items on a teacher questionnaire that was administered several times during the course of the three-year study. The variable was labeled CKT-M for content knowledge for teaching mathematics. Reading content items were also developed. Sample items are listed in Appendix B. The cumulative score on all questionnaire administrations were used to form a teacher's CKT-M score for the 115 first-grade and third-grade teachers participating in the study. Student achievement data were obtained from eight students in each classroom. Participating schools, however, did not constitute a nationally representative sample in two deliberate ways. First, selected schools were engaged in instructional improvement, and second, high-poverty and non-rural schools were overly represented, with a large percentage (68%) located in large and small cities. A control

group with similar characteristics was also included. Results of the study supported the importance of mathematical knowledge for teaching, as student achievement rose significantly by two points for every standard deviation in teachers' content knowledge –this translated to about one half to two thirds of a month of student growth in mathematics achievement (p. 396). When considering factors that affect student achievement in mathematics, mathematical knowledge for teaching seems to matter.

Another prominent study extending Shulman's work, is pertinent to the present discussion. Krauss and others (2008) identified three dimensions of PCK for mathematics: (1) knowledge of mathematical tasks as instructional tools; (2) knowledge of students' thinking and assessment of understanding; and (3) knowledge of multiple representations and explanations of mathematical problems. Applying these knowledge types, the researchers developed both a content knowledge assessment and a pedagogical content knowledge assessment. To measure the task dimension of PCK, teachers were asked to identify multiple solution paths for a problem. For the student dimension, teachers were assessed on their ability to recognize student misconceptions and difficulties. For the instructional dimension, teachers were given 10 vignettes of classroom situations and asked to provide possible learning solutions. Baumert and others (2010) used these data and student mathematics achievement data obtained from PISA 2009 for grade 9, and a separate mathematics test for grade 10, to analyze the effects of a teachers' content knowledge and PCK on student achievement. Results again were clear; the teacher PCK accounted for 39% of the variance in student achievement between classes in an HLM regression model. The study also found that PCK influenced the learning support dimensions of instructional quality (such as monitoring the learning

process, individual feedback, and adaptive instruction), but was independent of classroom management measures.

Teacher knowledge and student achievement relationships have also been analyzed according to *cognitive type*, which refers to “the kind of teacher content knowledge and thinking processes required to accomplish a task successfully” (Tchoshanov, 2011, p. 141). In this research, a teacher’s mathematical knowledge (content knowledge only) was divided into three categories: Type 1 – knowledge of facts and procedures; Type 2 knowledge of concepts and connections; and Type 3 knowledge of models and generalizations. Correlation analysis showed a significant relationship between student achievement and the Type 2 teacher knowledge, though Types 1 and 3 were not found to be correlated with student outcomes.

To close this section, a final caution about the process-product research mentioned above can help explain the limitations in solely analyzing teacher knowledge and student achievement. Brophy (1986) noted,

Even trained and experienced teachers vary widely in how well they organize the classroom, articulate achievement expectations and objectives, select and design academic tasks, and instruct their students. Those who do these things successfully produce significantly more achievement than those who do not, but doing them successfully demands a blend of energy, motivation, subject matter knowledge, and pedagogical skills that many teachers, let alone ordinary adults, do not possess. Thus, if we as a nation are serious about wanting good teaching in our schools, we will need not only to improve pay and working conditions so as to recruit and retain

talented people but also to arm them with the knowledge and skills they need to do their jobs effectively. Contemporary research on teaching is contributing by identifying the knowledge and skills that are needed.

(p. 1076)

The accumulated body of process-product literature shows that teacher knowledge is not the complete picture. What a teacher *does* with that knowledge is perhaps the deciding factor. Brophy added that if we are serious about wanting good teachers, we must “arm them with the knowledge and skills they need to do their jobs effectively.” This support includes arming them with the mathematical knowledge for teaching which has been proven to have a positive relationship with student achievement.

Teacher Efficacy, Teacher Knowledge, and Student Achievement

Research showing the combined effects of teacher efficacy and teacher knowledge on student outcomes is scarce. The few studies that involve both constructs offer only correlations between the two without study level analysis. In a comprehensive review, Fives (2003) found only nine studies that examined the relationship between teacher efficacy and teacher knowledge. Fives divided these studies into categories according to how the construct of teacher knowledge was measured. The first group used educational level as a proxy for teacher knowledge. In a study focusing on the organizational health of schools, Hoy and Woolfolk (1993) found a significant positive relationship between educational level and Personal Teacher Efficacy (PTE) ($r = 0.21, p < .01$). This construct was the only personal variable that had a positive relationship to PTE in a regression analysis. In 1996, Campbell (as cited in Fives, 2003, p. 83) also

found that educational level affected differences in PTE scores using three education categories: pre-Bachelor, Bachelor, and post-graduate levels. Results showed higher efficacy scores for teachers with post-graduate work than for the other two categories. The final study compared teacher efficacy across six groups: students, student teachers, students in advanced education courses, student teachers, practicing teachers, teacher education faculty, and non-faculty student teaching supervisors (Benz et al., 1992). The authors found that postsecondary teachers tended to show higher personal teaching efficacy than the other groups. The authors admit, however, that the generalizability of the result is minimal due to lack of rigorous controls (p. 284). Moreover, though college faculty may possess more mathematics content knowledge than K-12 teachers, postsecondary teaching situations are vastly different from K – 12 classrooms. Therefore, comparing efficacy scores across groups may not be an accurate method to compare knowledge and efficacy constructs. The results of the three studies above demonstrated a correlation between higher levels of education and higher levels of teacher efficacy.

Fives also reviewed studies that used completed coursework as a knowledge measure. Enochs, Scharmann, and Riggs (1995) used the number of high school and college science courses taken by elementary teachers to compare efficacy levels measured by the STEBI-B. Both high school and college courses were significantly correlated with teacher efficacy ($r = 0.22, p < .05$, and $r = -0.21, p < .05$, respectively). Negative correlations, as explained by Enoch and colleagues, were caused by poor science instruction models given by high school and post-secondary school teachers. Again, comparing teacher knowledge and teacher efficacy was only a tangential analysis

in this study and the studies mentioned above. None of these studies measured teacher knowledge directly, especially in ways that view teacher knowledge as defined and categorized in current educational research.

However, a few studies have incorporated more direct measures of teacher knowledge. In analyzing the effect of a mathematics methods course on teacher knowledge, Evans (2011) acquired teacher efficacy, attitudes towards mathematics, and mathematics content knowledge scores before and after the methods course. Evans claimed the mathematics content test measured “the combination of knowledge, skills, and understanding of mathematical concepts held by teachers” (p.228). This definition aligned with the types of knowledge outlined by Ball, on whose theoretical framework Evans grounded his research. Results showed that although the teachers’ content knowledge increased after the methods course, their self-efficacy beliefs did not. The somewhat high efficacy scores produced at the beginning of the course may have limited the possibility of the growth in efficacy.

Additional research measuring teacher knowledge more directly yielded more positive knowledge-efficacy correlations. Preservice science teachers who showed higher knowledge of alternative science conceptual knowledge scored higher on the Science Teaching Efficacy Belief Instrument (STEBI-B) (Schoon & Boone, 1998). Also, elementary teachers who demonstrated deeper knowledge of ADHD reported higher teacher efficacy beliefs (Sciutto, Terjesen, & Bender Frank, 2000). To measure teacher efficacy, participants rated the extent to which they felt they could effectively teach an ADHD student (on a 7-point scale). In both of these studies, specific content knowledge

was matched with efficacy measurements within the same content area. These studies give evidence of a positive correlation between the two constructs.

The research mentioned above shows that teacher knowledge, measured either indirectly using proxies such as education level and coursework, or more directly by content specific tests, tended to be positively correlated with teacher efficacy, though results are not always consistent. Post-secondary educators also showed higher efficacy beliefs than K-12 teachers. However, these studies did not incorporate data at the student level, which is necessary to show the influence of these variables on effective teaching.

CHAPTER 3

This chapter describes the research design, participants, variables, and measuring instruments used in this study. Instrument development, psychometrics, and reliability and validity issues are also presented. Finally, the limitations regarding research design and threats to validity are discussed.

Research Design

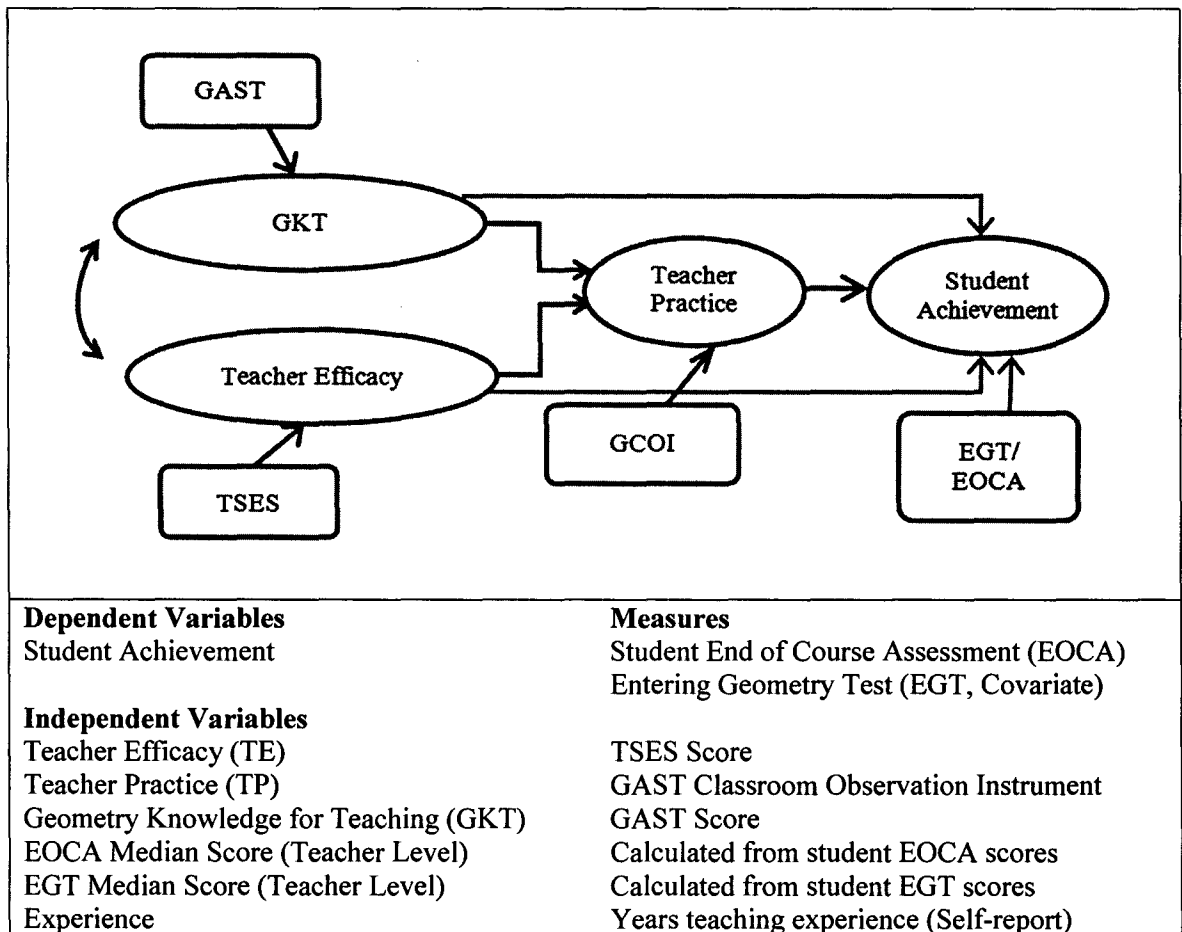
This study explored the possible predictive relationships of teacher efficacy (TE), geometry knowledge for teaching (GKT), and the cognitive complexity of teacher practices (TP) on student achievement. The relationships among the teacher-level variables were also analyzed. Table 1 below describes the basic research design of the study including variables, measures, and data analysis techniques. As reported by Klassen and colleagues (2011), 76% of research on teacher efficacy has used quantitative approaches. This study continued in that tradition. The teacher efficacy construct was measured with a 24-item self-report teacher efficacy survey—the Teacher Sense of Efficacy Scale (TSES). Geometry knowledge for teaching (GTK) was measured with the Geometry Assessment for Secondary Teachers (GAST), specifically designed to measure geometry teacher knowledge. Also incorporated from the GAST project was the GAST Classroom Observation Instrument (GCOI), used to capture the DOK level of teacher practices.

Given that the data had a two-level structure with students clustered in classrooms, Hierarchical Linear Modeling (HLM) was chosen as the principle method of

analysis. To give support and direction to the HLM analysis, correlation and regression analyses were performed to investigate relationships among variables, and the possible predictive relationships of teacher efficacy and geometry knowledge for teaching on teacher practices. For the HLM analysis, first-level student variables included a modified End-of-Course Assessment (EOCA) with an Entering Geometry Test (EGT) used as a covariate. As a later addition, the teacher's experience (number of years) was also added to the three principle teacher-level variables as a possible predictor of student achievement.

Table 1

Study Design and Research Questions



Research Question 1.

Are teacher efficacy, geometry knowledge for teaching, and the cognitive levels of teacher practice correlated?

Data Analysis. Bivariate Correlation.

Research Question 2.

Does teacher efficacy or geometry knowledge for teaching predict the cognitive levels of teacher practice?

Data Analysis: OLS Regression

$$c_1(\text{GKT}) + c_0 = \text{TP}$$

$$c_2(\text{TE}) + c_0 = \text{TP}$$

Research Question 3.

Does teacher efficacy, geometry knowledge for teaching, or the cognitive levels of teacher practice predict student achievement?

a. Correlation Analysis

- i. TE-EOCA (Median)
- ii. GKT-EOCA (Median)
- iii. TP-EOCA (Median)

b. Multiple Regression (Teacher Level)

$$\text{EOCA} = B_0 + B_1 (\text{TE}) + B_2 (\text{GKT}) + B_3 (\text{TP})$$

$$\text{EOCA} = B_0 + B_1 (\text{TE}) + B_2 (\text{GKT}) + B_3 (\text{TP}) + B_4 (\text{Med. EGT}) + B_5 (\text{YRS_EXP})$$

c. HLM Model

First Level

$$\text{EOCA}_{ij} = \beta_{0j} + \beta_{1j} * (\text{EGT}_{ij}) + r_{ij}$$

Second Level

$$\beta_{0j} = \gamma_{00} + \gamma_{01} * (\text{TE}_j) + \gamma_{02} * (\text{GKT}_j) + \gamma_{03} * (\text{TP}_j) + \gamma_{04} * (\text{YRS_EXP}_j) + u_{0j}$$

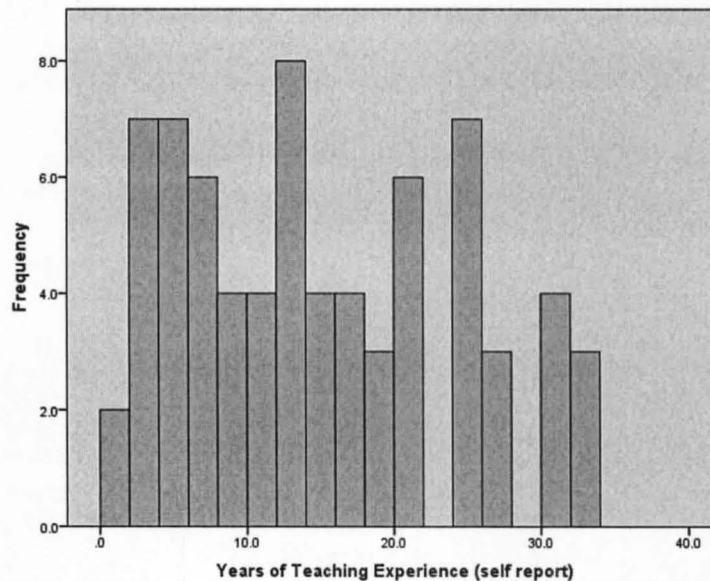
$$\beta_{1j} = \gamma_{10}$$

Population and Sample

The intended population for this study was high school geometry teachers, employed in either public or private high schools in Kentucky. The sample representing

this population was a convenience sample: the 72 teachers that participated in the NSF-funded Geometry Assessment for Secondary Teachers project. These teachers were employed in either urban or rural school districts across the state of Kentucky with a concentration in the Lexington and Louisville areas. The teachers' years of experience ranged from 1 to 33 years, with an average of 14.6 years and a standard deviation of 9.3 years. Fifty-four of the teachers had earned a masters degree, nine achieved Rank 1 (60 hours of graduate credit including a master's degree), and four had earned bachelor's degrees as their highest degree obtained. One teacher earned a Ph.D. in a field outside mathematics or secondary education. The distribution of teacher experience is shown in Figure 7.

Figure 7. Teacher Experience Distribution



Sampling Procedures

Participating teachers were solicited personally by GAST personnel at district professional development workshops or by telephone. Additional solicitations were

made through emails sent by district administrators, university faculty, and mathematics listservs. Eligible teachers must have taught high school geometry for at least one full year, have been scheduled to teach at least one high school geometry class of twenty or more students during the school year, and their students must have remained in the same teacher's classroom for the entire year. Forty-four teachers were accepted to participate in the 2010-2011 cohorts, and twenty-eight participated the following year, 2011-2012.

Instrumentation

Reliability and Validity of the Teacher Sense of Efficacy Scale

The Teacher Sense of Efficacy Scale (originally called the Ohio State Teacher Efficacy Scale, (Tschannen-Moran & Hoy, 2001)) was developed by a team of Ohio State teacher educators and practicing teacher doctoral students, all of whom had teaching experience ranging from 5 to 28 years. Drawing from their experience in the classroom, the members of the team designed the scale specifically to capture important and frequent tasks and capabilities integral to a teacher's work life. After several pilot tests and revisions, a 24-item survey emerged. Factor analysis revealed three subscales: efficacy for instructional strategies, efficacy for classroom management, and efficacy for student engagement. Internal consistency reliabilities for the subscales were high: $\alpha = .91$ for instruction, $\alpha = .90$ for management, and $\alpha = .87$ for engagement. Inter-correlations between these subscales were 0.60, 0.70, and 0.58 respectively ($p < .001$). When subjected to further factor analysis using only inservice teacher responses ($N = 255$), the three factors accounted for 54% of the variance of the efficacy score. In a confirmatory factor analysis, a principal axis factoring on a sample of 183 inservice teachers revealed one strong factor accounting for 75% of the variance. These results,

along with the moderate positive correlations of the three subscales, provided evidence that the instrument can be used to measure the underlying efficacy construct along with the three subscale scores. Loadings for each item on the efficacy factor ranged from .49 to .76, and from .50 to .78 on the subscale factors. Survey items, factor loadings, and subscale factor correlations can be found in Appendix E.

A shortened version of the TSES was also developed. An earlier 18-item version of the TSES produced a weak classroom management factor, so much so that Roberts and Henson (2001) recommended dropping this factor entirely. But convinced that classroom management constituted an essential element of teaching, the Ohio State team chose instead to introduce more items to further capture this important construct yielding the current 24-item survey. Because of the strong reliabilities of the three subscales, the four items with the highest loadings on each scale were selected to form a 12-item scale which showed nearly identical psychometric properties as the 24-item scale. However, the shortened survey did not show strong subscale factors when given to preservice teachers. Heneman, Kimball, and Milanowski (2006) attempted to further verify the psychometric properties of the TSES 12-item scale (and by extension the 24-item scale) by administering the short form to 1075 elementary, middle, and high school teachers. The factor analysis supported the three factor structure of the TSES and provided “strong replication of the Tschannen-Moran and Hoy results and extend those results to all three levels of teachers” (p.10). The high reliability was substantiated for the elementary, middle, and high school teacher data, and the psychometric properties reported by Tschannen-Moran were repeated, including the results of confirmatory and factor analyses, scale reliabilities, inter-correlations, means, and standard deviations.

To evaluate construct validity, the correlations between TSES measures and existing measures of teacher efficacy were analyzed by Tschannen-Moran and Hoy (2001). Survey respondents not only completed the long (and short) version of the TSES, but also the Rand Items and the Hoy and Woolfolk (1993) adaptation of the Gibson and Dembo Teacher Efficacy Scale (1984). Results are shown in Appendix E. Positive correlations were found between the TSES and Rand Item 2 and the Personal Teaching Efficacy (PTE) construct ($r = .53, p < .01$ and $r = 0.64, p < .01$) and also between Rand Item 1 and General Teaching Efficacy (GTE) ($r = .18, p < .01$ and $r = .16, p < .01$). The lower correlations with the GTE measure is not surprising, as other instruments have also shown lower correlations with this measurement, suggesting that the Gibson and Dembo GTE scale is “the least efficient in capturing the essence of teacher efficacy” (Tschannen-Moran & Hoy, 2001, p. 801). The authors also add,

Positive correlations with other measures of personal teaching efficacy provide evidence for construct validity. But the OSTES [now called TSES] moves beyond previous measures to capture a wider range of teaching tasks. Both the Rand and Gibson and Dembo instruments focused on coping with student difficulties and disruptions as well as overcoming the impediments posed by an unsupportive environment. Lacking were assessments of teaching in support of student thinking, effectiveness with capable students, creativity in teaching, and the flexible application of alternative assessment and teaching strategies (p. 801).

The TSES includes items that assess a broader range of teaching tasks in hopes of capturing the true essence of teacher efficacy. Klassen and others (2011) also supported

the use of the TSES due to its strong theoretical underpinnings regarding construct validity. These researchers claimed the TSES measure “shows considerably more congruence with self-efficacy theories than many of the other measures” among the hundreds of studies they reviewed (p. 40).

Reliability and Construct Validity of the Geometry Assessment for Secondary Teachers (GAST)

The development of the Geometry Assessment for Secondary Teachers (GAST) began in October of 2008 as part of an NSF grant designed to study high school geometry teacher’s geometry knowledge, teaching practices, and their relationship to student achievement. The test was designed to be predictive in nature— a high score on the GAST assessment was meant to signify the likelihood of student success in the high school classroom. As requested by the NSF, geometry content of the GAST assessment was restricted to four main areas: similarity, congruence, area, and volume. The intent of the request was to make the scope of the project less overwhelming by concentrating on these specific areas, which are so fundamental to geometry that results may still be generalized to represent high school geometry content in general. To enhance construct validity of the assessment, GAST team members first analyzed secondary and college geometry textbooks, state standards, and national assessments (e.g., NAEP, ACT) to better understand the current geometry used for teaching. Following this analysis, a team of mathematics educators, mathematicians, high school mathematics teachers, and doctoral students developed a blueprint for the assessment, incorporating ideas from Webb’s depth of knowledge framework. The blueprint of the assessment included three principal content areas: Teacher Knowledge of Mathematics (30%), Teacher Knowledge

of Geometric Reasoning and Problem Solving (25%), and Teacher Knowledge of Student Learning (45%). The test did not specifically measure all of the mathematical knowledge for teaching domains as categorized by Ball and colleagues, but the three major sub-domains—Common Content Knowledge, Knowledge of Content and Students, and Knowledge of Content and Teaching were well represented. Example items can be found in Appendix D. After item production, regional reviewers were assembled to validate items for section, domain, topic, item type, and DOK level. To further insure construct validity for the test, both the blueprint and test items were reviewed by a panel of national experts in geometry and mathematics teaching.

After pilot testing, two versions of the assessment were produced (Form A and Form B) which contained the same blueprint content and score points. Each version consisted of 11 open-response items (four 4-point items, and seven 2-point items), and 24 multiple choice items. Form A was administered to 44 teacher participants in Fall of 2010 and 28 in the Fall of 2011, and Form B was administered to the same participants in Spring of the same school years. The 2010-11 data from the administration of Form A produced an acceptable internal consistency reliability, with Cronbach's coefficient alpha of $\alpha = 0.74$, whereas Form B data produced questionable reliability, $\alpha = .63$. The lower reliability produced by the second set of data could be attributed to the urgencies of the end of the school year and its influences on teacher concentration and effort. Other than being used as a predictor of teacher effectiveness, the test may eventually be promoted as pre-post Professional Development assessment, or as a measurement to predict the possible success of preservice teachers. At this time, the predictive nature of the test is still under analysis and further test development and refinement is ongoing.

The GAST Classroom Observation Instrument

To evaluate the teaching practices used in the classroom, the GAST team incorporated a comprehensive mathematics teaching framework developed by Cooney, Davis, and Henderson (1975). This framework centers on teaching *moves*, which are “the kinds of logical things teachers do in teaching mathematics” (Davis, 1978, p. 13). These moves include defining concepts, giving examples and counterexamples, justifying formulas, demonstrating procedures—the typical building blocks of mathematics teaching. Teachers use moves to present or clarify mathematical concepts, generalizations, and skills. A complete glossary of moves for each of these three categories is listed in Appendix F, along with the DOK level assigned to each move. Using this framework, GAST observers coded teacher practices by categorizing each move on the GAST Classroom Observation Instrument (GCOI, See Appendix G). Classroom observations were scheduled to capture teacher practices when new content was introduced. For example, using the GCOI observers coded whether the teacher simply stated a definition, or whether the concept was further developed using such moves as counterexamples and comparisons with other concepts. As the classroom activity unfolded, observers recorded the teacher’s exposition of concepts, generalizations, and skills. Student and teacher tasks were also coded and categorized according to cognitive complexity. From this information, a single DOK level was calculated as a weighted average of DOK levels of both the moves and tasks (the total number of DOK 3 moves or tasks was multiplied by 3, DOK 2 by 2, etc.). This strategy was adopted so that the final teacher practice measurement was not deliberately or inadvertently biased by a teacher’s particular style of teaching. In other words, if a

teacher chose to use only tasks to teach a lesson, as long as the tasks required high levels of cognitive thinking, the teacher was not penalized for not using moves. By using a weighted average of moves and tasks, the resulting value reflected not so much the methods used as the overall DOK level of classroom activity. This methodology also eliminated the temptation to label activity as an end in itself. One well thought out DOK 3 level task can be more effective than numerous moves and tasks which do not reach beyond DOK level 1.

To help ensure inter-rater reliability of the coding, GAST observers underwent a two-day training session. During this time, two videotaped classroom sessions were coded and reviewed by the group together to establish uniform interpretations of teacher actions and coding methods. Two additional classroom videos were then coded independently by each observer. Results were tabulated and showed that all observers had produced scores that were within 50% of a benchmark score determined by the mean and median of observer scores. This training was repeated once more before the second year of data collection.

Student End-of-Course Post-Test

To measure student achievement, a reduced version of a Kentucky Geometry End-of-Course Assessment was given to the students of all GAST teacher participants at the end of the school year. The original assessment was designed by a team of educators led by Dr. William Bush at the University of Louisville (U of L) in collaboration with the Kentucky Department of Education (KDE). To develop the assessment, KDE and U of L staff selected a team of 13 professionals: mathematics educators (2), mathematicians (2), secondary school teachers (7) and secondary mathematics resource administrators (2).

The team created blueprints based on national and state standards documents (e.g., Achieve, NCTM, NAEP, TIMSS, Kentucky Program of Studies, and Kentucky Core Content). Team members wrote test items focusing on specific content areas and depths of knowledge according to the blueprint and Webb's DOK scale. The resulting assessment was pilot tested with approximately 3,200 Geometry students across Kentucky. After psychometric analysis and item revision, two versions of the assessment were developed and field tested with approximately 2,600 students. Psychometric results of these tests found high reliabilities for both sets of data ($\alpha = .92$ for Form 1 and $\alpha = .88$ for Form 2).

The reduced version of the End-of-Course assessment used for the GAST project consisted of items from the original version, but content areas were restricted to four major areas of geometry: similarity, congruence, area, and volume. Concentration in these areas was requested by NSF in accepting the proposal to develop GAST, and the student assessment was revised accordingly. The resulting test consisted of thirty items: 24 multiple-choice and six open response (three 4-point items and three 2-point items). At the end of the 2010-2011 school year, the test was administered to 952 high school geometry students. Psychometric analysis of the data produced a moderately high internal consistency reliability ($\alpha = .84$).

Student Entering Geometry Test

To control for student differences, a geometry "readiness" test was administered to the students of GAST teacher participants at the beginning of the school year. This test was used by the Cognitive Development and Achievement in Secondary School Geometry Project (CDASSG) conceived by Zal Usiskin and Sharon Senk of the

University of Chicago to study van Hiele level theories in geometry understanding (Usiskin, 1982). The assessment, originally named the “Entering Geometry Test”, was designed to measure the incoming geometry knowledge of students enrolled to take high school geometry. Usiskin’s data produced a K – R20 internal consistency reliability of 0.77. Because of the original purpose for which the test was created, the assessment aligns well in the present research design as a covariate with student achievement. Of the twenty items in the assessment, six focus on concepts of similarity, five focus on congruence, and three focus on area/volume. The other items were typically DOK level 1, definition items, which were necessary to develop the concepts of these four main content areas.

IRB Process

The procedures for communicating with the teachers, securing informed consent, collecting and using both pre-existing and new data were submitted to the University of Louisville Institutional Review Board (IRB) for approval. The data collection procedures described below were approved by the IRB committee on the condition that the preamble for consent be presented on the first page of the survey, instead of being included on the invitation letter or in prior notification emails or future email reminders. This guaranteed that subjects filling out the survey must have been exposed to the conditions of consent along with the purposes of the study.

Data Collection Procedures

Data for the GAST teacher assessment, teacher practice score, Student Entering Geometry Test, and Student End-of-Course Assessment were obtained from pre-existing GAST data. For teacher efficacy data, participants were invited to take an online version

of the TSES in the Spring of 2012 (which was one year after the first cohort participated in the GAST study). As an incentive to take the survey, participants were given \$5 in compensation. The administration of the survey followed the four-step procedure outlined below, which is a modified version of the method used by Dillman, Smyth, Christian, and Dillman (2009). Dillman's method was modified due to the specific population of GAST teachers, who had already collaborated with the GAST research project and personnel.

1. Prenotice email. Each teacher was contacted through a short email introducing the study and advising them that a follow-up email and letter would be sent explaining the study more fully. Teachers were given the option of taking the survey at this point if desired.

2. Invitation. A follow-up email was sent to each teacher, in addition to a letter containing an explanation about why a response was important, instructions on how to fill out the survey, and a \$5 token incentive. Consent was obtained through the teacher's willingness to participate in the survey, with consent details fully explained on the first page of the survey.

3. Thank you. A thank you email expressing appreciation for participation in the survey was sent out one week later. Teachers who had not yet responded also received this email and were again invited to take the survey.

4. Phone contact. Teachers who had not responded one week after the invitation were contacted by phone and politely invited to take the survey.

At the end of this process, 73 teachers had taken the survey. Responses from one teacher were removed because the teacher had not fully completed participation in the

GAST project. This return resulted in a 100% response rate from teachers who had completely fulfilled the requirements of the GAST project.

Limitations

Several pertinent limitations of the study must be noted. This section presents possible threats to validity using the structure outlined by Shadish, Cook, and Campbell (2002), including construct validity, internal and external validity, and statistical validity. These limitations will also be discussed in the context of the results of the study.

Construct Validity Threats

Each of the three latent variables teacher efficacy, geometry knowledge for teaching, and teacher practice, are strongly supported by research to be major contributors to student achievement gains. Measurement concerns, nevertheless, may threaten the validity of results found in the present study. Of the three, teacher efficacy is perhaps the most inherently difficult construct to measure, and no easy remedies to this problem seem to exist. As has been stated, the measurement of efficacy should be as task specific as possible, which in turn will limit the generalizability of the study. Beliefs in capabilities must be understood and measured as capabilities of *this task*. The task at hand can then be so specifically defined as to make results useless. For example, one might measure “the task of teaching of high school geometry, chapter 6 section 1 of textbook A, at a rural school with a 55% free and reduced lunch percentage, in a class with 27% English Language Learners during period 5 after the school pep-rally.” Though measuring the beliefs of the teacher at this time might be possible, generalizing the result certainly is not. Still, the procedures outlined in this study follow what previous research has predominantly used (Klassen et al., 2011). Efficacy was measured

with a self-report instrument which attempts, as much as possible, to define teaching tasks without limiting external validity. In its development process, the TSES was specifically created to address the most important tasks present in everyday teaching. For example, instead of eliminating the classroom management subscale when factor analysis showed poor results, more items were written to capture this factor, as researchers felt classroom management simply could not be left out. Nevertheless, the inherent difficulty of measuring efficacy can never be completely overcome (Pajares, 1996) and must be considered a threat to the construct validity of the TSES measurement.

The GAST instrument for measuring geometry knowledge for teaching also presented difficulties. Since this instrument was in its infancy, no additional research was available to support the reliability and validity of the measure. Though mathematical knowledge for teaching has been a major focus in mathematics educational research, much of this research has been carried out at the primary or middle school levels, and the concept of what mathematical knowledge for teaching is, continues to develop. Preliminary results validating the GAST instrument are currently being analyzed.

The last measurement, teacher practice, attempted to capture what is theoretically an extremely interesting and important aspect of teaching, the cognitive level of complexity of a teacher's daily practice. As might be expected, this construct presented a difficult playing field for any observation instrument. The different approaches to teaching can vary greatly across teachers as each teacher has his/her own background, preferences, personality, and experiences that all influence the choices of methods and pedagogical strategies. Even with the same teacher and students, classroom activities and teaching content vary from one day to the next. To increase measurement reliability,

the researcher conducted three observations for each teacher, with content area focused specifically on similarity, congruence, area, and volume. Still, producing an instrument that accurately measures the cognitive levels for such an enormous amount of teaching strategies and situations is no easy task. Construct validity must be considered when analyzing data from this type of study.

Internal Validity Threats

As has been discussed, a person's self-efficacy beliefs are intimately tied to the nature of the specific teaching task, making measurement and generalization difficult to reconcile. Raudenbush and colleagues (1992) using HLM techniques suggested that teachers report higher teacher efficacy scores when teaching high-track classes. This suggestion supports the idea that teacher efficacy fluctuates from year to year, and collecting data one year later for the first cohort would not be valid. But Raudenbush and his colleagues also found that this effect nearly disappeared when the teachers' perceptions of student engagement were considered. In addition, they used only one item to measure teacher efficacy, a possible threat to construct validity of their study. To build upon their findings, the current study incorporated a teacher efficacy instrument created by teachers to specifically address the essential aspects of the teaching profession. Student engagement ranks highly among these, so much so that 8 of the 24 items directly focus on student engagement issues, creating its own subscale. If high-track class effects are negated by a teacher's perceptions of student engagement, these effects will be accounted for by the TSES.

Still, the timing of the TSES administration was not optimal. In this study, the first cohort of teachers may have taught different classes with different students and

perhaps different courses when they completed the TSES survey, and some may have switched schools or were not teaching at all. The different setting might have changed their self-efficacy beliefs and rendered measurement inaccurate, thus compromising the internal validity of the study. To address this issue, the findings of Tschannen-Moran and Hoy (2007) are particularly relevant. They explored the sources of teacher efficacy beliefs, including key school resources and supports in the teaching context, and compared differences between novice and experienced teachers (experienced teachers had more than three years of experience). Results showed that the contextual variable of *school setting* was unrelated to the self-efficacy beliefs of both novice and career teachers, though the availability of resources did affect novice teachers. Neither group seemed to base their self-efficacy beliefs on the support of administrators. As can be expected, mastery experience made the biggest impact, especially for novice teachers. The authors wrote:

This variable [mastery experiences], assessed as satisfaction with teaching performance “this year,” was especially strong for novice teachers.

Career teachers, with an abundance of mastery experiences, may have a fairly stable sense of efficacy whether they are happy with how the current school year is going or not. (p. 954)

The findings of this study confirm what Bandura (1997) had suggested—that efficacy beliefs are more likely to change early in learning, and that once established, become resistant to change.

This finding supported the validity of the efficacy data to be collected one year after the GAST data were collected. Of the 44 teachers in the first cohort, only four

teachers had three years or less teaching experience. The range of experience for the remaining teachers in the first cohort spanned from 4 to 31 years, with a mean of 14.7 years of experience. Both current self-efficacy theory and empirical evidence support for stability of teacher efficacy beliefs for the vast majority of the study sample, thus measuring teacher efficacy beliefs one year later seemed a valid procedure. Still, efforts were made to collect additional data to account for any drastic changes to a teacher's professional life. Additional data collection items are listed in Appendix E.

Attrition of student participants may also be a factor leading to reduced internal validity. Of the original 1673 students who took the Entering Geometry Test, only 80.8% of them also took the End of Course Assessment. Likewise, of the 1576 students who took the EOCA, 14.8% did not take the EGT. It was not possible for the researchers to track these student absences or account for confounding teacher effects caused by student migration.

To help recuperate some of the lost data, EGT scores were imputed for students who produced valid EOCA results. The possible threat to internal validity resulting from this procedure is twofold. First, the EGT scores that each of these students would have produced are replaced by the class median. Though it seems reasonable that an individual would not score far outside the scores of his or her peers, this was not necessarily the case, and in the end there is no substitute for the true value. Secondly, the fact that the student did not have an EGT score led to the question of how long the student was exposed to the teacher's instruction and for what content areas. It was precisely this effect that the study attempted to measure. A significant loss of exposure to

the teacher's classroom instruction would definitely lessen the impact of this effect, threatening the internal validity of the study.

Finally, the multitude of sources that influence student performance was complex. School and district level effects, along with other classroom and student level variables such as curriculum, socio-economic status of both student and classroom averages, available resources, can have confounding effects on the variables of the current study. Although research supports the stronger influence of the teacher, other factors nevertheless could diminish the size of the effects, making relationships among variables harder to detect.

Statistical Validity Threats

Perhaps the greatest threat to the statistical validity of this study was the possible loss of power due to the sample size of the second-level variable. McCoach (2010) states that a minimum of 10 second-level clusters are required for a multi-level analysis, but adverse consequences arise if higher sample sizes are not achieved. For sample sizes with less than 30 second-level clusters, level 2 variance components can be overestimated while standard errors for all parameter estimates can be underestimated. For samples over 30, but less than 100, standard errors for the level 1 components can be reasonable, but level 2 variance components still might be underestimated. So although this study had the required 30 second-level clusters, it may be the case that the standard errors for the second-level variables were underestimated due to the small number of second level clusters (p. 129).

Another statistical validity threat stems from the restricted range of the efficacy measure. The TSES measured the teacher efficacy variable on a scale from 0 to 9, but

the range of teacher scores only spanned the top half of the scale, 5.29 – 8.67, with a standard deviation of only 0.73. This might be attributed to the fact that the teacher sample consisted of volunteer teachers. Teachers willing to allow observers to enter their classroom and collect data about their teaching would likely have positive beliefs about their teaching abilities. Although this did not alter the design of the study, the restricted range of the teacher efficacy construct made the relationship between teacher efficacy and student achievement harder to detect.

Finally the unreliability of measures may have contributed to a final statistical validity threat. As has been stated, both the GAST assessment and the GAST Classroom Observation Instrument (GCOI) were in their developing stages and no prior research existed to support the reliability of these measures. Possible measurement reliability issues, especially with the GCOI, again made teacher level influences on student achievement harder to detect within the context of the study design.

External Validity Threats

The projected population of this study was high school mathematics teachers employed in public and private schools. The study sample, however, consisted of volunteer teachers from both public and private schools in Kentucky. The fact that the sample contained only volunteers not only led to the restricted range problem mentioned above, but also added an external validity threat. Perhaps the relationships between teacher and student-level variables were only present in the type of teacher willing to participate in this kind of research, but would not be present in others. The schools were adequately mixed between public and private and urban and rural school districts, so theoretically threats to external validity for the student population were minimal.

CHAPTER 4

This chapter presents a description of the findings of the study. The descriptive statistics of the major teacher-level and student-level variables are presented first, followed by the results of the statistical analysis of the data including correlation analysis, ordinary least squares regression, multiple regression, and hierarchical linear modeling.

Descriptive Statistics

The descriptive statistics for the major teacher variables including geometry knowledge for teaching, teacher efficacy, and teacher practices are listed in Table 2.

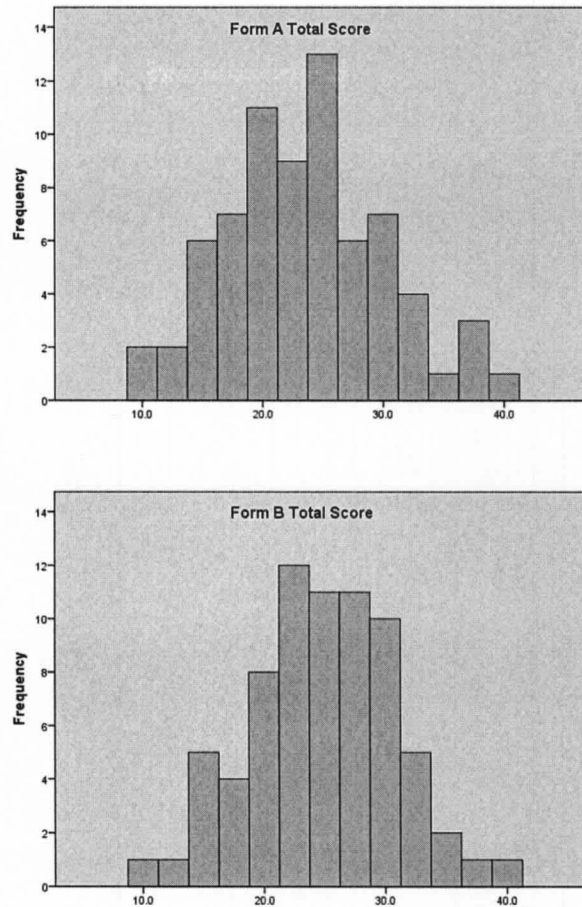
Table 2
Descriptive Statistics of Teacher Variable Measurements

	GKT (43 total pts)	GKT (43 total pts)	Teacher Efficacy (1 – 9)	Teacher Practice (DOK 1 - 3)
Measurement	GAST A	GAST B	TSES Score	Observation Instrument
n	72	72	72	72
Mean	23.74	24.82	7.10	1.66
Median	23.25	24.75	7.21	1.68
Std. Deviation	6.6	5.93	0.73	0.19
Range	29.0	31.0	3.38	0.93
Minimum	10.0	10.0	5.29	1.17
Maximum	39.0	41.0	8.67	2.10

GAST Form A assessments were administered at the beginning of the year, Form B assessments were administered within one month of the end of the geometry class teaching session. Score averages between the two forms were within 1 score point, with Form B slightly higher than Form A, and for both Form A and Form B, the means and

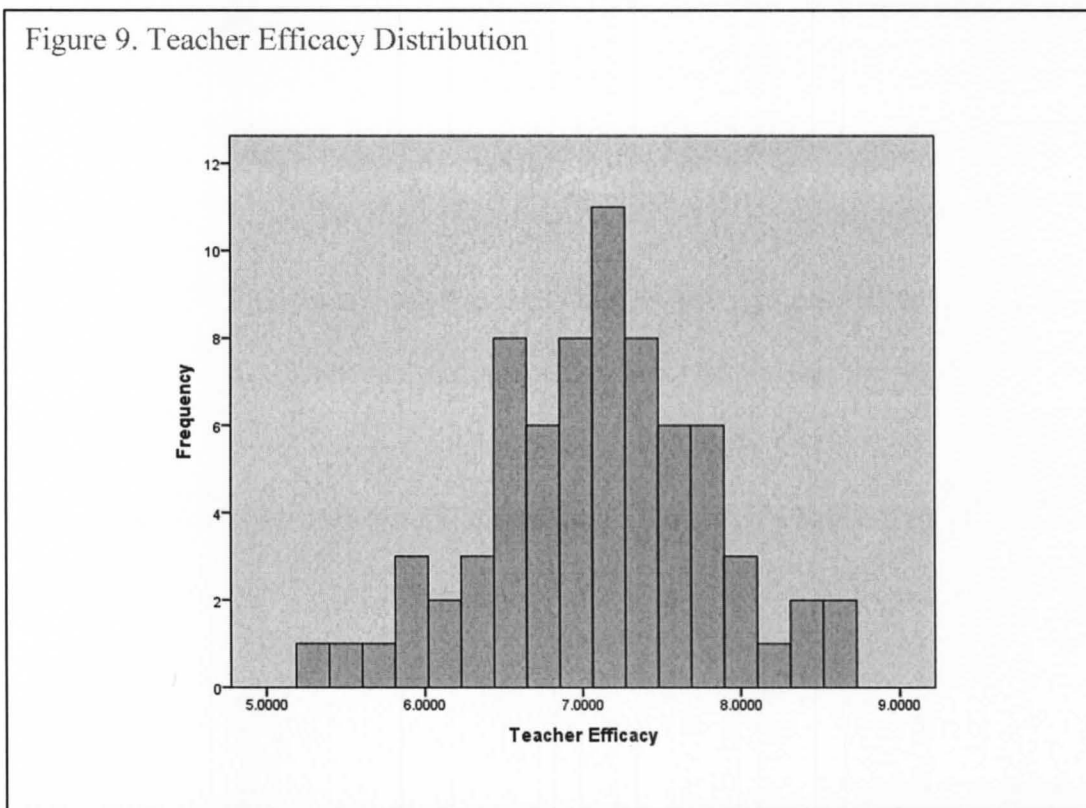
medians were almost identical for the same version of the test. The distributions of the two versions were close to normal, as shown in Figure 8.

Figure 8. GAST Assessment Distributions for Form A and Form B



Teacher Efficacy

The data from the Teacher Sense of Efficacy Scale survey produced a somewhat normal distribution (see Figure 9), although the center of the distribution is located at the higher end of the scale (the mean was $m = 7.1$ out of 9, with a standard deviation of $s = 0.73$). This shift towards the high end of the scale showed that the teachers in the study possessed high levels of teacher efficacy. An exploratory factor analysis of the data using a varimax rotation procedure produced a five factor structure, instead of the three factors designed by the authors of the instrument. Confirmatory factor analysis did not fully support the three-factor structure, although the statistics approached a reasonable fit. The full description of both analyses, including factor loadings of the five factors compared with the original specification, is described in Appendix H. The data collected with the TSES showed high reliability, yielding a Cronbach's alpha of $\alpha = .93$.



The last item on the survey included questions regarding any life changes that may have influenced teacher efficacy scores of the first cohort. A large proportion of first cohort teachers (30 of 44) reported that they felt their beliefs in their teaching abilities were about the same as they were last year. Nine teachers experienced a change of administration, but only one of those nine felt that this change affected their beliefs about teaching capabilities. One teacher no longer taught.

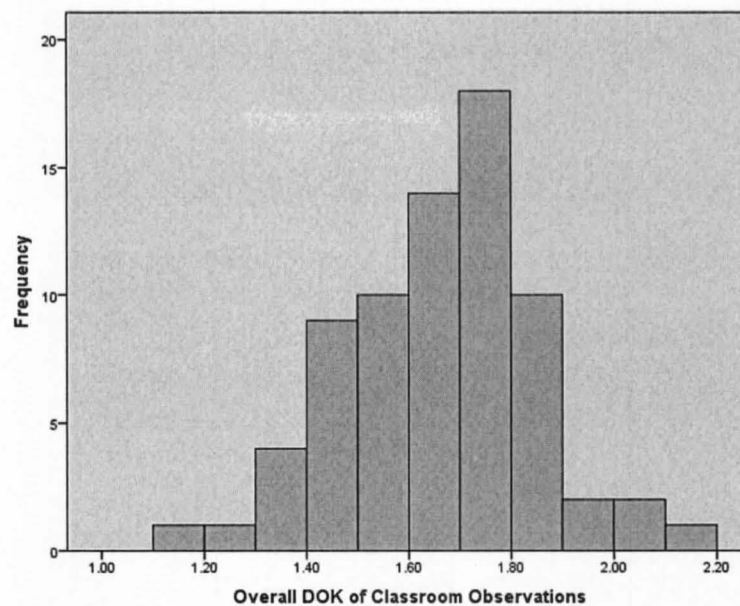
Differences between cohorts were also analyzed. Results showed no significant difference between the teacher efficacy means $F(1,70) = .000, p = .992$. Both cohorts produced nearly identical averages, and distributions were similar. When correlations with the other variables of the study were analyzed for both cohorts separately, results showed the same findings as the entire group overall; namely, no significant correlations with either the EGT or EOCA scores or the teacher level GKT and TP variables were found. Descriptive statistics and distributions for the two cohorts can be found in Appendix J.

Teacher Practice

Teacher practice data, measured by the GAST Observation Instrument, produced the distribution shown in Figure 10 with a mean of $m = 1.66$ and standard deviation of $s = 0.19$. Given that possible values ranged from 1 to 3, much of the teachers' classroom activity was on the lower end of cognitive complexity at levels 1 and 2. In fact, only three of the 72 teachers taught at an average cognitive complexity at or above DOK level 2 across three class periods. Within teacher statistics were also calculated. The average

standard deviation across the three observations for each teacher was 0.15 and the average difference between the highest DOK class to the lowest DOK class was 0.35.

Figure 10. Distribution of Teacher Practice Scores



Student Level Variables

In the 72 classrooms, 1895 students participated in the study with an average class size of 26 students. Of these students, 1673 took the Entering Geometry Test (EGT), 1576 took the End-of-Course Assessment (EOCA), and 1352 students took both tests. A student test was considered valid if there was at least one valid response (either correct or incorrect) on the test. This resulted in 71.3 % the total student sample that took both the EGT and the EOCA. The descriptive statistics for both assessments are shown in Table 3.

Table 3

Descriptive Statistics of the EGT and EOCA Variables

	EGT	EOCA
N Valid	1673	1576
N Missing	222	319
Mean	14.05	15.77
Median	15.00	15.00
Std. Deviation	4.056	7.435
Minimum	1	1
Maximum	20	37
Total Possible	20	40

The distributions of the Entering Geometry Test and the End of Course Assessment are shown in Figure 11 and Figure 12 respectively. Note that these two distributions are visually skewed from the normal distribution—the EGT is negatively skewed and the EOCA is positively skewed.

Figure 11. Distribution of the Entering Geometry Test

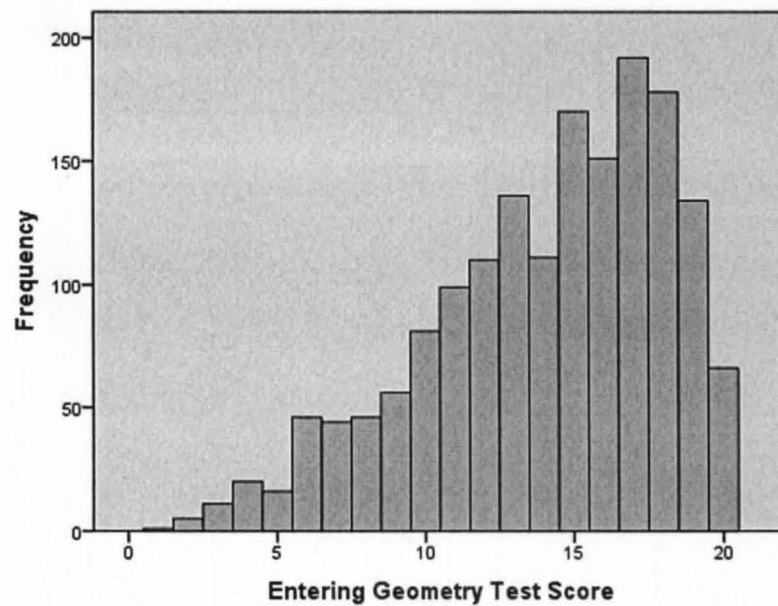
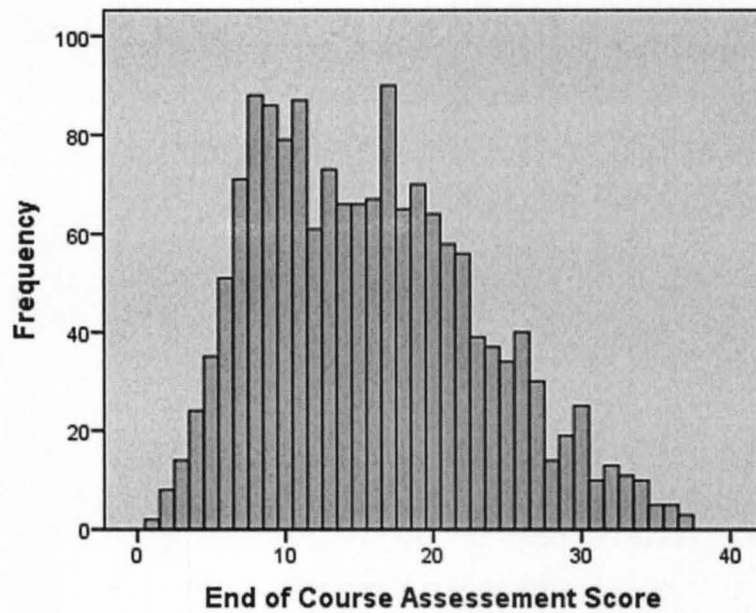


Figure 12. Distribution of Student End of Course Assessment Scores



Correlation Analysis

Before regression analyses were initiated, the researcher first explored correlations between each teacher variable (GKT, TE, and TP) and aggregated student level variables. Student level variables were aggregated using median class values for both the End-of-Course Assessment and Entering Geometry Test. Results are shown in Table 4. None of the teacher level variables were significantly correlated with each other, and neither GKT nor TE had any sizable or significant correlation to the student outcome variable EOCA. Teacher practice, however, showed positive significant correlations to student achievement, although the correlations were small. When teacher

practice was further disaggregated, the number of DOK 3 moves showed a positive correlation to student achievement, as did the total number of DOK 3 moves and tasks performed by each teacher. The years of experience of the teacher (YRS_EXP) showed the highest correlation with student achievement among teacher level variables.

Table 4

Teacher Level Bivariate Correlations

	Median EOCA Score	Median EGT Score	GAST A	GAST B	TE	TP	No. of DOK3 Teacher Moves	Total No. of DOK3 Moves &Tasks	YRS EXP
Median EOCA	1	.832	.122	.071	.093	.262*	.298*	.287*	.493**
Median EGT		1	.010	-.029	.089	.174	.184	.213	.339**
GAST A			1	.733**	.191	.078	.124	.064	.267*
GAST B				1	.076	-.033	.096	-.007	.226
TE					1	-.125	-.194	-.275*	.081
TP						1	.581**	.608**	-.017
No. of DOK3 Teacher Moves							1	.898**	.096
Total No. of DOK3 Moves &Tasks								1	.149
Years Experience									1

** $p < 0.01$ * $p < 0.05$

Student-level correlations were also performed. Student EGT readiness test scores were highly correlated with the EOCA scores $r(1354) = .68, p < .01$, and aggregated class median EOCA scores were highly correlated to the aggregated median EGT scores $r(71) = .85, p < 0.01$). These statistics support the use of the EGT student readiness test as a covariate with student achievement.

Regression Analysis

Correlation analyses revealed that the teacher practice variable showed promise of having a predictive relationship to student achievement. To investigate this relationship

further, ordinary least square regression analysis was performed for the teacher practice variable as a first step in looking at coefficient sizes and significance. (The lack of correlation between teacher efficacy and geometry knowledge for teaching with student achievement ruled out the possibility of a significant OLS regression equation for these two variables). Also, to guard against possible skewness of the EOCA student scores, the class median scores, and not the mean scores of the EOCA, were used as the dependent variable representing student achievement (at the class level).

As displayed in Table 5, teacher practice showed a significant predictive relationship to student outcomes, accounting for 5.5% of the variance in student achievement. The small Adjusted R was not surprising, given the small correlations between dependent and independent variables at the outset. The teacher practice coefficient was calculated to be relatively large at 8.03. Though the 3.53 standard error for this estimate is also large, the teacher practice relationship with student achievement was still positive.

Table 5
Ordinary Least Squares Regression Results

$$\text{Median EOCA} = B_0 + B_1(\text{TP})$$

Coefficients	B	Std. Error	Beta	t	Sig.
Intercept, B_0	1.464	5.917		.247	.805
Teacher Practice	8.026	3.538	.262	2.268	.026

Model Summary

R	R ²	Adj R ²	Std Error	R ² Change	F Change	df1	df2	Sig. F Change
.262	.068	.055	5.70100	.068	5.145	1	70	.026

Since the teacher variables were not significantly correlated, it was not possible to use ordinary least squares regression to analyze the predictive nature of teacher efficacy

and mathematics knowledge on teacher practice. No significant relationships were revealed among these variables.

Multiple Regression Models

Two multiple regression models reinforced what was found in the correlation analysis and the OLS equations; namely, teacher practice was a significant predictor of student outcomes while teacher efficacy and geometry knowledge for teaching were not. The first regression model incorporated only the three principle teacher level variables:

$$EOCA = B_0 + B_1 (TE) + B_2 (GKT) + B_3 (TP) \quad (1)$$

To run these analyses (including subsequent models), variables were entered using stepwise procedures because all three teacher variables theoretically have strong influences on student achievement, with no preference of one over the other. In this way, the order of entry was driven by the strength of the correlations of each independent variable with the dependent variable using criteria of $p \leq .05$ to include variables and $p \geq .10$ to remove variables. When the model was analyzed in SPSS, both the TE and GKT variables were excluded by the stepwise procedure. Only the teacher practice variable remained in the model, resulting in the same regression equation of the OLS analysis above, with the same coefficients and p-values. Again, this result follows logically from the correlation analysis conducted from the outset.

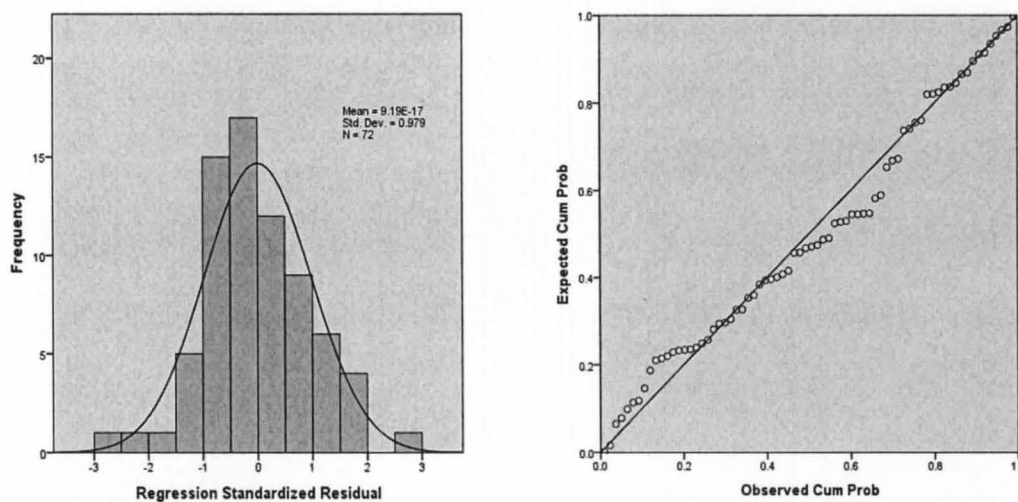
For the second regression model, the aggregated median Entering Geometry Test covariate was added, which is an important factor when analyzing the true relationship between the teacher variables and student achievement. Also included was the years of

experience of the teacher. This yielded the final model represented by the equation below:

$$\text{EOCA} = B_0 + B_1 (\text{TE}) + B_2 (\text{GKT}) + B_3 (\text{TP}) + B_4 (\text{Med. EGT}) + B_5 (\text{YRS_EXP}) \quad (2)$$

The additional two variables reduced the statistical power of the model, but according to Stevens (2009, p. 71), the data should contain at least 15 subjects per predictor to produce sufficient power. With 72 subjects, the current data set provided sufficient power for use with a five predictor model, though ideally more subjects would have been better. The other assumptions required by the model were also met. Both the histogram of residuals and the P-P plot (Figure 13) showed evidence of satisfying the normality of errors requirement.

Figure 13.
Histogram of Standardized Residuals & PP Plot

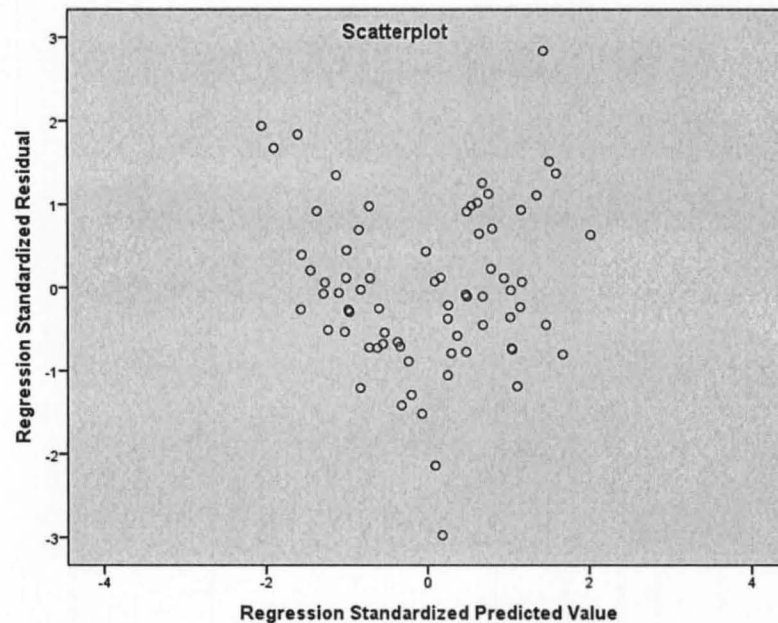


The linearity and homoscedasticity between predicted dependent variable scores and residual errors were verified by the scatterplot presented in Figure 14, which shows a random distribution about the zero axis. Regarding multicollinearity, SPSS collinearity

statistics revealed little collinearity between the predictors in the final model (Tolerance levels were between 0.852 and 1.000 and VIF scores were much less than 10). This finding was not surprising as correlation analysis revealed no evidence of any relationships between the predictors. One caveat should be noted regarding the independence of the aggregated student scores. Though theoretically the aggregated student-level scores are likely independent, an influence on student scores at the school level where two teachers taught at the same school was possible. Still, this influence would seem small compared to the much stronger influence of the teacher.

Figure 14.

Scatterplot of Residuals vs. Predicted Errors



Once assumptions were verified the model was analyzed, again using stepwise entry for independent variables. The procedure resulted in three variables entering the equation in the following order: the covariate EGT, YRS_EXP, and teacher practice (see Model 3, Table 6). This model was statistically significant $F(3, 68) = 71.970, p < .001$

with an Adjusted R² of .750 (using the Wherry formula). Coefficient statistics are listed in Table 7.

Table 6
Model Summaries

Model	R	R ²	Adj.R ²	Std. Er.	R ² Chg	F Chg	df1	df2	Sig. F Chg	F	Sig
1	.832	.691	.687	3.281	.691	156.841	1	70	.000	156.841	.000
2	.861	.742	.734	3.024	.050	13.398	1	69	.000	99.009	.000
3	.872	.760	.750	2.933	.019	5.366	1	68	.024	71.970	.000

Model 1: Predictors: (Constant), Median EGT

Model 2: Predictors: (Constant), Median EGT, Yrs. Experience

Model 3: Predictors: (Constant), Median EGT, Yrs. Experience, Teacher Practice

Excluded Variables: GKT, TE

Table 7
Coefficient Statistics

Independent Variable	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	-17.203	3.332		-5.162	.000		
EGT	1.568	.140	.722	11.235	.000	.852	1.173
Yrs. Exp.	.157	.040	.250	3.952	.000	.879	1.138
TP	4.296	1.855	.140	2.316	.024	.963	1.038

The resulting regression equation becomes:

$$EOCA = -17.203 + 4.296 (TP) + 1.57 (Med. EGT) + 0.157 (Yrs. Exp) \quad (3)$$

which accounts for about 75% of the variance of the student achievement scores. The teacher practice coefficient has dropped in magnitude but still holds a positive predictive relationship with student achievement despite the standard error.

HLM Analysis

Because the GAST data was naturally clustered by students within classes, Hierarchical Linear Modeling (HLM) was an especially appropriate method to look for linear predictive relationships. Using HLM not only allowed the teacher level effects to

vary across classes, but also enabled the variance levels among first and second level components to be explored. Missing data at the student level, however, presented a possible threat to statistical power. In the final dataset, missing student data points made up a significantly large percent of the student sample (28.7%). For this reason, the missing EGT scores were imputed with the median EGT from the students' corresponding class, adding an extra 12% of the data back into the sample. This resulted in 83.1% of the student sample as usable first level units. The addition of these data brought with it the possibility of deflating the teacher's influence, depending on how long the student was in the class and what content area was covered during this time. Nevertheless, some degree of teacher influence was necessarily present in each of these student scores and the extra power added by the increased sample size would help detect this influence if present.

The HLM analysis incorporated a three-stage procedure to explore possible second-level teacher effects on student achievement. The analysis began with the unconditional model to determine the amount of second-level variance contained in the student achievement scores. Then the regression analyses results were used to inform the HLM methodology: first the three teacher level variables were entered individually, and then the teacher's years of experience variable was added which had earlier proved to be significant. As will be shown, much of what was found in the correlation and multiple regression analysis was substantiated by the HLM exploration. Finally, all teacher variables were entered as both intercept and slope coefficients at the second level to explore any possible coefficient significance that had not yet been discovered. These HLM models were analyzed using HLM 7 software which incorporates a restricted

maximum likelihood method of estimation. For all of the HLM models, uncentered variables were used since within group differences were not the focus of study and interpretation with raw scores was straightforward.

As mentioned above, the unconditional model (also called the One-Way ANOVA with Random Effects model; see Raudenbush and Bryk, 2002, p. 23) was used to determine variance at the group level. This model removes all independent variables and covariates from the equation as shown in Table 8. HLM results showed that 49.9% of the variance of student scores was found at the teacher level:

$$\rho = \frac{\tau_{00}}{\tau_{00} + \sigma^2} = \frac{28.700}{28.700 + 28.770} = .499$$

The grand mean of the dependent variable EOCA was $\gamma_{00} = 15.289$, with a standard error of 0.645.

Table 8
Results from One-Way ANOVA, Unconditional Model

$$EOCA_{ij} = \beta_{0j} + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

where

Y_{ij} = Student EOCA score
 β_{0j} = mean EOCA score for group j

final equation

$$EOCA = \gamma_{00} + u_{0j} + r_{ij}$$

<i>Fixed Effect</i>	<i>Coefficient</i>	<i>Se</i>	<i>t-ratio</i>	<i>P Value</i>
Average EOCA score (grand mean), γ_{00}	15.289	0.645	23.7	<0.001
<i>Random Effect</i>	<i>Variance</i>	<i>df</i>	<i>X²</i>	<i>p Value</i>
Group mean, u_{0j}	$\tau_{00} = 28.700$	71	1535.300	<0.001
Level-1 effect, r_{ij}	$\sigma^2 = 28.770$			
Deviance = 9981	Reliability estimate, $\beta_0 = 0.94$			

After the unconditional model analysis, variables were introduced to identify the level 1 and 2 sources of variance similar to the methodology of the regression analysis. First, each of the three principle teacher variables were entered as the only second level variable both with and without covariate EGT scores on level one. This method was used in hopes that the HLM analysis might detect teacher effects not found with the aggregated regression model. The HLM equation using the teacher practice variable and the EGT score used the form given below. The slope β_{1j} was entered as a fixed effect because prior analysis revealed the variability of the random effect u_{1j} to be insignificant. This was true in every model in the analysis.

HLM Equations with the Covariate EGT and TP as an Intercept Predictor

Level-1

$$EOCA_j = \beta_{0j} + \beta_{1j}*(EGT_{ij}) + r_{ij}$$

Level-2

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}*(TP_j) + u_{0j} \\ \beta_{1j} &= \gamma_{10} \end{aligned}$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{01}*TP_j + \gamma_{10}*EGT_{ij} + u_{0j} + r_{ij}$$

After the three teacher-level predictors were entered separately, neither geometry knowledge for teaching nor teacher efficacy produced significant coefficients in the HLM calculations. The teacher practice variable, however, was found to be a significant predictor ($p = 0.03$), and produced a coefficient multiplier of 5.168 for each DOK level of a teacher's instruction. The standard error of the TP coefficient was a relatively large 2.236. Despite the standard error, the TP 95% confident interval (0.785, 9.551) does not

cross zero, supporting the claim that teacher practice is a positive predictor of student achievement. Model statistics are shown in Table 9 below.

Table 9

HLM Model Statistics with Teacher Practice as an Intercept Variable

<i>Fixed Effect</i>	<i>Coefficient</i>	<i>Se</i>	<i>t ratio</i>	<i>dF</i>	<i>p Value</i>
For Intercept, β_{0j}					
Intercept, γ_{00}	-7.377	3.761	-1.92	70	0.054
TP, γ_{01}	5.168	2.236	2.25	70	0.024
For EGT slope, β_{1j}					
Intercept, γ_{10}	0.991	0.058	17.006	1503	<0.001
<i>Random Effect</i>	<i>Variance</i>	<i>df</i>	<i>X²</i>	<i>p Value</i>	
Group mean, u_{0j}	$\tau_{00} = 11.565$	70	838.930	<0.001	
Level-1 effect, r_{ij}	$\sigma^2 = 21.442$				
Deviance = 9477	Reliability estimate, $\beta_0 = 0.904$				

The total unexplained variance of 57.470 from the unconditional model has now dropped to 33.007. (For a summary of variance values, see Table 12 below.) The addition of the Entering Geometry Test and the teacher practice variables explained 42.6% of the original variance and 59.7% of the original variance at the teacher level. Of the remaining unexplained variance, 35.0% [$\tau_{00} / (\tau_{00} + \sigma^2)$] remains at the teacher level. However, most of the drop in variance can be attributed to the EGT covariate which when entered without TP, reduced second level variance to 12.357 compared to 26.779 produced when the TP variable was added alone (26.779 is a drop of 3.34% from the original variance component). Both models are listed in Appendix I along with model statistics. These models again substantiate what was found earlier in the multiple regression analysis which showed that the largest amount of variance was explained by the EGT variable, and the TP variable only explained an additional 1.9% more in the final regression model.

Teacher practice effects were also significant when included in the slope coefficient for the Entering Geometry Test covariate. The equations for this model are given along with coefficient statistics in

Table 10. Though the magnitude of the γ_{11} coefficient is a relatively small 0.427, the $\gamma_{11} * TP_j * EGT_{ij}$ term in the mixed model equation becomes $(0.427) * TP_j * EGT_{ij}$, which can be rearranged as $TP_j * (0.427 * EGT_{ij})$. If the student scores well on the Entering Geometry Test, the increase in student achievement resulting from the DOK level of classroom activity could be substantial. The average of the EGT was 14.05, so the TP effect for an average student would be $TP * (0.427) * 14.05$, or $6.00 * TP$. Effects on student achievement would vary from this according to student performance above or below the average EGT. Still, this equation indicates that teacher practice had substantial interaction effects with readiness levels on student achievement. Variance components remained similar to the TP intercept model discussed above.

HLM Equations with Teacher Practice as EGT Slope Variable

Level-1

$$EOCA_j = \beta_{0j} + \beta_{1j} * (EGT_{ij}) + r_{ij}$$

Level-2

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11} * (TP_j) \end{aligned}$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{10} * EGT_{ij} + \gamma_{11} * TP_j * EGT_{ij} + u_{0j} + r_{ij}$$

Table 10

HLM Model Statistics with Teacher Practice as EGT Slope Variable

<i>Fixed Effect</i>	<i>Coefficient</i>	<i>Se</i>	<i>t ratio</i>	<i>df</i>	<i>p Value</i>
For Intercept, β_{0j}					
Intercept, γ_{00}	1.273	0.716	1.779	71	0.080 0.024
For EGT slope, β_{1j}					
Intercept, γ_{10}	0.273	0.272	1.003	1502	0.316
TP, γ_{11}	0.427	0.158	2.701	1502	0.007
<i>Random Effect</i>	<i>Variance</i>	<i>df</i>	<i>X²</i>	<i>p Value</i>	
Group mean, u_{0j}	11.646	70	838.930	<0.001	
Level-1 effect, r	21.349				
Deviance = 9476	Reliability estimate, $\beta_0 = 0.905$				

The final models added teacher experience to the two teacher practice models shown above. The intercept model shown in Table 11 again produced significant coefficients and helped explain 4.6% more of the original unexplained variance of student achievement scores. On the teacher level, 68.8% of the original second level variance was explained by the model. The TP coefficient remained unchanged from the previous model, and the teacher experience coefficient, though relatively small, still provided evidence that teacher experience had a positive influence on student scores. Deviance for this model had dropped from the previous models, showing a progressively

better model fit. Ultimately this model verified the results of the multiple regression analysis, where TP and teacher experience were meaningful predictors of achievement.

Table 11

Final Model, TP and Yrs_Exp as Intercept Variables

Level 1

$$EOCA_{ij} = \beta_{0j} + \beta_{1j}*(EGT_{ij}) + r_{ij}$$

Level 2

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(TP_j) + \gamma_{02}*(YRS_EXP_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{01}*TP_j + \gamma_{02}*YRS_EXP_j + \gamma_{10}*EGT_{ij} + u_{0j} + r_{ij}$$

<i>Fixed Effect</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>t-ratio</i>	<i>Approx. d.f.</i>	<i>p-value</i>
For INTRCPT1, β_0					
Intercept, γ_{00}	-10.181	3.338	-3.049	69	0.003
TP, γ_{01}	5.339	2.002	2.667	69	0.010
YRS_EXP, γ_{02}	0.179	0.043	4.144	69	<0.001
For EGTI slope, β_1					
Intercept, γ_{10}	0.984	0.059	16.817	1503	<0.001
<i>Random Effect</i>	<i>Variance</i>	<i>df</i>	<i>X²</i>	<i>p Value</i>	
Group mean, u_{0j}	8.957	69	658.858	<0.001	
Level-1 effect, r_{ij}	21.435				
Deviance = 9465	Reliability estimate, $\beta_0 = 0.88$				

As with the earlier teacher practice model, the two teacher level variables TP and YRS_EXP were shifted to the EGT slope coefficient equation. While both the TP and YRS_EXP coefficients were significant, the teacher experience coefficient was close to zero, and no additional variance reduction was found at either level. The Deviance level was also not significantly reduced. Statistics of this model can also be found in

Appendix I.

Table 12

HLM Model Variances

	Unconditional Model	EGT + TP Intercept Model	EGT + TP + YRS_EXP Model
Teacher-level Unexplained Variance, INTRCPT1, u_0	28.700	11.565	8.957
Student-level Unexplained Variance, r	28.770	21.442	21.435
Total Unexplained Variance	57.470	33.007	30.392
% Variance at 2 nd Level	49.9%	35.0%	29.5%
%Total Variance Explained		42.6%	47.1%
%Teacher Level Variance Explained		59.7%	68.8%
% Additional Variance explained by adding YRS_EXP to TP Model			4.6%
% Additional Teacher Level Variance explained by adding YRS_EXP to TP Model			9.1%

Exploratory Models

To fully investigate any further coefficient significance or combinations of variables which may collectively predict student achievement, a model with all teacher level variables were entered simultaneously for both the intercept and slope equations as shown below.

Level-1 Model

$$EOCA_{ij} = \beta_{0j} + \beta_{1j}*(EGT_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(GKT_j) + \gamma_{02}*(TE_j) + \gamma_{03}*(TP_j) + \gamma_{04}*(YRS_EXP_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}*(GKT_j) + \gamma_{12}*(TE_j) + \gamma_{13}*(TP_j) + \gamma_{14}*(YRS_EXP_j) + u_{1j}$$

The results of this model yielded no new information, but reinforced what had been previously found—that the teacher practice and teacher experience variables were significant predictors of student achievement. The full statistics of this model can be found in Appendix I. Numerous other models were examined, especially models which

involved the number of DOK3 moves and tasks, which had previously been shown to have significant correlations with student achievement (measured at the aggregate level). These models produced very small coefficient sizes (typically near 0.10) but nevertheless, were found to be statistically significant. Models exemplifying the DOK variables are listed in Appendix I.

CHAPTER 5

Discussion

Improving student mathematics scores is one endeavor in the country's political, social, cultural, and economic life that is supported by the population across all boundaries. Both TIMSS 2007 and PISA 2009 verified that recent generations of Americans are more deficient in mathematics skills when compared to other industrialized nations. Improving this situation is of utmost importance. Research reveals that the strongest influence of student achievement besides student characteristics, are teachers (Coleman et al., 1966; Darling-Hammond, 2007; Sanders, 2004). However, how to determine the measurable teacher qualities that contribute most to student learning is still unknown (Boyd et al., 2007; Goldhaber, 2002). For example, although teacher efficacy has a positive relationship with student achievement (Ashton & Webb, 1986; Berman et al., 1977; McLaughlin & Marsh, 1978; Moore & Esselman, 1992; Ross, 1994), one group of researchers questioned whether improving teacher efficacy is enough on its own to make a difference (Raudenbush et al., 1992). This study examined whether or not teachers' efficacy, along with their geometry knowledge for teaching (GKT) and the level of cognitive complexity of their classroom practices (TP), predict student performance on mathematics achievement tests. In addition, correlations among these variables, along with the predictive nature of teacher efficacy and geometry knowledge for teaching on classroom cognitive complexity levels were investigated. The results of

this study will be discussed by analyzing each teacher variable, its relationship to student achievement, and possible factors that may have influenced these relationships.

Teacher Efficacy and Student Achievement

Previous research has found teacher efficacy to be positively related to student achievement scores (Anderson et al., 1988; Ashton & Webb, 1986; Ross, 1992) and to with-in teacher variables, such as a willingness to work longer with struggling students (Gibson & Dembo, 1984) and a willingness to implement instructional change (Ashton & Webb, 1986). The results of this study did not support these findings. Teacher efficacy was not correlated to either student achievement or the geometry knowledge for teaching (GKT) or teacher practice (TP) variables. Furthermore, neither the regression analyses nor the HLM techniques revealed relationships between teacher efficacy and the other variables in the study. In fact, both correlation results and regression coefficients were insignificant, frequently with very large p-values. The only significant teacher efficacy correlation was found between the teacher efficacy variable and the teacher's total number of DOK 3 moves and tasks, which produced a small negative correlation ($r = -.275, p = .02$). These findings do not necessarily confirm that no relationship exists between teacher efficacy and student achievement. It may be the case that the particular instruments and methods used in this study were not capable of detecting a significant relationship.

Perhaps the most logical explanation of these results is the statistical validity threat caused by the low number of second-level clusters. The fact that only 72 teachers (units at the second level) participated in the study may have weakened the statistical power of the study so that significant relationships between the dependent and

independent variables were not detected. This low power was present in all of the findings of the study, including the relationships between GKT and TP and student achievement. Moreover, selection bias may also have been a validity threat due to a ceiling effect, with most participants scoring near the top of the range. Because participating teachers were not randomly selected and had volunteered to take part in the study, they might have had higher than expected efficacy beliefs. As a result, the TE scores might not have had the necessary variance to reveal linear relationships. The data seemed to support this conjecture. The average TE score for this group of teachers was 7.1 on 9-point scale. This group of teachers felt strongly about their abilities to promote student learning, a result that was not surprising given their proactive nature to volunteer to participate in the study. Also, self-report measures can be biased by social desirability, which was a concern in earlier teacher efficacy measures (Ashton et al., 1982). Teachers may possess low self-efficacy tendencies, yet might not admit this shortcoming on a self-report survey. This tendency might have added to the restricted range problem. The resulting variance of the TE scores was quite small, $s = 0.73$, thus 68% of the teachers had TE scores between 6.37 and 7.83. With this restricted range, detecting an accurate linear relationship between teacher efficacy and student achievement was likely diminished.

The internal validity threat due to the delayed administration of the TSES also might have contributed to an inaccurate measurement of teacher efficacy. The research of Tschannen-Moran and colleagues (2007) supported the methodology used here because of the high representation of experienced teachers in the first cohort (90.9%). However, the optimal time to administer the instrument might have been at the end of the

same year that the GAST test was administered to each group. Moreover, the inherent difficulty of measuring teacher efficacy due to the task specific-versus-generalizability problem may have been increased by the late administration for the first cohort.

Nevertheless, a large proportion of the first cohort (30 of 44) reported that they felt their beliefs in their teaching abilities were about the same as they were the year before.

Regarding the teachers who felt their beliefs changed, one might suspect that changes in belief from one year to the next would be relatively small when measured by the TSES, given the experienced population of the sample. As with any quantitative study in the social sciences, latent variable measurements can be difficult, and the complications of the late administration and the inherent difficulties of measuring the teacher efficacy construct may have caused the student achievement relationships to go undetected.

Future research might investigate adjustments to the frequency and timing of the administration of teacher efficacy instruments. For this study, a more accurate measurement of the true efficacy might have been obtained by having teachers complete the TSES immediately prior to the classroom observations and then averaging the three TE scores for each set of classes, especially considering the same procedure was used to collect the TP variable data. Without specifying any particular task, the beliefs and attitudes of the teacher regarding the particular content area might have been captured. This method also may be more likely to identify any further underlying linkages to the TP and GKT constructs while also capturing a better picture of the teacher's sense of efficacy in general.

Geometry Knowledge for Teaching and Student Achievement

The results of the geometry knowledge for teaching analysis were also inconclusive. Although previous research supported the positive relationship of teachers' mathematical knowledge for teaching on student achievement (Hill et al., 2005), this study did not reveal substantial relationships. No significant correlations were found between teachers' geometry knowledge for teaching and student achievement, nor were any significant correlations between teacher efficacy or the DOK level of a teacher's instruction and student achievement found. In addition, neither regression analyses nor HLM modeling produced any significant findings.

Again, the first place to look for possible reasons why no significance was detected is lack of statistical power. The relatively small number of clusters may simply not have produced enough statistical power to detect GKT predictive relationships with the other variables. Upon analyzing the descriptive statistics of the GAST scores, I found nothing that would cause concern regarding the possible selection bias caused by incorporating volunteer teachers as the sample population. Average scores were near the middle of the 43-point test (GAST A mean = 23.74, GAST B mean = 24.82) with standard deviations of 6.6 and 5.93 points, respectively. The range of scores was near 30 points for both tests, and the distributions were relatively normal and centered about the mean. These statistics suggest that restricted range was not an issue in the analysis—the data points themselves simply did not produce significant relationships with the other variables.

Because this study represented the first administration of the GAST test, no previous research was available to support construct validity of the test outside the GAST

project itself. Possibly the assessment did not capture fully the underlying construct it was intended to measure. Still, the blueprint did not appear to contain deficiencies that might cause problems in capturing geometry knowledge for teaching. Although the blueprint design included three subscores (Teacher Knowledge of Mathematics, Teacher Knowledge of Geometric Reasoning and Problem Solving, and Teacher Knowledge of Student Learning) these factors were not weighted equally. The third factor involving student learning, contributed 45% of the total score, and this factor might have been the most likely to capture pedagogical elements of GKT. Nevertheless, the third factor itself showed no significant correlations to the teacher-level or student-level variables. Ultimately, the blueprint was created and reviewed by experts in the field, and the design would not likely have glaring flaws, although individual item analysis might be used to refine the test further.

Selection bias may have also caused internal threats to validity—not only because of the selection of teachers, but also because of the schools included in the study. Although a mix of rural and urban districts were represented, the sample selection was not random, leaving room for bias. To account for differences in student/classes, the study used the Entering Geometry Test as a covariate, which similar to the GAST test, produced reasonable descriptive data. Since this test was designed precisely to capture the aptitude of students beginning high school geometry, it seemed an appropriate measure for use in this study. The same was true of the End-of-Course Assessment which was designed by content experts to specifically capture students' depth of knowledge in geometry. However, the reduced EOCA test which included only the four content areas of similarity/congruence and area/volume may have in some way altered

the psychometric properties of the assessment. This change may have skewed student achievement scores in ways not anticipated.

Another area that may have accounted for the absence of GKT relationships was the large amount of missing data at the student level. Of the 1,576 students who took the EOCA, 14.8% did not have covariate EGT scores, leading to the decision to impute the data with class median scores to gain statistical power. Not only did the imputed data threaten internal validity, but the fact that 14.8% did not have an EGT score raised concerns. If this missing data was caused by a significant amount of student migration, many students would not have received the full impact of the teacher's knowledge and teaching skills, perhaps resulting in a decreased effect size for the teacher level variables. In addition, the degree to which teacher effects were reduced would depend not only on the time the migration occurred, but also on which geometry content was covered during students' absences. These two factors were impossible to track in the study.

Finally, the many other factors that influence student learning at the student, class, and school level cannot be completely determined. This study incorporated one student-level independent variable, the covariate. Perhaps if other first-level variables were included such as socio-economic status or parent education, along with the interaction effects of these variables and second-level variables such as teacher preparation, other variances not previously identified could be uncovered.

Teacher Practice and Student Achievement

Although teacher efficacy and geometry knowledge for teaching data did not produce significant results, the data revealed that the cognitive level of teacher practices was a significant predictor of student achievement. The remainder of this chapter will

focus on this finding, along with the effects of the additional teacher-level variable, years of teaching experience.

The descriptive statistics of the TP variable did not reveal high cognitive levels of instruction. The top half of the range of classroom DOK levels was barely breached by the teacher data, which showed a maximum individual teacher score of 2.10 out of a range of 1 – 3 (see Figure 10 for the distribution of scores). In fact, only three teachers averaged DOK level 2 or higher. Fifteen teachers scored below 1.5, indicating that many of the teachers at least tended more towards DOK 1 level moves and tasks than DOK 2 level moves and tasks.

If the GAST Observation Instrument accurately measured what it was intended to measure, this finding alone is worth addressing. When international examinations are constructed to compare the mathematics achievement levels across industrialized nations, test designers likely did not create many DOK 1 items. Mathematics tasks at the DOK 1 cognitive complexity focus little on mathematics comprehension and more on memorization and retention skills. One could argue that even a DOK level 2 of cognitive complexity does not capture the essence of what makes mathematics what it is. While students must learn more than basic skills and concepts, what really exemplifies *doing* mathematics is the understanding and creative/exploratory element that comes from students thinking, elicited by a teacher through high cognitive level teaching moves and student tasks. If teacher practice data show that the tasks that teachers give students are, on average, at DOK level 2 or lower, then students will not likely be capable of really *doing* mathematics at DOK level 2. In this case, students were not using higher levels of cognitive thinking because they were not exposed to it by their teachers. The descriptive

statistics of the teacher practice variable alone are perhaps the most useful finding of this study, and the picture painted by these data of the future of mathematics education is not an encouraging one.

The conclusions above assume that the GAST Observation Tool accurately captured the cognitive complexity levels of the teacher's moves and tasks presented in the classroom. This assumption is not without concerns. Though GAST classroom observers participated in extensive training to insure the reliability of scores, the sheer number of ways a particular teaching move may be recorded may have presented reliability threats. For this reason, the method of interpreting observation results for this study adopted a "global perspective" by trying to capture only the overall DOK level of the teacher's instruction. In other words, by using this method, it made no difference whether a teacher gave a counterexample for a concept (assigned DOK level 2 by the rubric) or assigned a DOK 2 level task, both were quantitatively identical. Though the logic behind this method makes sense, whether or not this method is optimal for capturing the DOK level of the class requires further research, especially considering the widely varying approaches to teaching and personal characteristics that teachers bring to their classrooms.

Additional research is also needed (1) to verify the construct validity of the instrument in general and (2) to investigate methods to help observers maintain high inter-rater reliability. The analysis of the "within teacher" scores across classroom observations was encouraging. The average standard deviation across the three observations for each teacher was a low 0.15, and the average difference from the highest DOK class to the lowest DOK class was 0.35. While these numbers do not necessarily

add reliability to the data, they do indicate that teachers' TP scores seemed to be consistent.

Besides the generally low DOK level of the overall teacher practice variable, the other significant finding was the relationship of this variable to student outcomes. All four data analysis techniques—the correlation analysis, regression, multiple regression, and hierarchical linear modeling techniques—supported the predictive nature of the teacher practice variable on student achievement. Not only was the regression coefficient statistically significant, but the coefficient magnitude was relatively large, ranging from 4.3 to 8.0 across the three regression results. Assuming the HLM technique was the most appropriate analysis for this study, one could predict that for every increase of one DOK level in a teacher's classroom activity, the EOCA score of a student in this classroom would increase by 5.34 points. Since the End-of-Course Assessment had a high score of 40, the increase predicted by the TP variable of one additional DOK level results in a 13.35% increase of the total scale (over 1/8th of the total possible points). This represents a meaningful increase in scores as predicted by the teacher-level variable. Since the range of DOK values is limited to DOK 1 through DOK 3, a teacher may not be able to increase this score by one full point, but the findings show that the DOK 3 level adds much to student learning compared to DOK 1. Not only did the cognitive complexity of teacher practices seem to make a difference in student outcomes, this difference was a large one.

Along with the teacher practice variable, the number of years of experience also revealed significant predictive coefficients, though the coefficient was not as large ($B = 0.18$). Therefore, a student taught by a teacher with 20 years of experience would be

expected to score 3.6 points higher on the EOCA than a student taught by a novice teacher. That a professional will become more effective in his/her career with more experience is widely accepted. The findings of this study provide empirical evidence for this assumption in the case of these teachers. When the effects of both the TP variable and the years of experience of the teacher were combined in the HLM model, the total unexplained variance of student achievement scores due to second-level variables was reduced by 68.8%. While both the HLM and multiple regression models suggested that most of the variance of student achievement still remained at the student level, the variance analysis was still informative. The teacher practice variable explained why some classes did better than others after controlling for student readiness. When searching for meaningful second-level variables to consider, the cognitive DOK levels of a teacher's instruction should be included as one of the contributing factors to improving student scores.

Conclusion

Finding ways to improve student achievement scores has been the focus of educational research for decades and has cost millions of dollars. At this point in history, educators are still hard pressed to determine which teacher qualities will lead to improved achievement scores and deeper comprehension levels for students of all economic levels. This study looked specifically at a teachers' beliefs in their own capabilities that, when combined with other teacher qualities such as teacher knowledge and teacher practices, might predict student achievement scores. Much research has been conducted on the individual parts of this picture, and this study attempted to add to that body of knowledge by analyzing these variables simultaneously. Although the teacher efficacy and teacher

knowledge components did not show positive results, the significant and relatively large coefficient of the teacher practice variable added further evidence that the DOK level of the teacher's activity in the classroom should be a focus of further research.

America is searching for solutions to deficiencies in the mathematical knowledge of recent generations of high school students and graduates. The results of this study yield two main conclusions, and the combined effect of these conclusions is somewhat alarming. The first is that the cognitive complexity of a teacher's classroom practices matter. Higher cognitive complexity of classroom activities lead to higher student achievement scores. This conclusion is perhaps not new, but has been verified by the particularly unique methodology and measures of this study. The second conclusion is that teachers are not adequately engaging students in high DOK levels of thinking. Only a fraction of the teachers in the sample reached a DOK 2 level of cognitive complexity in their classroom activities, and most fell below this. Student achievement in mathematics on the national level is obviously lacking, but the precise reasons for this performance are not so evident. This study has empirically verified one reason scores may be low, and along with that, has identified one area for focused improvement. Future research is warranted both on how to bring awareness to teachers of the importance of high cognitive complexity classroom practices as well as methods to actually get them to make the changes necessary to their own teaching practices.

Why is it that teachers do not teach mathematics at cognitively high levels? It is likely that the teachers themselves do not have a high cognitive understanding of mathematics or, more appropriately, mathematical knowledge for teaching (the average GAST score was about 55% for the teachers in the sample). Perhaps the traditional "drill

and kill” methods of mathematics instruction that focus predominately on skills still pervade the educational structures that produce future mathematics teachers. If a student’s understanding of mathematics depends on the depth of their teacher’s understanding, then how will students ever break out of this cycle? How will a student who is to become a future teacher somehow “outperform” the teacher when it comes to high cognitive thinking? It’s possible that even the sharpest mathematics students learn how to solve a myriad of highly complex problems without having the faintest idea of the rich, theoretical concepts that underlie even the simplest of mathematical operations.

How might this cyclic process be broken? Appropriate, cognitively challenging mathematics problems is one answer. At what level of mathematics instruction? At all levels. But teachers themselves are the first place to start. Today’s teachers need to be made aware of the importance of engaging students in high cognitive tasks. They need to become aware of the current low level of mathematics learning of which they themselves are products, and they need to recognize how the current system is a closed one, and it will not change unless they themselves change and the cycle is broken. They must be taught what high cognitive tasks look like and how they differ from low level procedures. This awareness and knowledge of high cognitive instruction must be delivered through professional development opportunities and through mathematics and mathematics education courses in their preparation programs. What underlies this cultural change is ultimately the recognition by mathematicians and mathematics educators that, for teachers, a knowledge domain called “mathematical knowledge for teaching” is necessary for quality teaching. This knowledge domain encompasses the knowledge of

how to identify and implement high cognitive tasks, recognize and appreciate the beauty of the mathematics, and understand the learning process itself that goes along with them.

No doubt some mathematicians or mathematics educators will neither desire nor have the time to incorporate these ideas alongside the typical old school way of doing things. And some district administrators will prefer the stand and deliver methods of teaching that seemed to work fine for them and their generation of learners. Parents will echo their sentiments. Finally, some school systems will no doubt prefer textbooks and curriculum materials that have become a trademark of their districts and believe that what produces low mathematics achievement are in fact the new curricula that deviate from them, “We have to hold on to the tried and true!” And in the meantime, the memorization cycle continues, as the other industrial nations’ mathematics competency gets further and further ahead and out of reach.

All of this again points to the teacher as the first spark in the catalyst of change. Teachers must recognize how the current system of low-level instruction actually hinders a learner’s progress and perpetuates this endless cycle of low level thinking. Then, by learning to recognize, create, and implement appropriate, interesting, and high level tasks, teachers may begin the process of change toward improved student achievement, and deeper, conceptual understanding, reasoning, and problem solving.

On a final note, refining the GAST Observation Instrument is especially important in the current educational climate as administrators and policy makers strive to find methods to evaluate teachers in a balanced way. In fact, the explosive conflict that led to the recent teacher strike in the Chicago public school district was principally about how teachers were evaluated (Tareen, 2012). The GAST Observation Instrument was never

intended to be used as an evaluative tool and should not be used for that purpose. Nevertheless, looking at a teacher's actions through the lens of the "teaching moves" structure proposed by Cooney, Davis, and Henderson (1975) coupled with Webb's DOK levels presents an interesting approach to capturing classroom activity. Perhaps this approach could add a richness and complexity not present in current evaluation methods, which rely heavily on student achievement gains rather than teachers' actions. Further research is especially warranted during this crucial time of economic and political unrest to produce an evaluation tool that truly captures effective teaching. As the beginning steps of that process, the design, methods, and results of this study are particularly apropos.

REFERENCES

- Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education, 17*(2), 86-95.
- Allinder, R. M. (1995). An examination of the relationship between teacher efficacy and curriculum-based measurement and student achievement. *Remedial and Special Education, 16*(4), 247-254.
- Anderson, R. N., Greene, M. L., & Loewen, P. S. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. *Alberta Journal of Educational Research, 34*(2), 148-165.
- Ashton, P., Buhr, D., & Crocker, L. (1984). Teachers' sense of efficacy: A self- or norm-referenced construct? *Florida Journal of Educational Research, 26*(1), 29-41.
- Ashton, P., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.
- Ashton, P., Webb, R. B., & Doda, N. (1982). A study of teachers' sense of efficacy. Final report. (Vol. 1). Gainesville: University of Florida, Foundations of Education.
- Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics* (pp. 83-104). Westport, CT: Ablex.
- Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. *Handbook of research on teaching, 4*, 433-456.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall Inc.

- Bandura, A. (1995). *Self-efficacy in changing societies*. Cambridge ; New York: Cambridge University Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W H Freeman.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133-180.
- Benz, C. R., Bradley, L., Alderman, M. K., & Flowers, M. A. (1992). Personal teaching efficacy: Developmental relationships in education. *The Journal of Educational Research*, 85(5), 274 - 285.
- Berman, P., McLaughlin, M. W., Bass, G., Pauly, E., & Zellman, G. (1977). Federal programs supporting educational change; factors affecting implementation and continuation (Vol. VII). Santa Monica, CA: RAND Corporation.
- Bloom, B. S. (1956). *Taxonomy of educational objectives; the classification of educational goals* (1st ed.). New York: Longmans, Green.
- Bloom, B. S. (1978). New views of the learner: Implications for instruction and curriculum. 35(7), 563-576.
- Boardman, A. E., Davis, O. A., & Sanday, P. R. (1977). A simultaneous equations model of the educational process. *Journal of Public Economics*, 7(1), 23-49.
- Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, 40(2), 119-156.
- Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *Future of Children*, 17(1), 45-68.
- Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K., & Wisenbaker, J. M. (1978). Elementary school social climate and school achievement. *American Educational Research Journal*, 15(2), 301-318.
- Brophy, J. E. (1986). Teacher influences on student achievement. *American Psychologist*, 41(10), 1069-1077.
- Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Whittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: Simon & Shuster.
- Bruce, C. D., Esmonde, I., Ross, J., Dookie, L., & Beatty, R. (2010). The effects of sustained classroom-embedded teacher professional learning on teacher efficacy and related student achievement. *Teaching and teacher education*, 26(8), 1598-1608.

- Burley, W., Hall, B., Villeme, M., & Brockmeier, L. (1991). *Annual meeting of the American Educational Research Association*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Campbell, J. (1996). A comparison of teacher efficacy for pre and in-service teachers in Scotland and America. *Education Unlimited*, 117, 2 - 12.
- Caprara, G. V., Barbranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, 44(6), 473-490.
- Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education*, 60(4), 323-337.
- Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Wienfield, F., & York, R. (1966). *Equality of educational opportunity*. Washington: U.S. Dept. of Health, Education, and Welfare; Office of Education; U.S. Government Printing Office.
- Cooney, T. J., Davis, E. J., & Henderson, K. B. (1975). *Dynamics of teaching secondary school mathematics*. Boston: Houghton Mifflin.
- Cousins, J. B., & Walker, C. A. (2000). Predictors of educators' valuing of systematic inquiry in schools. *Canadian Journal of Program Evaluation*, 25-52.
- Cuoco, A. (2001). Mathematics for teaching. *Notices of the AMS*, 48(2), 168-174.
- Dale, A., Phillips, R., & Sianjina, R. (2011). *Influences of instructional leadership, transformational leadership, and the mediating effects of self-efficacy on student's achievement*. Paper presented at the American Institute of Higher Education - The 6th International Conference, Charleston, SC.
- Darling-Hammond, L. (2007). Building a system for powerful teaching and learning. In R. Whelting (Ed.), *Building a 21st century US education system* (pp. 65-74). Washington DC: National Commission on Teaching and America's Future.
- Davis, E. J. (1978). A model for understanding understanding in mathematics. *The Arithmetic Teacher*, 26(1), 13-17.
- Dillman, D. A., Smyth, J. D., Christian, L. M., & Dillman, D. A. (2009). *Internet, mail, and mixed-mode surveys : The tailored design method* (3rd ed.). Hoboken, N.J.: Wiley & Sons.
- Enochs, L. G., Scharmann, L. C., & Riggs, I. M. (1995). The relationship of pupil control to preservice elementary science teacher self-efficacy and outcome expectancy. *Science Education*, 79(1), 63-75.
- Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science and Mathematics*, 100(4), 194-202.

- Evans, B. R. (2011). Secondary mathematics teacher differences: Teacher quality and preparation in a new york city alternative certification program. *Mathematics Educator*, 20(2), 24-32.
- Fives, H. (2003). *Exploring the relationships of teachers' efficacy, knowledge, and pedagogical beliefs: A multimethod study*. (Ph.D. 3107210), University of Maryland, College Park, MD.
- Gage, N. L. (1978). *The scientific basis of the art of teaching*. Oxford England: Teachers College Press.
- Gardner, D. P. (1983). *A nation at risk*. Washington: National Commission on Excellence in Education.
- Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, 76(4), 569-582.
- Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy of Management Review*, 17(2), 183-211.
- Glickman, C. D., & Tamashiro, R. T. (1982). A comparison of first-year, fifth-year, and former teachers on efficacy, ego development, and problem solving. *Psychology in the Schools*, 19(4), 558-562.
- Goldhaber, D. (2002). Mystery of good teaching: Surveying the evidence on student achievement and teachers' characteristics. *Education Next*, 2, 50-55.
- Goldhaber, D., Brewer, D. J., & Anderson, D. J. (1999). A three-way error components analysis of educational productivity. *Education Economics*, 7(3), 199-208.
- Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66(3), 361-396.
- Greenwood, G. E. (1990). Relationships between four teacher efficacy belief patterns and selected teacher characteristics. *Journal of Research and Development in Education*, 23(2), 102-106.
- Guskey, T. R. (1981). Measurement of the responsibility teachers assume for academic successes and failures in the classroom. *Journal of Teacher Education*, 32(3), 44-51.
- Guskey, T. R. (1987). Context variables that affect measures of teacher efficacy. *Journal of Educational Research*, 81(1), 41-47.
- Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education: An International Journal of Research and Studies*, 4(1), 63-69.

- Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, 31(3), 627-643.
- Haney, J. J., Wang, J., Keil, C., & Zoffel, J. (2007). Enhancing teachers' beliefs and practices through problem-based learning focused on pertinent issues of environmental health science. *Journal of Environmental Education*, 38(4), 25-33.
- Hanushek, E. A. (1996). A more complete picture of school resource policies. *Review of Educational Research*, 66(3), 397-409.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19(2), 141-164.
- Hanushek, E. A. (2011). Recognizing the value of good teachers. *Education Week*, 30(27), 34-35.
- Harbison, R. W., & Hanushek, E. A. (1992). *Educational performance of the poor: Lessons from rural northeast brazil*. Oxford, England: Oxford University Press.
- Heneman, H. G., Kimball, S., Milanowski, A., & Wisconsin Center for Education Research, M. (2006). The teacher sense of efficacy scale: Validation evidence and behavioral prediction. Madison, WI: Wisconsin Center for Education Research.
- Henson, R. K. (2002). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist*, 37(3), 137-150.
- Hiebert, J. (1986). *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ Lawrence Erlbaum Associates, Inc.
- Hiebert, J. (1989). The struggle to link written symbols with understandings: An update. *Arithmetic Teacher*, 36(7), 38-44.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the national council of teachers of mathematics*. (pp. 65-97). New York: Macmillan.
- Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, 30(2), 393-425.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.
- Hoy, W. K., & Woolfolk, A. (1993). Teachers' sense of efficacy and the organizational health of schools. *The Elementary School Journal*, 93(4), 355-372.

- James, W. (1890). *The principles of psychology*. New York: Holt.
- Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, 23(1), 21-43.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (Third ed.). New York: Guilford press.
- Krauss, S., Brunner, M., Kunter, M., Baumert, J., Neubrand, M., Blum, W., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, 100(3), 716-725.
- Lee, V. E., Dedrick, R. F., & Smith, J. B. (1991). The effect of the social organization of schools on teachers' efficacy and satisfaction. *Sociology of Education*, 64(3), 190-208.
- Ma, L. (1999). *Knowing and teaching elementary mathematics : Teachers' understanding of fundamental mathematics in china and the united states*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Maguire, K. (2011). *The role of teacher efficacy in student academic achievement in mathematics*. Walden University.
- Maslow, A. H. (1954). *Motivation and personality* (1st ed.). New York: Harper.
- McCoach, D. B. (2010). Hierarchical linear modeling. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences*. New York: Routledge.
- McLaughlin, M. W., & Marsh, D. D. (1978). Staff development and school change. *Teachers College Record*, 80(1), 69-94.
- Meijer, C. J. W., & Foster, S. F. (1988). The effect of teacher self-efficacy on referral chance. *Journal of Special Education*, 22(3), 378-385.
- Metzler, J., & Woessmann, L. (2010). *The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation*. Mimeo: University of Munich.
- Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, 81(2), 247-258.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13(2), 125-145.

- Moore, W., & Esselman, M. (1992). *Teacher efficacy, power, school climate and achievement: A desegregating district's experience*. Paper presented at the Annual Meeting of American Educational Research Association, San Francisco.
- Mueller, D. J. (1986). *Measuring social attitudes: A handbook for researchers and practitioners*. New York: Teachers College Press.
- Muijs, D., & Reynolds, D. (2002). Being or doing: The role of teacher behaviors and beliefs in school and teacher effectiveness in mathematics, a sem analysis. *Journal of Classroom Interaction*, 37(2), 3-15.
- Mullens, J. E., Murnane, R. J., & Willet, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from belize. *Comparative Education Review*, 40(2), 139-157.
- Murray, H. B., & Staebler, B. K. (1974). Teacher's locus of control and student achievement gains. *12(4)*, 305-309.
- Newmann, F. M., Rutter, R. A., & Smith, M. S. (1989). Organizational factors that affect school sense of efficacy, community, and expectations. *Sociology of Education*, 62(4), 221-238.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578.
- Pajares, F. (2002). Overview of social cognitive theory and of self-efficacy. Retrieved October 11, 2011, from <http://www.emory.edu/EDUCATION/mfp/eff.html>
- Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis*, 30(2), 111-140.
- Parkay, F. W., Greenwood, G., Olejnik, S., & Proller, N. (1988). A study of the relationships among teacher efficacy, locus of control, and stress. *Journal of Research & Development in Education*, 21(4), 13-22.
- Peterson, P. E., Woessmann, L., Hanushek, E. A., & Lastra-Anadon, C. X. (2011). Globally challenged: Are u.S. Students ready to compete? The latest on each state's international standing in math and reading: Program on Education Policy and Governance, Harvard University.
- Podell, D. M., & Soodak, L. C. (1993). Teacher efficacy and bias in special education referrals. *Journal of Educational Research*, 86(4), 247-253.
- Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education*, 65(2), 150-167.

- Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education, 74*, 625-638.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.
- Roberts, J. K., & Henson, R. K. (2001). *A confirmatory factor analysis of a new measure of teacher efficacy: Ohio state teacher efficacy scale*. Paper presented at the annual meeting of the American Educational Research Association, Seattle.
- Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcome. *Journal of Educational Research, 74*(3), 185-190.
- Ross, J. A. (1992). Teacher efficacy and the effects of coaching on student achievement. *Canadian Journal of Education, 17*(1), 51-65.
- Ross, J. A. (1994). *Beliefs that make a difference: The origins and impacts of teacher efficacy*. Paper presented at the Annual Meeting of the Canadian Association for Curriculum Studies.
- Ross, J. A., Hogaboam-Gray, A., & Hannay, L. (2001). Effects of teacher efficacy on computer skills and computer cognitions of canadian students in grades k-3. *Elementary School Journal, 102*(2), 141-156.
- Rotter, J. B. (1954). *Social learning and clinical psychology*. Englewood Cliffs, NJ Prentice-Hall.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General & Applied, 80*(1), 1-28.
- Sanders, W. L. (2004). *A summary of conclusions drawn from longitudinal analyses of student achievement data over the past 22 years (1982-2004)*. Paper presented at the Governors Education Symposium, Asheville, NC.
- Schoon, K. J., & Boone, W. J. (1998). Self-efficacy and alternative conceptions of science of preservice elementary teachers. *Science Education, 82*(5), 553-568.
- Sciutto, M. J., Terjesen, M. D., & Bender Frank, A. S. (2000). Teachers' knowledge and misperceptions of attention-deficit/hyperactivity disorder. *Psychology in the Schools, 37*(2), 115-122. doi: 10.1002/(sici)1520-6807(200003)37:2<115::aid-pits3>3.0.co;2-5
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

- Silver, E. A., & Stein, M. K. (1996). The quasar project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education*, 30(4), 476-521.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research*, 75(3), 417-453.
- Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 1 - 7.
- Skemp, R. R. (1987). *The psychology of learning mathematics* (Expanded American ed.). Hillsdale, NJ: L. Erlbaum Associates.
- Smylie, M. A. (1988). The enhancement function of staff development: Organizational and psychological antecedents to individual teacher change. *American Educational Research Journal*, 25(1), 1-30.
- Soodak, L. C., & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *The Journal of Special Education*, 27(1), 66-81.
- Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, 33(2), 455-488.
- Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, 2(1), 50-80.
- Stein, M. K., Lane, S., & Silver, E. A. (1996). *Classrooms in which students successfully acquire mathematical proficiency: What are the critical features of teachers' instructional practice*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.
- Stein, M. K., Smith, M. S., Henningsen, M. A., Silver, E. A., & National Council of Teachers of Mathematics, I. R. V. A. (2000). *Implementing standards-based mathematics instruction: A casebook for professional development*. New York, NY: Teachers College Press.
- Stevens, J. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York: Routledge.
- Strauss, R. P., & Sawyer, E. A. (1986). Some new evidence on teacher and student competencies. *Economics of Education Review*, 5(1), 41-48.
- Tareen, S. (2012). Teacher evaluations at the center of chicago strike, *Associated Press*. Retrieved from <http://bigstory.ap.org/article/teacher-evaluations-center-chicago-strike-0>

- Tatto, M. T., Nielsen, H. D., Cummings, W., Kularatna, N. G., & Dharmadasa, K. H. (1993). Comparing the effectiveness and costs of different approaches for educating primary school teachers in Sri Lanka. *Teaching and Teacher Education, 9*(1), 41-64.
- Tchoshanov, M. A. (2011). Relationship between teacher knowledge of concepts and connections, teaching practice, and student achievement in middle grades mathematics. *Educational Studies in Mathematics, 76*(2), 141-164.
- Tschannen-Moran, M., & Barr, M. (2004). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in Schools, 3*(3), 189-209.
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783-805.
- Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education, 23*(6), 944-956.
- Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202.
- Usiskin, Z. (1982). Van Hiele levels and achievement in secondary school geometry (*Final report of the Cognitive Development and Achievement in Secondary School Geometry Project*). Chicago: University of Chicago, Department of Education.
- Usiskin, Z. (2001). Teachers' mathematics: A collection of content deserving to be a field. *The Mathematics Educator, 6*(1), 86-98.
- Watson, J. (1930). *Behaviorism*. New York: W.W. Norton & Company.
- Watson, S. (1991). *A study of the effects of teacher efficacy on the academic achievement of third grade students in selected elementary schools in South Carolina*. South Carolina State College.
- Webb, N. L. (1997). Determining alignment of expectations and assessments in mathematics and science education. *National Institute for Science Education Brief, 1*(2), 1-8.
- Webb, N. L. (1999). Alignment of science and mathematics standards and assessments in four states (*Research Monograph No. 18*). Madison, WI: University of Wisconsin-Madison: National Institute for Science Education.
- Webb, N. L. (2002). *An analysis of the alignment between mathematics standards and assessments for three states*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461-481.

Woolfolk, A., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology*, 82(1), 81-91.

Woolfolk, A., Rosoff, B., & Hoy, W. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and teacher education*, 6(2), 137-148.

APPENDICES

Appendix A

Table A1. Relationship between Self-Efficacy and Self-Concept (Pajares, 2002)

Construct	Attributes	Definitions
Self-Efficacy	Definition	The <i>belief</i> in one's capabilities to organize and execute the courses of action required to manage prospective situations.
	Effects	It influences choices, effort, persistence, and feelings.
	Sources	Mastery experience Vicarious experience Verbal persuasion Physiological state
Self-Concept	Definition	A cognitive appraisal, integrated across various dimensions that individuals attribute to themselves, typically accompanied by self-evaluative judgment of self-worth.
Contrasts with Self-Concept/Esteem	<u>Self-Efficacy</u> Context specific Assessment of competence About performance of a specific task	<u>Self-Concept/Esteem</u> Integrated across domains Cognitive appraisal About being/feeling

Appendix B

Examples of Items Measuring Content Knowledge for Teaching Mathematics

1. Mr. Allen found himself a bit confused one morning as he prepared to teach. Realizing that 10 to the second power equals 100 ($10^2 = 100$), he puzzled about what power of 10 equals 1. He asked Ms. Berry, next door. What should she tell him? (Mark [X] One answer.)

- a) 0
- b) 1
- c) Ten cannot be raised to any power such that 10 to that power equals 1
- d) -1
- e) I'm not sure

2. Imagine that you are working with your class on multiplying large numbers. Among your students' papers, you notice that some have displayed their work on the following ways:

Student A	Student B	Student C
35	35	35
<u>x 25</u>	<u>x 25</u>	<u>x 25</u>
125	175	25
<u>+ 75</u>	<u>+ 700</u>	150
875	875	100
		<u>+ 600</u>
		875

Which of these students would you judge to be using a method that could be used to multiply any two whole numbers?

	Method would work for all whole numbers	Method would NOT work for all numbers	I'm not sure
Method A	1	2	3
Method B	1	2	3
Method C	1	2	3

(Hill et al., 2005, p. 402)

Appendix C

Final Blueprint for the Geometry Assessment for Secondary Teachers (GAST)

Section 1: Teacher knowledge of mathematics		% items per form
		30
1	The teacher recognizes and describes appropriate demonstrations, interpretations, analogies, and justifications to introduce and develop mathematical skills and procedures.	10
2	The teacher recognizes and describes appropriate definitions, representations, examples, distinguishing examples, non-examples, counter-examples, and necessary and sufficient conditions to introduce and develop mathematical concepts.	10
3	The teacher recognizes and describes meaningful connections (lateral, upward, downward) within and among mathematics content.	5
4	The teacher recognizes and constructs meaningful mathematical models of real world situations.	5
Section 2: Teacher knowledge of geometric reasoning and problem solving		% items per form
		25
1	The teacher solves non-routine problems, including real world applications, in geometry.	10
2	The teacher analyzes and constructs synthetic, transformational, and analytical proofs; and recognizes valid and invalid arguments (e.g., reasoning by converse, proofs by contradiction, negating, arguing the contrapositive, non-example).	10
3	The teacher analyzes and justifies geometric formulas.	5

Section 3: Teacher knowledge of student learning		% items per form
		45
1	The teacher recognizes and describes strategies and activities that promote student reasoning and problem solving (e.g., questioning, posing a problem, offering a conjecture, describing an application).	10
2	The teacher anticipates, recognizes, describes, assesses, and addresses correct and incorrect elements of student responses (e.g., skills, concepts, reasoning).	10
3	The teacher recognizes, describes, and assesses critical student prerequisite knowledge..	10
4	The teacher recognizes and constructs mathematics assessment tasks at different cognitive levels.	5
5	The teacher recognizes and describes advantages and limitations of using digital technologies (e.g., interactive geometry software, graphing calculators, virtual manipulatives, other internet resources) to foster student learning.	5
6	The teacher recognizes and describes advantages and limitations of using physical models (e.g., solids, paper folding) and tools (e.g., compass, straight edge, protractor) to foster student learning.	5

Appendix D

Example GAST Items with Possible MKT Classifications

1. Identify four ways a teacher can use these images of pyramids to introduce the topics of surface area and volume. (Knowledge of Content and Teaching/Knowledge of Content and Students)



2. In promoting reasoning about triangle congruence, a teacher might do the following classroom activities with students: (Knowledge of Content and Teaching)

1. Have students draw a triangle with two sides of given length and included angle of given measure. Have students cut out the triangle and compare it to other students' triangles and describe what they notice about the triangles.
2. Draw two congruent triangles and ask students to explore the triangles to discover the relationship between the triangles.
3. Have students explore pairs of triangles to identify which three corresponding congruent measurements guarantee triangle congruence.
4. Draw two congruent triangles on a coordinate grid, and ask students to measure the length of three sides in one triangle and the corresponding sides in the other triangle to illustrate SSS congruence. Ask students to describe the relationship about the corresponding angles of the triangles.

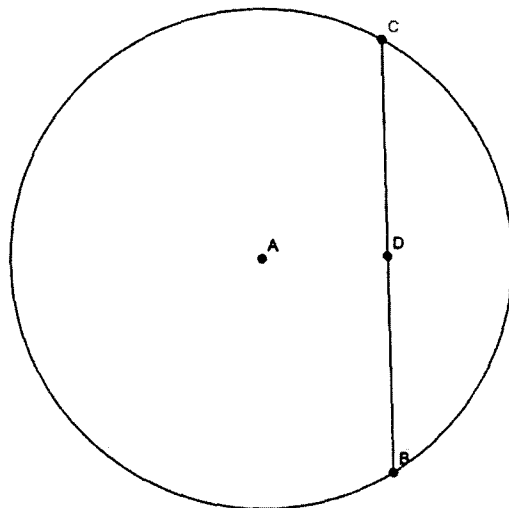
How should these activities be ordered so that they represent a developmentally appropriate learning sequence for students?

- A. 1, 2, 3, then 4
- B. 2, 4, 1, then 3
- C. 3, 2, 4, then 1
- D. 4, 1, 2, then 3

3. Explain three prerequisite concepts and skills that students need in order to find the area of a sector of a circle. Describe a formative assessment task that a teacher could use to determine students' readiness for this skill. (Knowledge of Content and Teaching)
4. Which topic would **NOT** be prerequisite knowledge for students beginning to learn triangle similarity? (Knowledge of Content and Teaching)
- Parallel lines
 - Properties of triangles
 - Pythagorean theorem
 - Solving proportions

5. Refer to the diagram below.

(Knowledge of Content and Students/ Knowledge of Content and Teaching)



The circle has a diameter of 12 inches. D is the midpoint of line segment \overline{BC} , and $|\overline{AD}|=3$. A student reasons that the area of the segment of the circle cut off by \overline{BC} must be one fourth of the area of the full circle.

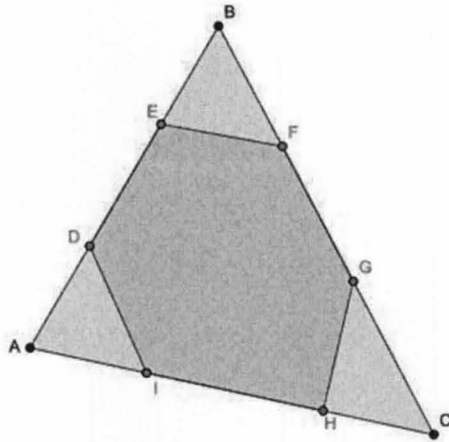
Describe one activity that a teacher could use to help the student understand and correct the error in the student's thinking.

6. A spherical planet has a spherical inner core with a diameter half of that of the planet.

What fraction of the volume of the planet is the volume of its inner core? (Common Content Knowledge)

- A. $\frac{1}{8}$
- B. $\frac{1}{4}$
- C. $\frac{1}{3}$
- D. $\frac{1}{2}$

7. Refer to the diagram below. (Common Content Knowledge)



$\triangle ABC$ has an area equal to 120 cm^2 . Each of the sides is trisected with points D, E, F, G, H, I.

- I. What is the area of hexagon DEFGHI?

- A. 30 cm^2
- B. 40 cm^2
- C. 80 cm^2
- D. 90 cm^2

Appendix E

Teacher Sense of Efficacy Scale – Long Form

TEACHER BELIEFS	How much can you do?								
<p>Directions: This questionnaire is designed to help us gain a better understanding of the kinds of things that create difficulties for teachers in their school activities. Please indicate your opinion about each of the statements below. Your answers are confidential.</p>	Nothing	Very Little	Some Influence	Quite a bit	A Great Deal				
1. How much can you do to get through to the most difficult students?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2. How much can you do to help your students think critically?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
3. How much can you do to control disruptive behavior in the classroom?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
4. How much can you do to motivate students who show low interest in school work?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
5. To what extent can you make your expectations clear about student behavior?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
6. How much can you do to get students to believe they can do well in school work?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
7. How well can you respond to difficult questions from your students?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
8. How well can you establish routines to keep activities running smoothly?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
9. How much can you do to help your students value learning?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
10. How much can you gauge student comprehension of what you have taught?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
11. To what extent can you craft good questions for your students?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
12. How much can you do to foster student creativity?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
13. How much can you do to get children to follow classroom rules?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
14. How much can you do to improve the understanding of a student who is failing?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
15. How much can you do to calm a student who is disruptive or noisy?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
16. How well can you establish a classroom management system with each group of students?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
17. How much can you do to adjust your lessons to the proper level for individual student?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
18. How much can you use a variety of assessment strategies?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
19. How well can you keep a few problem students from ruining an entire lesson?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
20. To what extent can you provide an alternative explanation or example when students are confused?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
21. How well can you respond to defiant students?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
22. How much can you assist families in helping their children do well in school?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
23. How well can you implement alternative strategies in your classroom?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
24. How well can you provide appropriate challenges for very capable students?	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

Additional Survey Item to Account for delayed administration of the TSES.
Please select any of the following which apply to you.

1. I am no longer teaching.
2. I have changed schools this year.
3. My principal has changed this year.
4. New administration (such as principal) has had a large effect on my teaching abilities this year.
5. I am teaching the same geometry class I taught last year.
6. The classes I am teaching this year are much more difficult or much easier to teach than last year.
7. I feel my beliefs in my teaching abilities are about the same as they were last year.

Table E1.
Validity Correlations for the TSES

Validity Correlations								
	TSES	Instruct	Manage	Engage	Rand 1	Rand 2	GTE	PTE
TSES		0.89**		0.87**	0.18**	0.53**	0.16**	0.64**
Instructional Strategies	0.84**			0.70**	0.07	0.45**	0.06	0.62**
Classroom Management	0.79**	0.46**		0.58**	0.29**	0.46**	0.30**	0.45**
Student Engagement	0.85**	0.61**	0.50**		0.11*	0.47**	0.06	0.58**
Rand 1	0.18**	0.08*	0.26**	0.11**		0.23**	0.65**	0.12*
Rand 2	0.52**	0.45**	0.39**	0.45**	0.23**		0.13*	0.65**
General Teaching Efficacy	0.16**	0.08**	0.26**	0.06**	0.65**	0.13*		0.07
Personal Teaching Efficacy	0.61**	0.60**	0.37**	0.56**	0.12*	0.65**	0.07	

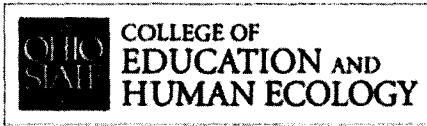
(Tschannen-Moran & Hoy, 2001, p. 802)

Table E2

Factor Loadings for the TSES Items

Teacher Efficacy Scale (TSES)	24 items	12 items
<i>Factor 1: Efficacy for instructional strategies</i>		
1. To what extent can you use a variety of assessment strategies?	0.72	0.73
2. To what extent can you provide an alternative explanation or example when students are confused?	0.70	0.75
3. To what extent can you craft good questions for your students?	0.68	0.63
4. How well can you implement alternative strategies in your classroom?	0.66	0.73
5. How well can you respond to difficult questions from your students?	0.66	
6. How much can you do to adjust your lessons to the proper level for individual students?	0.59	
7. To what extent can you gauge student comprehension of what you have taught?	0.57	
8. How well can you provide appropriate challenges for very capable students?	0.55	
<i>Factor 2: Efficacy for classroom management</i>		
9. How much can you do to control disruptive behavior in the classroom	0.78	0.83
10. How much can you do to get children to follow classroom rules?	0.69	0.66
11. How much can you do to calm a student who is disruptive or noisy?	0.66	0.63
12. How well can you establish a classroom management system with each group of students?	0.66	0.61
13. How well can you keep a few problem students from ruining an entire lesson?	0.62	
14. How well can you respond to defiant students?	0.61	
15. To what extent can you make your expectation clear about student behavior?	0.53	
16. How well can you establish routines to keep activities running smoothly?	0.50	
<i>Factor 3: Efficacy for student engagement</i>		
17. How much can you do to get students to believe they can do well in schoolwork?	0.75	0.75
18. How much can you do to help your students value learning?	0.70	0.69
19. How much can you do to motivate students who show low interest in schoolwork?	0.66	0.64
20. How much can you assist families in helping their children do well in school?	0.63	0.62
21. How much can you do to improve the understanding of a student who is failing?	0.57	
22. How much can you do to help your students think critically?	0.56	
23. How much can you do to foster student creativity?	0.50	
24. How much can you do to get through to the most difficult students?	0.47	

Permission to use the TSES.



ANITA WOOLFOLK HOY, PH.D.

**PROFESSOR
PSYCHOLOGICAL STUDIES IN EDUCATION**

Dear

You have my permission to use the *Teachers' Sense of Efficacy Scale* in your research. A copy of both the long and short forms of the instrument as well as scoring instructions can be found at:

<http://www.coe.ohio-state.edu/ahoy/researchinstruments.htm>

Best wishes in your work,

Anita Woolfolk Hoy, Ph.D.
Professor

Appendix F

Glossary for GAST Lesson Analysis

Mathematics Concepts

A kind of mathematical subject matter; the most basic learnable object. A decision rule which, when applied to the description of an object, specifies whether or not a name can be applied (Hunt, Marin, and Stone, 1966). A concept has a *name* that represents a set of objects with common attributes. Examples of mathematics concepts include: rational number, rectangle, theorem, micrometer, sine function, pi, assumption, similarity, volume.

Moves for Teaching Concepts

Definition: Teacher provides students a definition of the concept or leads students to know a definition of the concept. Precision is often a critical criterion for definitions of mathematical concepts. **(DOK 1 if stated, DOK 2 if developed)**

Example: Teacher gives or asks for an example of the concept. The example is a member of the set denoted by the concept. **(DOK 2)**

Nonexample: Teacher gives or asks for a nonexample of the concept. The nonexample often has some properties of the concept, but not all. **(DOK 2)**

Counterexample: Teacher provides or asks a student to provide a counterexample to an incorrect assertion or definition. **(DOK 2)**

Sufficient Condition: Teacher states or asks students to state a sufficient condition of the concept. Terms and phrases like *if, provided that, because, since, or for the reason that* often signal statements of sufficient conditions. **(DOK 2)**

Necessary Condition: Teacher states or asks students to state a necessary condition of the concept. Terms like *has to or must* often signal statements of necessary conditions. It is often used to prevent misconceptions. **(DOK 2)**

Comparing or Contrasting: Teacher identifies or asks students to identify similarities or differences between the concept and other concepts. **(DOK 2)**

Mathematics Generalizations : Mathematics statements that hold true over a set of mathematical objects or concepts. Generalizations include definitions, statements of necessary and/or sufficient conditions, axioms, theorems, corollaries, propositions, formulas, or rules. The statements may use words, symbols, or combinations of both.

Moves for Teaching Generalizations

Assertion: Teacher asserts or asks students to assert the generalization in writing, words, or symbols. **(DOK 1)**

Application: Teacher applies or asks to students to apply the generalization. This move requires deduction in that the teacher or student must analyze a situation or problem and decide which generalization or generalizations are relevant. The teacher, by means of questions, exercises, or problems seeks to get the students to apply the generalization either by itself or with other generalizations. **(DOK 2)**

Instance: Teacher employs or asks students to employ an example of the generalization. It usually requires replacing variables or words by constants or numbers. Instances of generalizations are almost always sentences, whereas examples of concepts are generally elements of a set. **(DOK 2)**

Paraphrase: Teacher states or asks students to state the meaning of the generalization in different words. **(DOK 2)**

Counterexample: Teachers provides or asks students to provide a counterexample to a false generalization. **(DOK 2)**

Analysis: Teacher discusses or asks students to discuss the components and logic of the generalization. The discussion may focus on the *if-then* or *if-and-only-if* propositions of the generalization. **(DOK 3)**

Justification: Teacher provides or asks students to provide evidence, through instances or reasons, that the generalization is true. Four kinds of justification moves can be used: acceptance by authorities (mathematicians); deductive argument; through instances; search for counterexamples. **(DOK 3)**

Mathematics Skills

Mathematics knowledge that reflects how to do something. Examples include squaring binomials, interpolating, solving equations, bisecting line segments. Student must know how to perform mathematics skills with proficiency and understanding. Proficiency in skill requires practice.

Assertion: Teacher asserts or asks students to assert how to perform the skill. The steps in performing the skill may be in words or written. **(DOK 1)**

Demonstration: Teacher demonstrates or asks students to demonstrate how to perform the skill. Teachers clarify the skill through demonstrations. **(DOK 1)**

Interpretation: Teacher explains or asks students to explain the meaning of the terms in the procedure or how to perform each step in the skill. Students often do not understand key concepts or remember facts or subskills within the skill. **(DOK 2)**

Analogy: Teacher compares or asks students to compare the skill to a previously learned skill. **(DOK 2)**

Justification: Teacher provides evidence or asks students to provide evidence that the procedures of the skill are valid. Possible methods for justification include: (1) always obtaining a correct result (Does it work?) and (2) steps are predicated on acceptable generalizations (Can I prove it always works?). **(DOK 3)**

GAST Classroom Observation Instrument

Teacher ID # _____	School _____	District _____	Date _____
Period _____	Textbook/Resources _____	Pages _____	
Reviewer _____	Date _____	Content (Circle one): Area/Volume Similarity/Congruence Other _____	
Chapter Title _____	Instruction Start Time _____		_____
Lesson Title _____	Instruction End Time _____		_____

Concept/ Instructional Moves						Generalization/ Instructional Moves						Skill/Procedure/ Instructional Moves						Teacher Tasks Demonstrated			
																		DOK 1	DOK 2	DOK 3	
Def	Ex	Cx	Nx	Prp	CAC	State	App	Exam	Para	Co Ex	Anlys	Just	Stat	Dm	App	Intrp	Anlgy	Just			
																		Student Tasks Attempted			
																		DOK 1	DOK 2	DOK 3	
Def	Ex	Cx	Nx	Prp	CAC	State	App	Exam	Para	Co Ex	Anlys	Just	Stat	Dm	App	Intrp	Anlgy	Just			
Def	Ex	Cx	Nx	Prp	CAC	State	App	Exam	Para	Co Ex	Anlys	Just	Stat	Dm	App	Intrp	Anlgy	Just			
																		Student Exers - Skills	Student Miscgn Concept	Student Exers Re/ FS	
																		Anticipated			
																		Demonstrated/Explained			
Def	Ex	Cx	Nx	Prp	CAC	State	App	Exam	Para	Co Ex	Anlys	Just	Stat	Dm	App	Intrp	Anlgy	Just	Identified in Student Work		
																		Representations/Models			
																		Pictorial Representations			
																		Physical Representations			
Def	Ex	Cx	Nx	Prp	CAC	State	App	Exam	Para	Co Ex	Anlys	Just	Stat	Dm	App	Intrp	Anlgy	Just	Real-world Representations		

Appendix H

Exploratory Factor Analysis

Descriptive statistics for the 24-item TSES are given in Table H1. Though 72 teachers responded to the survey, cases with missing data were deleted listwise for this analysis, leaving only 68. The KMO showed the data factorable, and Bartlett's test verified strong correlations among items. Principal Component Analysis (Table H2) produced 5 factors which accounted for 67.4% of the variance. This is also supported by the Scree Plot. Factor loadings on these five factors are shown below in Table H3.

Table H1

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Q1	6.03	1.486	68
Q2	7.06	1.105	68
Q3	7.37	1.233	68
Q4	5.91	1.484	68
Q5	8.35	.958	68
Q6	6.94	.991	68
Q7	8.01	.938	68
Q8	7.90	1.248	68
Q9	6.62	1.037	68
Q10	7.65	.989	68
Q11	7.71	1.008	68
Q12	6.49	1.139	68
Q13	7.47	1.000	68
Q14	6.43	1.150	68
Q15	6.96	1.139	68
Q16	7.49	1.086	68
Q17	7.00	1.197	68
Q18	7.03	1.315	68
Q19	6.96	1.202	68
Q20	8.09	.958	68
Q21	7.06	1.314	68
Q22	5.99	1.607	68
Q23	6.88	1.240	68
Q24	7.46	1.190	68

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Meas. of Sampling Adequacy.		0.851
Bartlett's Test of Sphericity	Appr. Chi-Square	989.724
	df	276
	Sig.	0

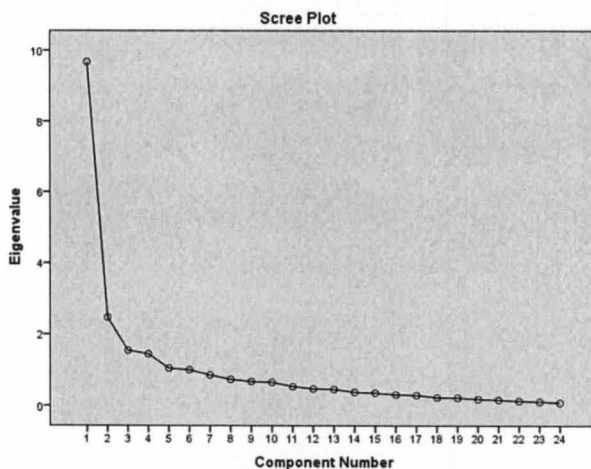


Table H2

Principle Component Analysis

	Total Variance Explained								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.676	40.315	40.315	9.676	40.315	40.315	4.475	18.647	18.647
2	2.462	10.259	50.574	2.462	10.259	50.574	3.529	14.704	33.351
3	1.542	6.425	56.999	1.542	6.425	56.999	3.336	13.899	47.251
4	1.442	6.009	63.008	1.442	6.009	63.008	2.500	10.416	57.667
5	1.044	4.349	67.357	1.044	4.349	67.357	2.326	9.690	67.357
6	.996	4.151	71.508						
7	.854	3.560	75.067						
8	.732	3.049	78.117						
9	.675	2.812	80.928						
10	.653	2.719	83.648						
11	.536	2.233	85.881						
12	.472	1.965	87.846						
13	.457	1.904	89.750						
14	.376	1.568	91.318						
15	.355	1.479	92.797						
16	.308	1.284	94.081						
17	.292	1.218	95.300						
18	.231	.963	96.263						
19	.222	.926	97.189						
20	.181	.753	97.941						
21	.163	.679	98.620						
22	.133	.554	99.174						
23	.113	.472	99.646						
24	.085	.354	100.000						

Extraction Method: Principal Component Analysis.

Table H3

Rotated Component Matrix

	Component				
	1	2	3	4	5
Q1	.179	.156	.182	.780	.208
Q2	-.026	.205	.334	.373	.635
Q3	.706	.259	.159	.039	.286
Q4	-.064	.111	.529	.448	.168
Q5	.348	.318	-.031	.085	.697
Q6	.217	.292	.717	-.131	.052
Q7	.258	.751	.160	.023	.201
Q8	.413	.279	-.119	-.054	.575
Q9	.380	.039	.635	.235	.254
Q10	.303	.068	.211	.495	.469
Q11	.171	.655	.189	.338	.318
Q12	.211	.266	.552	.251	.086
Q13	.697	.064	.221	.042	.347
Q14	.257	.213	.077	.816	-.049
Q15	.881	.029	.044	.271	.011
Q16	.675	.094	.157	.108	.462
Q17	.296	.401	.524	.347	.105
Q18	.099	.666	.270	.224	.129
Q19	.820	.253	.176	.125	.034
Q20	.196	.797	.155	.066	.113
Q21	.763	.288	.153	.188	-.003
Q22	.064	.114	.708	.077	-.096
Q23	.213	.634	.573	.204	.051
Q24	-.023	.434	.454	.140	.354

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Original Design

Efficacy in Student Engagement: Items 1, 2, 4, 6, 9, 12, 14, 22

Efficacy in Instructional Strategies: Items 7, 10, 11, 17, 18, 20, 23, 24

Efficacy in Classroom Management: Items 3, 5, 8, 13, 15, 16, 19, 21

Exploratory Factor Loadings $\geq .6$

Factor 1 Q3, Q13, Q15, Q16, Q19, Q21

Factor 2 Q7, Q11, Q18, Q20, Q23

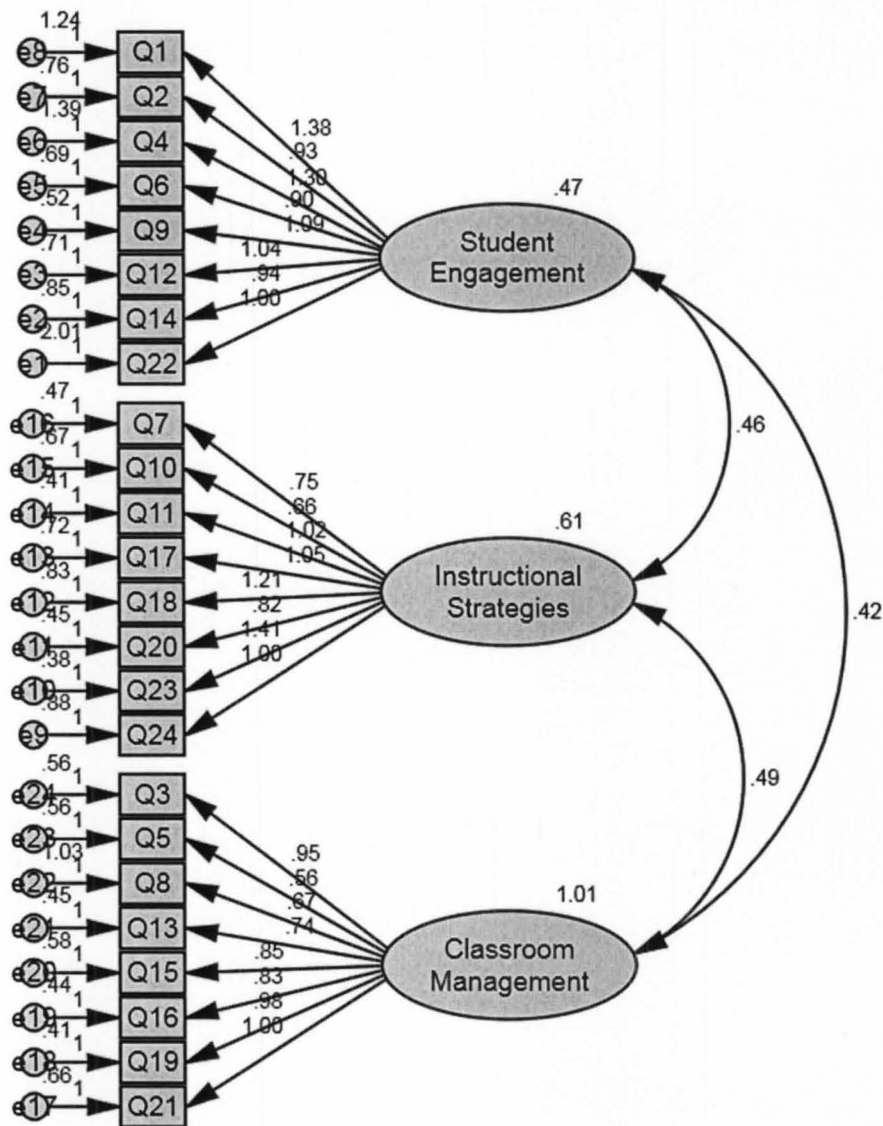
Factor 3 Q6, Q9, Q22 (Q4 = .529, Q12 = .552, Q17 = .524, Q24 = .454)

Factor 4 Q1, Q14

Factor 5 Q2, Q5

Confirmatory Factor Analysis

The model below represents the three factor structure of the TSES. Model fit statistics are listed on the following page. Although opinions differ about which statistics are required to test for model fit, Kline (2011, pp. 209-211) recommends reporting the Chi-squared test, the RMSEA, the CFI, and the SRMR. This model yielded a $\chi^2(249) = 1.605, p = .000$ rejecting the fit hypothesis. The CFI = .830 (< .90), and the SRMR = .0887 (> .80), both of which suggest a poor fit. Finally, the RMSEA = .092, with 90% confidence interval (.075, .109). Both the lower and upper bounds suggest poor fit (.075 > .05, and .109 > .100).



Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	51	399.691	249	.000	1.605
Saturated model	300	.000	0		
Independence model	24	1160.189	276	.000	4.204

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.120	.694	.632	.576
Saturated model	.000	1.000		
Independence model	.493	.224	.157	.206

Baseline Comparisons

Model	NFI	RFI	IFI	TLI	CFI
	Delta1	rho1	Delta2	rho2	
Default model	.655	.618	.835	.811	.830
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.902	.591	.748
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	150.691	100.040	209.261
Saturated model	.000	.000	.000
Independence model	884.189	782.741	993.170

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	5.629	2.122	1.409	2.947
Saturated model	.000	.000	.000	.000
Independence model	16.341	12.453	11.025	13.988

RMSEA				
Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.092	.075	.109	.000
Independence model	.212	.200	.225	.000

AIC				
Model	AIC	BCC	BIC	CAIC
Default model	501.691	557.126	617.801	668.801
Saturated model	600.000	926.087	1283.000	1583.000
Independence model	1208.189	1234.276	1262.829	1286.829

ECVI				
Model	ECVI	LO 90	HI 90	MECVI
Default model	7.066	6.353	7.891	7.847
Saturated model	8.451	8.451	8.451	13.043
Independence model	17.017	15.588	18.552	17.384

HOELTER		
Model	HOELTER	HOELTER
	.05	.01
Default model	51	54
Independence model	20	21

Minimization: .016
 Miscellaneous: .359
 Bootstrap: .000
 Total: .375

Appendix I

Additional HLM Models

HLM Model, Teacher Practice with No Covariate

Level-1 Model

$$EOCA_{ij} = \beta_{0j} + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01} * (TP_j) + u_{0j}$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{01} * (TP)_j + u_{0j} + r_{ij}$$

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	2.009724	5.334792	0.377	70	0.708
TP, γ_{01}	7.985866	3.156241	2.530	70	0.014

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	5.17484	26.77902	70	1414.19395	<0.001
level-1, r	5.36378	28.77017			

Deviance = 9973.277780

HLM Model, Covariate Only

Level-1 Model

$$EOCA_{ij} = \beta_{0j} + \beta_{1j} * (EGT_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{10} * EGT_{ij} + u_{0j} + r_{ij}$$

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	1.192523	0.758084	1.573	71	0.120
For EGT slope, β_1					
INTRCPT2, γ_{10}	0.992523	0.058337	17.014	1503	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	3.51531	12.35737	71	910.24508	<0.001
level-1, r	4.63053	21.44184			

Deviance = 9487.537707

Final Model with TP and YRS_EXP as EGT Slope Predictors

Level-1 Model

$$EOCA_{ij} = \beta_{0j} + \beta_{1j} * (EGT_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} * (TP_j) + \gamma_{12} * (YRS_EXP_j) + u_{1j}$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{10} * EGT_{ij} + \gamma_{11} * TP_j * EGT_{ij} + \gamma_{12} * YRS_EXP_j * EGT_{ij} + u_{0j} + u_{1j} * EGT_{ij} + r_{ij}$$

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	0.462755	0.663651	0.697	71	0.488
For EGTI slope, β_1					
INTRCPT2, γ_{10}	0.228641	0.225222	1.015	69	0.314
TP, γ_{11}	0.369530	0.129290	2.858	69	0.006
YRS_EXP, γ_{12}	0.011562	0.002737	4.225	69	<0.001

Final estimation of variance components					
Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	2.68124	7.18904	71	109.04307	0.003
EGTI slope, u_1	0.28861	0.08330	69	131.54580	<0.001
level-1, r	4.54989	20.70154			

Deviance = 9436.392489

Full Exploratory Model

Level-1 Model

$$EOCA_{ij} = \beta_{0j} + \beta_{1j}*(EGT_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(GKT_j) + \gamma_{02}*(TE_j) + \gamma_{03}*(TP_j) + \gamma_{04}*(YRS_EXP_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}*(GKT_j) + \gamma_{12}*(TE_j) + \gamma_{13}*(TP_j) + \gamma_{14}*(YRS_EXP_j) + u_{1j}$$

Final Estimation of Fixed Effects

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, θ_0					
INTRCPT2, γ_{00}	5.980	7.939	0.754	67	0.454
GKT, γ_{01}	0.003	0.104	0.034	67	0.973
TE, γ_{02}	-0.139	0.855	-0.162	67	0.872
TP, γ_{03}	-2.193	3.493	-0.628	67	0.532
YRS_EXP, γ_{04}	-0.084	0.091	-0.923	67	0.359
For EGT slope, θ_1					
INTRCPT2, γ_{10}	-0.318	0.663	-0.480	67	0.633
GKT, γ_{11}	-0.001	0.008	-0.098	67	0.922
TE, γ_{12}	0.031	0.071	0.433	67	0.666
TP, γ_{13}	0.536	0.271	1.979	67	0.052
YRS_EXP, γ_{14}	0.017	0.007	2.553	67	0.013

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	3.061	9.370	67	109.225	0.001
EGT slope, u_1	0.310	0.096	67	132.921	<0.001
level-1, r	4.550	20.703			

Model Incorporating the Total Number of DOK moves

Level-1 Model

$$EOCA_{ij} = \beta_{0j} + \beta_{1j}*(EGT_j) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(TE_j) + \gamma_{02}*(YRS_EXP_j) + \gamma_{03}*(TOTDOK1_j) + \gamma_{04}*(TOTDOK2_j) + \gamma_{05}*(TOTDOK3_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}*(TE_j) + \gamma_{12}*(YRS_EXP_j) + \gamma_{13}*(TOTDOK1_j) + \gamma_{14}*(TOTDOK2_j) + \gamma_{15}*(TOTDOK3_j)$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{01}*TE_j + \gamma_{02}*YRS_EXP_j + \gamma_{03}*TOTDOK1_j + \gamma_{04}*TOTDOK2_j + \gamma_{05}*TOTDOK3_j + \gamma_{10}*EGT_{ij} + \gamma_{11}*TE_j*EGT_{ij} + \gamma_{12}*YRS_EXP_j*EGT_{ij} + \gamma_{13}*TOTDOK1_j*EGT_{ij} + \gamma_{14}*TOTDOK2_j*EGT_{ij} + \gamma_{15}*TOTDOK3_j*EGT_{ij} + u_{0j} + r_{ij}$$

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-2.735918	6.667344	-0.410	66	0.683
TE, γ_{01}	0.479423	0.906008	0.529	66	0.598
YRS_EXP, γ_{02}	-0.055622	0.088251	-0.630	66	0.531
TOTDOK1, γ_{03}	0.010484	0.016621	0.631	66	0.530
TOTDOK2, γ_{04}	0.018886	0.022145	0.853	66	0.397
TOTDOK3, γ_{05}	-0.091929	0.098945	-0.929	66	0.356
For EGT slope, β_1					
INTRCPT2, γ_{10}	1.008088	0.518042	1.946	1498	0.052
TE, γ_{11}	-0.015887	0.071628	-0.222	1498	0.825
YRS_EXP, γ_{12}	0.016224	0.006363	2.550	1498	0.011
TOTDOK1, γ_{13}	-0.002582	0.001268	-2.036	1498	0.042
TOTDOK2, γ_{14}	-0.002123	0.001601	-1.326	1498	0.185
TOTDOK3, γ_{15}	0.015483	0.006939	2.231	1498	0.026

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	2.77539	7.70281	66	545.65585	<0.001
level-1, r	4.61205	21.27103			
Deviance = 9503.47963					

Model Incorporating Number of DOK 3 Level Moves

Level-1 Model

$$EOCA_{ij} = \beta_{0j} + \beta_{1j} * (EGT_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01} * (YRS_EXP_j) + \gamma_{02} * (MDOK3_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

Mixed Model

$$EOCA_{ij} = \gamma_{00} + \gamma_{01} * YRS_EXP_j + \gamma_{02} * MDOK3_j + \gamma_{10} * EGT_{ij} + u_{0j} + r_{ij}$$

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-1.969494	0.892710	-2.206	69	0.031
YRS_EXP, γ_{01}	0.167872	0.043288	3.878	69	<0.001
MDOK3, γ_{02}	0.121152	0.060125	2.015	69	0.048
For EGT slope, β_1					
INTRCPT2, γ_{10}	0.985394	0.058685	16.791	1503	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	3.00772	9.04636	69	660.39634	<0.001
level-1, r	4.62957	21.43294			

Deviance = 9473.921259

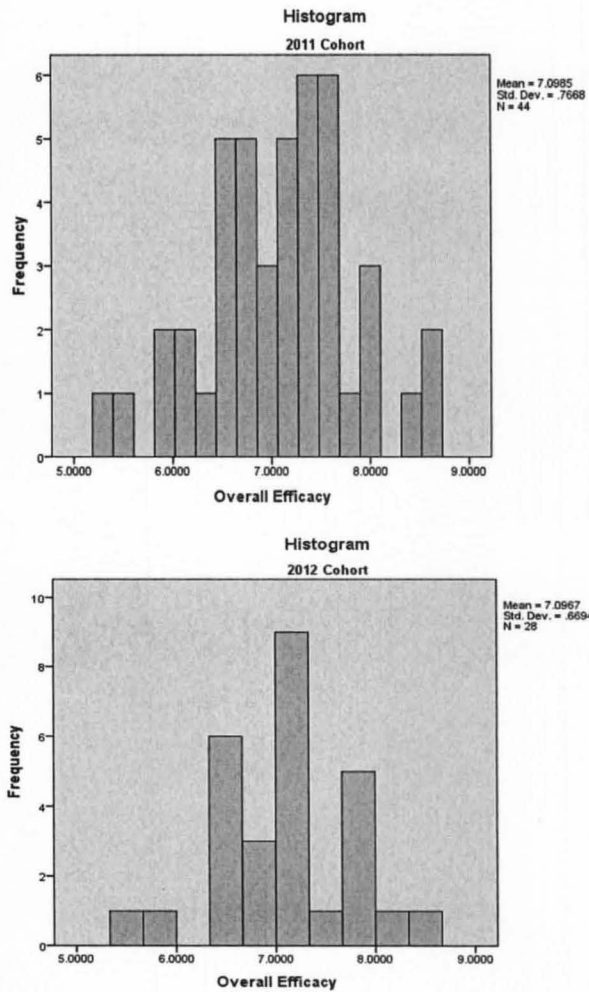
Appendix J

Distributions and Descriptive Statistics for the TSES Scores

Table J1
Descriptive Statistics for TSES Scores

Cohort	2011	2012
N	44	28
Mean	7.098	7.097
Median	7.229	7.104
Std. Dev.	0.767	0.669
Min	5.292	5.625
Max	8.667	8.500

Figure J1. Distribution of TSES Scores for 2011 and 2012 Cohorts



One-way ANOVA results showed no significant differences between the mean TE scores of both 2011 and 2012 cohorts.

ANOVA

Teacher Efficacy

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	1	.000	.000	.992
Within Groups	37.379	70	.534		
Total	37.379	71			

Number of affirmative responses to the additional item on the TSES.

1. I am no longer teaching. (1)
2. I have changed schools this year. (5)
3. My principal has changed this year. (9)
4. New administration (such as principal) has had a large effect on my teaching abilities this year. (1)
5. I am teaching the same geometry class I taught last year. (30)
6. The classes I am teaching this year are much more difficult or much easier to teach than last year. (9)
7. I feel my beliefs in my teaching abilities are about the same as they were last year. (30)

Correlations analysis was performed for the 2011 and 2012 cohorts separately with similar results to the overall group. No significant correlations were found.

CURRICULUM VITAE

PAUL J. KLEIN

1600 Ruth Ave, Apt 6

Louisville, KY 40205

cell (502) 387-9709

e-mail: pklein15@yahoo.com

EDUCATION

- Ph.D. Mathematics Education, University of Louisville, Louisville, KY 2012
- M. Ed. Xavier University, Cincinnati, OH 2002
- Diploma in Bioethics, PURA (Pontifical University Regina Apostolorum), Rome, Italy 2000
- B.A. (Theology) PURA, Rome, Italy 2000
- B.A. (Philosophy) PURA Rome, Italy 1995
- M. Eng. (EE) University of Louisville, Louisville, KY, 1990
- B.S. University of Louisville, Louisville, KY, 1988

EXPERIENCE

Evaluation Specialist, School of Medicine, University of Louisville. Jan. 2012 – Present. Responsible for management of student and faculty databases and associated databases used to evaluate medical school students, courses, clerkships, and faculty. Conducted statistical analysis and prepared reports for medical school Deans and Department Chairs involving course evaluation, student grade levels, interdepartmental instruction, assessment results, and yearly annual medical education reports.

Evaluation Specialist, Department of Accountability, Research & Planning. Jefferson County Public Schools, 2011 – 2012. Collected quantitative and qualitative data and conducted mixed methods analysis for the evaluation of district-wide programs and projects. Prepared statistical reports using demographic, school, and district level data for executive review and decision making. Developed surveys and program evaluation reports for district, state and federal educational intervention programs after careful data preparation and analysis.

Graduate Research Assistant. University of Louisville, 2007-2011. Contributing investigator for the NSF funded Geometry Assessment for Secondary Teachers project (GAST), a research effort seeking to link a teacher's depth of knowledge of geometry with student achievement outcomes. Participated in all aspects of the project including textbook review, assessment blueprint development, item writing, observation instrument development, classroom observation, database management, advanced statistical analysis and publication writing.

Engineering Math Assistant Professor, University of Louisville, 2007 Instructed engineering students of all levels from freshmen to upper classmen. Courses ranged from Precalculus, to Linear Algebra and Numerical Analysis. Developed visual instructional techniques to explain multi-dimensional concepts using Maple mathematical software.

Sales Engineer. Automatic Controls Co. Louisville, KY 2003- 2007 Solicited sales calls and provided technical support and engineering expertise on a variety of process control equipment and utility metering devices for Kentucky manufacturing facilities. Cultivated and maintained major sales accounts including major manufacturing plants along with refineries, distilleries, hospitals, and municipal water and wastewater treatment facilities.

Education Consultant. Kindred Healthcare, Louisville, KY 2002-2003 Designed nationwide web-based process and procedure instructional courses for internal pharmaceutical employees using Intellinex e-Learning Development Systems. Created course structure and evaluation reviews to track employee progress. Developed educational storyboards, incorporating procedural instructions, computer simulations, and graphical interfaces.

Statistics Professor, Xavier University, Cincinnati, OH 2001-2002 Instructor of undergraduate statistics course with class sizes averaging 28 students from freshmen to senior levels. Topics included descriptive statistics, probability distributions, confidence intervals, hypothesis testing, regression, correlation, and Chi-square tests analysis of variance.

Missionaries of the Sacred Heart Seminary Cheshire, CT. 1990-2001

Assistant to the Director/Student Counselor. (Our Lady of Thornwood College, Thornwood, NY 1997-1998) Counseled philosophy and theology students, assessing student growth and development. Placed student body in charitable work positions and tracked progress. Supervised teams responsible for maintenance and operation of dining halls and dormitories, guest room preparations, and audio/video projects.

High School Instructor/Assistant Prefect of Studies. (Immaculate Conception High School, Center Harbor, NH, 1995-1997) Supervised faculty and instructed courses in Latin, Greek, Spanish, and History. Editor/designer of school publications utilizing Adobe Photoshop and PageMaker desktop publishing software.

SUPPORTING DATA

- Fluent in Spanish and Italian
- Mastery of SPSS Statistical software and Microsoft Excel
- Personal interests: running, cycling, sailing, cooking