Distribution Agreement

Mathematical Knowledge for Teaching, Teacher Practices, and Student Learning within Urban Learning Contexts

by

Rubye K. Sullivan
Doctor of Philosophy

Educational Studies
Emory University

____
Robert J. Jensen
Chair

____
Yuk Fai Cheong
Committee Member

____
George Engelhard, Jr.
Committee Member

Accepted:

____
Lisa A. Todesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

____
Date

Mathematical Knowledge for Teaching, Teacher Practices, and Student Learning within Urban Learning Contexts

By

Rubye K. Sullivan
B.S., Boston University, 1994
M.Ed., Emory University, 1997

Advisor: Professor Robert J. Jensen, Ed.D.

An abstract of
A Dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies Division of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In Educational Studies
2010

Abstract

Mathematical Knowledge for Teaching, Teacher Practices, and Student Learning within Urban Learning Contexts

By Rubye K. Sullivan

This study explored how the staffing of the classroom with a teacher in possession of mathematical knowledge for teaching (MKT) or a teacher implementing standards-based, reform-oriented (SB – RO) teaching practices might relate to classroom and school-level aggregates of student race, class status, and prior learning in mathematics. In addition, I examined whether mathematical knowledge for teaching and the use of standards-based, reform-oriented teaching practices relate to student learning in mathematics. A correlational analysis along with a multi-level regression analysis, specifically hierarchical linear modeling, were employed. The teacher level variables are teacher knowledge (MKT) and teaching practices in mathematics (SB-RO). Student level data consisted of two administrations of the state's criterion-reference test (CRT). Students' grade two and grade three mathematics scale scores were included along with the student contextual variables, race and class status. The sample consisted of 531 grade three students nested in the classrooms of 35 grade three teachers in 17 elementary schools, approximately two teachers per school, from a large urban district in the southeastern United States. Results indicated that although mathematical knowledge for teaching and the use of standards-based, reform oriented practices are positively correlated (r (33) =.30, $p < .01$), neither exhibited statistically significant contributions to the prediction of student learning in mathematics. Additionally, the possession of MKT and the use of SB-RO practices were not related to the proportion of African American students, the proportion of students eligible for free or reduced lunch, or the prior learning in mathematics at the class or school level. The teachers studied, however, possessed lower levels of MKT than the average elementary teacher in the nation.

Mathematical Knowledge for Teaching, Teacher Practices, and Student Learning within Urban Learning Contexts

By

Rubye K. Sullivan
B.S., Boston University, 1994
M.Ed., Emory University, 1997

Advisor: Professor Robert J. Jensen, Ed.D.

A Dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies Division of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
In Educational Studies
2010

Acknowledgements

Writing a dissertation can be a lonely and isolating experience at times. This was only true for me in the rarest of occasions. My husband Yusef and my children, Elijah and Zola, were partners in my work, keeping the loneliness and isolation to a minimum. My uncle Dow provided my writing sanctuary, the lake house, for which I will always be grateful.

My committee chair, Dr. Robert Jensen, was both supportive and encouraging. He let me make my own mistakes and I am a better scholar and professional for experiencing the cognitive struggles resulting from my gung-ho approach to this work.

Dr. George Engelhard, Jr. will always be a collaborative partner in my research trajectory. The work that we have done and will continue to do using the Rasch measurement model both excites and exhausts me because the options are many.

Dr. Yuk Fai Cheong's work facilitating my understanding of hierarchical linear models (HLM) was invaluable. This research study would not have been actualized without the ability to model the nested nature of students within classrooms. I remain excited about further applications of HLM in my future work understanding teacher effectiveness.

Without my friends and neighbors, those on Elmira and off, who listened to my struggles and my triumphs, this work would not have been possible. Members of my Emory student cohort, the Uncontrolled Variables, were recipients of the tears and the laughter that accompanied the struggle of a doctoral program. We will always be a family. Thank you all.

Table of Contents

List of Tables

List of Figures

Chapter One: The Problem

The Education Trust, a nonprofit organization working to close achievement gaps that separate students of color and low-income students, documents the inequitable distribution of effective teachers. Students of color and low-income students tend to learn in classrooms staffed by undereducated teachers (Haycock,1998). Undereducated teachers, in turn, decrease the opportunities to learn and may therefore contribute to the achievement gap. William James (2001, 1899) stated, “the teachers of this country, have its future in its hands” (p.1). Our nation is struggling to provide equitable educational opportunities for the students that will one day lead this nation. In mathematics and science, the problem is exacerbated by the need for our nation to compete in a global economy fueled by technology innovation. The issue, teacher effectiveness, seems simple. Given that teachers can affect the future, how might we develop teachers with the qualities necessary to ensure student learning? How can we provide the professional development for existing teachers, specifically teachers in urban districts, to potentially close the academic achievement gap that persists for students of color and low-income students?

Darling-Hammond & Sykes (2003) state, “qualified teachers are a critical national resource that requires federal investment and cross-state coordination” (p. 33). This decade, we have witnessed a federal focus on teacher quality in the reauthorization of the Elementary and Secondary Education Act (ESEA) as the No Child Left Behind (NCLB) Act (2001). NCLB required states to develop teacher quality criteria. More recently, President Obama and his Secretary of Education, Arne Duncan, have required states to remove legislation impeding the use of student learning outcomes in the evaluation of teachers to even be considered for the *Race to the Top* grants being awarded as a part of the Stimulus Act (2009). In Washington, D.C., Michelle Rhee, the Chancellor of D.C. Public Schools has recently proposed an evaluation system requiring teachers to demonstrate their influence on student learning gains during one academic year. In return, teachers will receive much higher rates of pay, but will have to give up their tenure. This controversial proposal in the nation's worst performing district illustrates the necessity of understanding a teachers' influence on student learning. How does a state, a university, or a school district determine teacher quality characteristics? How does each entity determine the curriculum and content of programs preparing new teachers, or inducting new teachers into the profession? If a teacher is not producing adequate student learning, what professional learning opportunities should be offered to remedy the situation?

Researchers have been examining the relationship between educational inputs (e.g. teacher knowledge or teaching practices) and educational outputs (e.g. learning) since Coleman (1966) released his controversial report, *Equality of Educational Opportunity* (EEO). The Coleman Report, as it became known, was the first in a line of sociological studies within the category *Educational Production Function* literature. The focus of these studies was on the relationship between resources and student achievement wherein teachers were included as one resource. As obvious and commonsensical as it may seem that an output such as student achievement is closely linked to an input such as teacher quality, capturing the quality of a teacher has been somewhat elusive. Coleman and his colleagues examined teacher degree levels, experience, and verbal scores as measures of teacher quality. The results indicated that neither teacher degree level nor experience seemed to relate to the student learning of most students, while verbal ability had a relatively small effect on student learning.

Another similar line of inquiry emerged in educational research at approximately the same time known as teacher *process-product* literature. Studies within this body of work focused on the relationship between teacher practices during their interaction with students (process) and student achievement (product). Shulman (1986) criticized the process-product literature for not examining teacher effects in the context of specific subject matter, such as mathematics. Shulman's critique led later researchers to traverse the same path of trying to connect teacher quality to student learning, specifically in the area of mathematics (Monk & King, 1994, Goldhaber & Brewer, 1997, Rowan, Chiang & Miller, 1997, Goldhaber & Brewer, 2000, Rowan, Correnti, & Miller, 2002, Hill, Rowan, & Ball, 2005, Kane, Rockoff, & Steiger, 2006).

Researchers considered early teacher production-function literature to be overly focused on teacher characteristics or "traits" (Brophy & Good, 1986). The initial focus on teacher traits, including intelligence, were poorly measured and were not directly related to the professional knowledge necessary for teaching (Rowan, 1999). An emerging interest in teaching as "expert" work has resulted in measures designed specifically for capturing the professional knowledge necessary for teaching. In the area of mathematics, Deborah Ball and her colleagues from the *Study of Instructional Improvement* worked with National Science Foundation funding to design measures to capture mathematical knowledge for teaching (Hill, Schilling, & Ball, 2005). Grounded in Shulman's (1986) seminal work defining pedagogical content knowledge, a groundbreaking measure was developed specifically for measuring the situated knowledge of teachers of mathematics. Figure 1 (Engelhard & Sullivan, 2007) illustrates the intersection of Shulman's dimensions of teacher knowledge, Deborah Ball's dimensions of teacher knowledge specific to mathematics, and the dimensions that can be measured using the *Study of Instructional Improvement* instruments. Knowing how knowledgeable a teacher may be only addresses one piece of the production-function puzzle. A new question emerges: How does the professional knowledge of a mathematics teacher translate into practice?

A teacher may be defined as effective or highly qualified based on traits, such as degree level or experience; however, if they are unable to effectively navigate the interactive phase of teaching, how could the teacher affect student learning? Teaching practices, or the intentional methods employed by the teacher during the interaction between teacher and student, seem to link directly to teacher quality. Teacher actions became the primary focus of process-product literature in response to the overemphasis on teacher traits in the production-function literature (Rowan, 1999). Thomas J. Cooney (1980) in a meta-analysis of research related to teaching and teacher education states, "the role of the teaching agent is to engage in behavior that gives rise to setting learner interactions that cause the student to learn. This behavior is called teaching" (p. 433). During the interactive phase of teaching in mathematics, intentional methods are typically based on two distinct philosophies, the more traditional rote-learning teaching practices and the practices encouraged by the National Council of Teachers of Mathematics (NCTM). This study deals specifically with one type of instructional practice. Endorsed by the NCTM in the Curriculum and Evaluation Standards for School Mathematics (1989, 2000), standards-based, reform-oriented teaching practices encourage conceptual meaning and understanding (NCTM 1989, 2000; Hierbert, Carpenter, Fennema, Fuson, Human, Murray, Olivier, & Wearne, 1996).

This study builds on the research of Hill, Rowan & Ball (2005) where the effect of a specific type of teacher knowledge, the mathematical knowledge for teaching, was statistically linked to student learning in mathematics. The authors found that mathematical knowledge for teaching was the strongest teacher-level predictor of student achievement gains, more than average instructional time spent on mathematics and other teacher input variables such as coursework and years of experience. Hill, Rowan, & Ball (2005) indicate "our findings help envision a new generation of process-product studies designed to answer questions about how teachers' mathematical behavior – in particular, the classroom explanations, representations, and interactions with students' mathematical thinking, might affect student outcomes" (p. 400). Additionally, findings from my own empirical research (Sullivan, 2009) guide the choices in teacher-level characteristics that are included in the analysis. The rationale for this study is to contribute to the process-product literature in mathematics education while intentionally focusing on how aggregates of student race, class, and prior knowledge may be related to the likelihood of being taught by a teacher possessing mathematical knowledge for teaching or using standards-based, reform-oriented practices. Brophy (1991) states, "where knowledge is more explicit, better connected, and more integrated, [the teacher] will tend to teach the subject more dynamically, represent it in more varied ways, and encourage and respond fully to student comments and questions" (p.352). Exploring the effects of teacher knowledge and teaching practices on student learning answers the call from Hill and her colleagues for "examining the interactions between teacher knowledge and instructional methods" (Hill, Rowan, & Ball, 2005, p. 401). Specifically, examining the ways in which teachers explain, represent, and interact with the students and the mathematics along with an analysis of mathematical knowledge for teaching and student learning, fills a void in the process-product literature.

One complication that is predominant in the process-product studies in education is the non-random assignment of students to both schools and classrooms. Society and economics play a vital role in determining placement, creating placements that are far from random. The ability of researchers to separate the effects of schools, classrooms, and teachers from student backgrounds in these nonrandomized situations is difficult (Raudenbush 2004). Decades of research attempting to examine the effect of schools and school personnel on student achievement have been conducted, often with findings that contradict one another, more often than not due to methodological choices related to how product variables, specifically student learning, are measured (Rowan, 2002). Because of the nested structure of these data, the analyses are conducted using Bryk & Raudenbush's (1992) approach to hierarchical linear modeling. In this study, students are nested within teachers. The variance in student learning is therefore estimated at two levels, the variance that is 1) between teachers within schools, and 2) between students within teacher. In this study, I am interested in the differences in student learning affected by teacher-level variables. As such, I employ a covariate adjustment model. Other studies have used gain scores; however, a recent report *Using statistics effectively in mathematics education research* (Scheaffer, et al., 2007) recommends the use of covariate adjustment models due to concerns around the reliability of gain scores. Specifically, students scoring at the top end of the scale on the pre-test do not have the same opportunity for a resulting gain score as students scoring at the bottom end of the scale on the post-test. Also, the assessments used in my study are not vertically equated such that a difference would be meaningful. Rogosa (1995) recommends growth curves for estimating student learning, but this approach requires at least three waves of student data. Two waves, a spring 2008, grade two administration and a spring 2009, grade three administration of a state criterion-referenced test are used in the study.

The practical implications of this study may influence current teacher practice, teacher preparation, teacher retention, and educational policy in related to providing all students with highly qualified teachers. Hill, Rowan, & Ball (2005) found that teacher knowledge is inequitably distributed across student race and class status. In particular, as the proportion of students of color and students receiving free or reduced lunch increase, the respective teacher's mathematical knowledge for teaching decreases. An inequitable distribution of teacher knowledge could result in an inequitable distribution of effective teaching practices, in turn resulting in lower student learning outcomes for the disadvantaged subset of students. This study may provide insight into contributing factors for the achievement gap so often discussed in mathematics based on race and class - an insight that should inform hiring and retention practices. The relationship between teacher knowledge and teaching practices and the ways in which each variable affects student learning should have far-reaching influence on the curriculum of teacher preparation programs and professional learning opportunities. Policy related to providing all students with highly qualified teachers will be provided a solid, empirical base on which to establish the metrics of teacher quality. Process-product research will also be influenced in that this study directly addresses the new direction of the literature called for in Hill, Rowan and Ball's (2005) study.

Conceptual Framework

I utilize Dunkin and Biddle's (1974) scheme for classifying variables as *presage, process, context,* and *product* to inform my conceptual framework. Presage variables refer to the characteristics of teachers that are in place prior to the teacher's interaction with students, but may still have an influence on the interactive phase of teaching. Process variables are specific to the interactive phase of teaching, or the moments in which students and teachers interact with curriculum and content. Context variables are defined as variables that can influence the educational outcome or condition the potential relationships between process variables and student learning outcomes. Product variables consist of the potential outcomes of teaching, or student learning in this study. A schema from 1974 may appear dated; however, this schema is still used as a conceptual framework in the literature examining teacher effects on student learning outcomes (Hill, Rowan, & Ball, 2005; Rowan 2002) and fits well within a theme of extending the process-product literature.

The purpose of this study is to explore the relationship among teacher knowledge, teaching practices and student learning in mathematics. Additionally, an explicit focus on how the concentration of African-American students, students in poverty, or lower performing students might be associated with being taught by teachers with high levels of mathematical knowledge for teaching or teachers who use standards-based, reform-oriented practices is examined. The context variables student race/ethnicity and class status are considered. Using Dunkin and Biddle's (1974) framework for understanding teaching as a conceptual framework, I examined how the teacher-level presage variable mathematical knowledge for teaching (*MKT)*, the teacher process variable standards-based, reform-oriented teaching practices (*SB-RO*), and the aggregates of student-level contextual variables *race/ethnicity* and *class status* might relate to student-level product variable *student learning in mathematics*. The following two research questions guide this inquiry:

1. Are there any associations between the ethnic, SES, and ability compositions of classrooms or schools and teachers' level of mathematical knowledge for teaching and use of standards-based, reform-oriented teaching practices?
2. What is the relationship among mathematical knowledge for teaching, the use of standards-based, reform-oriented teaching practices, and student learning outcomes in mathematics?

Research question one extends the empirical study (Sullivan, 2009) to include the student context variables Race (*African-American)*, class status *(eligibility for free or reduced lunch)*, and *prior mathematics achievement* aggregates at the classroom and the school levels. Given the belief that students of color and in poverty are less likely to be taught by high quality teachers, I would like to know if African American students, students who are eligible for the free or reduced lunch program, or less mathematically proficient students are less likely to be taught mathematics by teachers who possess mathematical knowledge for teaching or who implement standards-based, reform-oriented teaching practices. Research question one will be approached using a correlational analysis. I expect that as the proportion of African American students, or students eligible for free or reduced lunch increases, or as the mean prior learning score in mathematics decreases, that teacher's level of mathematical knowledge for teaching and his or her use of standards-based, reform-oriented practices will decrease. In research question two, given the previous work done by Hill, Rowan, & Ball (2005) in relating a teacher's mathematical knowledge for teaching to student learning outcomes in mathematics, I expect mathematical knowledge for teaching to have a significant partial effect on student learning in mathematics. Based on my own work relating teacher knowledge and teaching practices (Sullivan, 2009), I expect teaching practices to have a significant partial effect on student learning in mathematics. Given the increased understanding of the quality of the psychometric properties of the student measures to be used in this study, I expect the ability of my student learning measure to better capture the latent variable student knowledge in mathematics. In turn, using two time points of a more reliable and valid instrument will provide a more precise view of student learning in mathematics than in my previous work using raw scores on a district-level assessment (Sullivan, 2009). Figure 2 provides an illustration of research question two, including each variable type and their respective levels of analysis.

Chapter Two: Literature Review

Teacher quality in the area of mathematics has been defined and measured using multiple methods since the advent of the education production function research. The literature review is organized thematically using Dunkin & Biddles'(1974) schema, including *presage*, *process*, *context*, and *product* variable types. Empirical evidence related to each of the variable types is included. I also discuss my methodological choices, hierarchical linear modeling and the use of a covariate adjustment model.

I used selection criteria to determine the studies that I would include in the review of literature. Table 1 provides the criteria levels that I used to determine the studies that would be included in my literature review. The criteria are organized by the focus, teacher or student, of the study, the variable type as it relates to my conceptual framework, and the variable name. First, I only included studies that are specific to mathematics where feasible. Second, I included literature that deals with at least two-levels of schooling, teachers and students. Finally, I included only articles that examine the relationship between either *presage* and *product* variables, *process* and *product* variables, or the influence of *context* variables on the relationship between *process* and *product* or *presage* and *product*. Each piece of relevant literature is listed in the table.

Presage: Teacher Knowledge Proxy Measures

Elbaz (1983) states, "the single factor which seems to have the greatest power to carry forward our understanding of the teachers role is the phenomenon of teacher knowledge" (p. 45). It seems commonsensical that a teacher must have knowledge of the content that they teach. However, linking a specific type of knowledge to student learning outcomes is rare (Fennema & Loef Frank, 1992). Not examining the types of knowledge that affect student learning limits our understanding of teacher quality. Proxy measures, or substitutes, of teacher knowledge dominate the literature. Primary proxy measures include years of experience teaching mathematics, the number of mathematics courses taken in college, and being degreed in mathematics (Monk & King, 1994; Goldhaber & Brewer, 1997; Rowan, Correnti, & Miller, 2002; Hill, Rowan & Ball, 2005).

Across the literature linking teacher proxy measures for knowledge with student learning outcomes, regardless of the measure, findings are mixed. The National Mathematics Advisory Panel's (2008) task group on teachers and teacher education conducted an extensive meta-analysis of research related to teacher knowledge. The group supported my own interpretation of the literature on proxy measures of teacher knowledge in mathematics, " across all studies, the findings are mixed" (p. 5-x). Additionally, since much of the research has been conducted based on proxy measures, a significant finding still does not provide much detail as to the specific knowledge and skill necessary to affect student learning.

I examined five studies specifically meeting the criteria for inclusion discussed previously. The specific proxy measure I examined in this set of literature is coursework completed in mathematics. Measures of coursework are operationalized in two ways: some studies examine the number of mathematics courses taken and others focus on the teachers' major. Three of the five studies measured teacher knowledge using the possession of a degree in mathematics, in other words, the authors examined teachers who majored in mathematics and the effect of the math major on student learning (Rowan, Chiang & Miller, 1997; Goldhaber & Brewer, 2000; Rowan, Correnti & Miller, 2002). Positive effects on student learning were found in the Rowan, Chiang, and Miller (1997) and Goldhaber and Brewer (2000) studies. A negative correlation was found between the possession of a math major and student learning in the Rowan, Correnti, and Miller (2002) study. Hill, Rowan, and Ball (2005) examined the relationship between the number of courses taken and student learning and found that no significant relationship was present. Monk and King (1994) measured both the number of courses taken and the possession of a math major and found a positive relationship between each measure and student learning. Due to the mixed results of the literature related to the proxy measure coursework attainment, I included the number of courses in mathematics taken at the university level in my own empirical study (Sullivan, 2009). My findings revealed that the number of courses did not have a statistically significant effect on student learning outcomes. Based on this finding, I am not including this proxy measure in this proposed study.

Certification levels are another dominant proxy measure for teacher knowledge in mathematics. I include four studies that examine potential effects of mathematics certification on student learning. Of the four, two found a positive effect (Goldhaber & Brewer, 1997; King Rice, 2003), one found a negative effect (Rowan, Correnti, & Miller, 2002), and one found no significant relationship between certification level and student learning (Hill, Rowan & Ball, 2005). Certification in mathematics is rare in elementary school. The elementary certification level in the state in which this study is conducted is that of generalist. Teachers are also given the opportunity to participate in a state- and district-supported mathematics endorsement course. Again, to add further data to this unclear issue I included a more general distinction in certification, traditional versus alternative certification, in my empirical study (Sullivan, 2009) and found that certification type did not have a statistically significant effect on student learning outcomes in mathematics. Additionally, the endorsement course was not found to have enough variability in my empirical sample to be included in the model.

Given the focus on proxy measures in previous research and my own exploration, I will not include proxy measures in the proposed dissertation study. A more narrowed focus on a direct measure of teacher knowledge will fulfill the *presage* variable type from the conceptual framework.

Presage: Direct Assessment of Teacher Knowledge

Direct assessment of teacher knowledge comes in many forms in the studies included in this review of the literature. Some researchers use pre-established forms while others create their own. A direct measure is often more attractive to both researchers and administrators, but great care must be taken to ensure both the validity and the reliability of the measure. Again, this approach to capturing the knowledge of teachers is plagued by mixed results; however, a promising measure emerges from the body of literature.

The direct measures used to capture teacher knowledge in mathematics included a one-item measure (Rowan, Chiang & Miller, 1997), the quantitative portion of the Scholastic Aptitude Test (Harris & Sass, 2007), the math portion of the teacher certification test (Mullens, Murnane & Willett, 1996), and a measure built specifically to capture the situated knowledge of mathematics teachers (Hill, Rowan & Ball, 2005). Issues of reliability and validity abound in the use of the instruments described. The results are mixed. Two studies did not find a significant relationship (Harbison &

Hanuchek, 1992; Harris & Sass, 2007), two found a positive relationship (Rowan, Chiang & Miller, 1997; Hill, Rowan, & Ball, 2005), and one study (Mullens, Murnane & Willett, 1996) found a positive relationship between teacher knowledge and students' understanding of advanced topics and no significant relationship between teachers' knowledge and students' understanding of basic mathematics. Given that two studies found a statistically significant effect, I examined the instruments used to measure teacher knowledge in both studies. The Rowan, Chiang, and Miller (1997) instrument contains one item, so I elected not to use that instrument due to reliability concerns. This review led me to further examine the work of Hill, Rowan, and Ball (2005) and their study of mathematical knowledge for teaching. In my own study (Sullivan, 2009) mathematical knowledge for teaching was not found to have a statistically significant effect on student learning in mathematics; however, it was found to be positively correlated (.31, $p \leq .05$) with the use of standards-based, reform-oriented teaching practices, which was found to be a statistically significant predictor of student learning. Given a better understanding of the student measures and the correlation found in the previous study, I am including mathematical knowledge for teaching in the proposed dissertation study.

Presage: The Evolution of Mathematical Knowledge for Teaching

In the area of mathematics, a great deal of progress has been made in both defining subject matter knowledge and in the development of a way to measure this construct. Shulman (1986) coined the phrase pedagogical content knowledge as a specialized body of knowledge that includes subject matter competence as well as the ways in which students may be interested in or engage with the content. Shulman's work was responsible for beginning the conversation on the existence of a specialized body of professional knowledge that teachers must possess in order to be effective that goes beyond the typical knowledge of an educated adult. Liping Ma (1999) described this body of knowledge as a "knowledge package" (p. 17). Her comparative study used interviews to distinguish the presence of this "knowledge package" in teachers in Shanghai, China and in the United States to determine a possible link to student achievement in the two countries. Deborah Ball (1988) began unpacking this knowledge according to specific content strands such as fractions and place value in an effort to distinguish the exact nature of this pedagogical content knowledge. She then joined with a noted mathematician, Hyman Bass, and examined the knowledge in the context of pure mathematics for accuracy and alignment of teacher knowledge with the field of mathematics (Ball & Bass, 2000). Ball's work eventually led to the content-specific extension of Shulman's work and described the body of knowledge as *mathematical knowledge for teaching* (MKT) (2007).

Ball and her colleagues have taken their research a step further by conducting a study that measures the effect of mathematical knowledge for teaching on student gains in mathematics (Hill, Rowan, & Ball, 2005). The authors used a multi-level model with gains as the outcome to examine how teachers' mathematical knowledge for teaching contributes to student achievement gains in first and third grades. The study conducted by Hill, Rowan, and Ball (2005) illustrates that a direct measure of this specialized body of knowledge supercedes the effects of proxy measures such as the number of courses taken or majors attained in the content area in its' ability to predict student learning. While multiple variables related to the student, teacher, classroom, and school were collected and included in this important study by Hill and her colleagues, no information was collected concerning the specific instructional practices of the teacher. The closest measure of instructional practice was the average classroom time spent daily in mathematics. The authors indicate that an important area for future research would be an examination of how teachers use their knowledge when interacting with students, and how that use may affect student learning. The knowledge base that has been established by Ball and her colleagues has implications for professional learning and pre-service teacher preparation programs as well as equity, in that a measure now exists that could assist educators in ensuring that "the intellectual resources are available to students across race and SES" (Hill, Rowan, & Ball, 2005, p. 400).

Hill, Rowan, and Ball (2005) found that a teacher's mathematical knowledge for teaching significantly predicted student gains in grades one and three modeling the data in a three-level hierarchical linear model with student gain scores during a one year period as the dependent variable. Mathematical knowledge for teaching was the strongest teacher-level predictor of student achievement gains, more than average instructional time spent on mathematics and other teacher background variables such as coursework and years of experience. In fact, the effect size in grade three "rivaled that of SES and student ethnicity" (p. 396). Coefficients in the model can be interpreted as the effect of a one-standard deviation increase in the independent variable on gains made by students during one academic year. The mathematical knowledge for teaching coefficient was 2.28, $p < .01$ while SES was 2.13, $p < .01$ (Hill, Rowan, & Ball, 2005). The authors cite lack of alignment between the content measured in the teacher knowledge measure, the Learning Mathematics for Teaching instrument, and the content measured in the student achievement measure (CTB/McGraw-Hill's Terra Nova Survey) as one limitation. Multiple areas for future research are discussed, including the need for examining the effects of mathematical instructional methods on student performance and "investigating whether and how the instructional practices of mathematically knowledgeable and less mathematically knowledgeable differ" (p. 401). Finally, the authors state that the ways in which mathematical knowledge for teaching affects instruction has yet to be studied and analyzed. It is the intent of this study to examine the ways in which mathematical knowledge for teaching relates to instruction by including the process variable teaching practices.

Process: Standards-Based Instructional Practices

Dunkin and Biddle (1974) consider *process* variables as the heartland of research on teacher effectiveness. In 1989, the National Council of Teachers of Mathematics (NCTM) emerged as the first to release national standards within a content area by publishing *Curriculum and Evaluation Standards for School Mathematics* (1989). The vision for the standards, now in their second iteration, is to provide access to engaging and high quality mathematics for all students (1989, 2001).

In 2002, Ross, McDougall, and Hogabaum-Gray conducted a meta-analysis of research related to the reform of mathematics. They examined NCTM documents and 154 empirical studies conducted from 1993 to 2000 on standards-based instruction in mathematics. From this review emerged nine dimensions of standards-based instruction in mathematics: (a) program scope – characterized by a broad view of mathematics with all students having access to all types of mathematics; (b) student tasks – characterized by complex, open-ended problems embedded in a real-life context; (c) discovery –

characterized by a focus on the discovery of mathematical ideas led by student discovery; (d) teacher's role – characterized as that of a co-learner and creator of a mathematical community within the classroom; (e) manipulatives and tools – characterized by ready access to and use of manipulatives and other mathematical tools such as calculators; (f) student-student interaction – characterized by the promotion of student-to-student interaction; (g) student assessment – characterized as authentic and embedded in everyday instruction; (h) teacher's perception of math as a discipline – characterized by an understanding of mathematics as a dynamic subject as opposed to a fixed body of knowledge; and (i) student confidence – characterized by a teacher's intent to boost mathematical confidence in students. Table 2 details the nine dimensions captured in this study.

Given the call for linking teacher knowledge and teaching practices in mathematics, I have elected to use the nine dimensions to capture and describe teaching practices. The nine dimensions, and the measure developed to capture the implementation of the dimensions, will allow for quantitative analysis and statistical linking with both teacher knowledge and student learning. The instrument will not only allow for a general understanding of a teachers' implementation of standards-based, reform-oriented practices, but will allow for future exploration into how the knowledge intersects with each of the nine dimensions.

Since the inception of the standards in mathematics, multiple curricular development projects have been funded, resulting in standards-based curricula that have been implemented in classrooms across our nation. A main component of each of these curriculum programs is that mathematics should be taught in such a way that both meaning and understanding are emphasized. This study is based on the reform-oriented, standards-based approach to mathematics instruction. Grounded in the belief that "instructional programs that emphasize conceptual development, with the goal of understanding, can facilitate significant mathematics learning without sacrificing skill proficiency" (Hiebert, 2003, p. 16), the notion of effective teaching practices is operationalized as the practices espoused by the NCTM standards.

In my own empirical study (Sullivan, 2009) I found the reported use of the nine dimensions of standards-based, reform-oriented practices to be both a statistically and practically significant predictor of student learning of mathematics. For every one standard deviation increase in the use of standards-based, reform-oriented practices, student scores increased by one raw score point. A limitation of the empirical study was the lack of information about the student level measures. In this study, I propose the use of state developed and mandated assessments hoping to better estimate student learning and in turn, better estimate teacher effects on the student learning.

Context: Student Race/Ethnicity and Class Status

Recently, scholars have debated the benefit of the dissection of achievement gaps in mathematics. Rochelle Gutierrez (2008) equates the prolonged "love affair" with various achievement gaps to the act of navel gazing, coining the term "gap-gazing." Lubienski (2008) respectfully disagrees and argues for a continued line of inquiry into better understanding the achievement gaps in mathematics. Both scholars agree on a new direction for the lines of inquiry. Gutierrez asks that mathematics educators study how the gaps can be positively affected with "more research on effective teaching and learning environments for Black, Latino/Latina, First Nations, English language learners, and working class students" (p. 362) in addition to a better understanding of intervention work including professional development. Lubienski recommends a move toward more complex analysis, such as the use of hierarchical linear modeling "to attend to such sample size conundrums as students nested within classrooms" (p. 354). Additionally, Lubienski recommends that studies of instructional practices should focus on typically underserved populations rather than students in general. This study responds to both authors call for research to better understand how the race and class-based achievement gaps might be affected by teacher quality in mathematics using complex modeling techniques to better estimate the effects of teacher quality on typically underserved students.

Research on the differences in the mathematics achievement between students of color and their White counterparts has been conducted since Coleman (1966). William Tate (1997) argues that there is a lack of research in mathematics education that addresses policy implications specifically for urban and rural school districts, specifically as it relates to the equitable distribution of high quality teachers across student race and poverty status. In general, Tate argues, that much of the research has focused on policy related to the development of new materials or procedures as opposed to policy that is specific to the schooling of students living in urban and rural communities. Tate calls on literature beyond the realm of mathematics education where "political and cultural dimensions of both low-SES and students of color are considered (e.g. Banks & Banks, 1995; Irvine, 1990; Ladson-Billings, 1990, 1995; and Ladson–Billings, & Tate, 1995)" (p. 674). I argue that understanding the exact nature of teacher quality in mathematics has the potential to address policy specific to the typically underserved populations consisting of students of color and students classified as low-SES.

Much of the research examining the mathematics achievement trends that are specific to students of color include comparative analysis, or gap analyses. Scholars have most notably been focused on trends in the race-based achievement gap. Secada (1992) found that while the Black - White gap had been narrowing, it seemed to only be narrowing on low-level items that required basic skills mastery. Similarly, Green, Dugoni, Ingels, and Camburn (1995) defined five levels of math proficiency using the NELS:88 data that ranged from an inability to perform simple arithmetic to the ability to solve complex, multi-step word problems. They found that 12% of the African American and 20% of the Hispanic students scored at or above advanced proficiency (Level 4) while 39% of the Whites and 45% of the Asians scored at or above advanced proficiency. Additionally, 50% of the African American and 42% of Hispanics scored low proficiency or below while 14% of their Asian counterparts and 21% of their White counterparts scored low proficiency or below proficiency. Rainski, Ingels, Rock, and Pollack (1993) used an item response theory (IRT) scaling procedure to compare the change in scores by ethnic group from the *High School & Beyond Data* in 1980 and the 1990 NELS:88 data. They defined their effect size as the difference between the mean score in 1980 and the mean score in 1990 divided by the pooled 1980/1990 standard deviations. Black students had the largest effect size (.35) followed by Hispanic students (.34), White students (.21) and Asians (.12). This study illustrated a narrowing of the race-based achievement gap in mathematics during the 1980s for Black and Hispanic students, but not specific to any one level of mathematical skill like the previously mentioned studies.

Tate (1997) states “poverty is more severely concentrated among African American and Hispanic students than it is among Whites” (p. 667). In a meta-analysis of studies examining the effects race and class have on student achievement in mathematics, Tate finds that “across the various assessments, a strong relationship between SES and mathematics achievement was evident. These test results demonstrate the need to raise mathematics achievement of low-SES students as a whole and, even more urgently, of low-SES minority students” (p. 667). Green, Dugoni, Ingels, and Camburn (1995) followed up their NELS:88 study using SES categories. High SES students were less likely than middle or low SES students to be classified as below proficiency (8.6%, 25.1%, and 46.1% respectively). Additionally, the authors examined the percent of students performing at each level by racial/ethnic group while controlling for SES. The authors found that achievement differences persisted even when controlling for SES. Secada (1992) found that middle SES and upper SES students enter school with higher achievement levels than students from lower SES.

The achievement gaps have been argued as having lifetime consequences for the underserved populations, limiting opportunities to learn, employment, and future earnings (Carnervale, 1999; Jencks, 1992; Murnane & Levy, 1996; Ogbu, 1994). Each of the studies included in this review might be categorized by Gutierrez as “gap-gazing” in that none addresses potential solutions for the problem. Rather they focus on the documentation of the gap. I intend my study to specifically move away from “gap gazing” and move toward research that directly applies to policy specific for urban schools responsible for schooling students of color and/or students classified as low-SES. Hedges and Nowell (1999) also examined changes in the race-based gap from 1965 to

1996. They stated that if the narrowing of the gap continued at the same rate, it would take 75 years to completely close. Ginsburg and Russell (1981) studied the pre-mathematical ability of preschoolers in Baltimore and found that upon entry into school, the pre-mathematical ability of both Black and White students were virtually the same. As schooling progressed, the differences grew larger. If, as a whole, the narrowing of the gap would take 75 years to be complete while the differences increase for cohorts of children as they progress through school, we have a moral imperative to study ways in which to address the differences in mathematics achievement for students of color, not just point to the problem. I argue that teacher quality may be the most important policy related issue that holds the promise of addressing context – based differences in the learning of mathematics.

Product: Student Learning

The majority of educational research related to teacher effects on student learning has focused on process-product variables (Rowan, 1997). Process variables can include the observable classroom interaction that occurs between students, teachers, and the content in which the instruction is focused. Cohen and Ball (1999) describe this phenomenon as an interactive view of instructional capacity. This view holds that while the curricula and teacher characteristics are important, they must be viewed in connection with students rather than in isolation. Instruction requires all three of the elements to interact in real time. This seemingly simple idea has been very influential in research examining school and classroom effects on student learning. In addition, this interactive process provides a framework for examining the effect of mathematical knowledge for teaching on instructional practice in the context of student learning.

For the purposes of the review of the literature related to student learning, I had to widen my criteria for inclusion beyond mathematics to better understand how teacher effects are empirically explored. In other words, studies that compare different methods for isolating teacher effects, not always specific to mathematics are included. One study specific to mathematics supports the use of hierarchical linear modeling in their discussion of the amount of variance in student learning attributable to the teacher. Nye, Konstantopoulos, and Hedges (2004) found that 11% of the total variability in student achievement gains in mathematics could be attributed to teachers. These findings were based on a meta-analysis of correlational studies relating teacher quality and student learning in mathematics. The authors also performed an analysis of the original Tennessee class size study and determined that teacher differences accounted for 12 to 14% of the total variability in student gains in mathematics.

Multiple methods exist for modeling teacher effects on student learning. The two most common approaches are the use of a covariate adjustment model and the student gains model (Rowan, Correnti, & Miller, 2002; McCaffrey, Lockwood, Koretz, & Hamilton, 2003). The student gains approach is supported as a more accurate view of student learning over time (Rogosa, 1995), but is lambasted over reliability issues (Schaeffer, et al., 2007). The covariate adjustment model is said to merely model student achievement (Rowan, Correnti, & Miller, 2002) while controlling for prior achievement. The best approach is considered to be the use of growth curves for individual students (Rowan, Correnti, & Miller, 2002). However, this approach requires at least three waves of student data. Given that I only have two waves of student data and the forms used at both time points are not equated, I have elected to use the covariate adjustment model.

Additionally, my decision is supported by the recommendations of the panel that wrote *Using statistics effectively in mathematics education research* (2007).

Progress has been made in the mathematics education community in relation to a more thorough understanding of mathematical knowledge for teaching and the introduction of a measure for capturing the latent variable. Additionally, reform-oriented, standards-based teaching practices has evolved, and been distilled into nine clearly defined dimensions. Statistical tools have emerged with a keener understanding of the most effective ways to isolate the effects of teachers and teaching on student learning. The advances in these three areas, mathematical knowledge for teaching, standards-based instructional practice, and student learning, provide the impetus for this study. Brophy (1991) discusses the teacher with less explicit and connected knowledge as more likely to teach the content as static and favor seat-based assignments, while the teacher with better connected knowledge uses multiple representations and treats the subject as dynamic. My study attempts to analyze this broad statement empirically and specifically for teachers of mathematics by exploring the teacher *presage* and *process* variables that might contribute to the *product* variable student learning while also paying special attention to student *context* variables.

Chapter Three: Methodology

Participants and Setting

The participants in this study were from a large urban district in the southeastern United States. The data were collected from the grade three students nested within 35 grade three teachers who are, in turn, nested within 17 public elementary schools. The district has a student population that is predominately African American and participating in the free and reduced lunch program. The sample, while primarily African American, also consists of Caucasian, Hispanic, Asian, and students identified as multi-racial. Approximately 66.1% of the students receive free lunch and 4% pay a reduced lunch rate. The concentration of students participating in the free or reduced lunch program varies by school along with the concentration of racial groups based on the communities and neighborhoods served by the school. The teacher sample was recruited using principals, teachers, math coaches, and district mathematics leadership personnel. The corresponding students of these teachers form the student sample component of my study.

Measures

Presage: Mathematical knowledge for teaching. The Learning Mathematics for Teaching (LMT) Instrument has been developed to measure a teacher's mathematical knowledge for teaching related to specific content strands taught in both elementary and middle schools such as number concepts and operations, algebraic reasoning, and geometry and measurement (Hill, Schilling & Ball, 2004). Assessment forms are further delineated into three constructs with unique test forms for each content strand assessed. The constructs are: (a) common content knowledge; (b) specialized content knowledge;

and (c) knowledge of content and students. Multiple forms were developed and equated for each content strand. Factor analysis was utilized to ensure that the items that were written to assess certain constructs within each content area did in fact fit together under that construct (Hill, Schilling, & Ball, 2004). After a full examination of the state curriculum objectives for grade three mathematics, I decided to use a measure of number concepts and numeration because it most closely aligned with the student curriculum objectives. My next decision related to the three types of constructs: (a) common content knowledge; (b) specialized content knowledge; and (c) knowledge of content and students. The 1-PL, or Rasch, reliability measures ranged from .58 to .67 for knowledge of content and students and from .74 to .81 for specialized content knowledge. Given the moderate reliability of the forms for knowledge of content and students, I decided to use a form for the construct specialized content knowledge. I narrowed my selection to one form given its' 1-PL reliability (.81) and the point at which max information may be obtained from the measure (-0.97). This means that the test is best suited for teachers that are slightly below the national norm established by the representative sample in the scale development studies. The measure selected provides the most information when properly targeted for teachers below the national norm. I selected the *2002 Form A Elementary Number Concepts and Operations – Content Knowledge* form for use in this study.

The raw scores are converted to standardized z-scores using a lookup table provided by the developers wherein each score is representative of the number of standard deviations away from the mean of the nationally representative sample used to scale the measure. A score of zero is therefore representative of a teacher with average mathematical knowledge for teaching. A score above zero is considered to be that of an above average teacher and a score below zero is that of a below average teacher in the area of mathematical knowledge for teaching. The greater the distance above zero in standard deviations, the deeper the mathematical knowledge for teaching of the assessed teacher. The form consists of 24 multiple-choice items. Of the 24 total items, 17 of the items are testlets with one stem and 4 to 5 questions connected to that one stem. The standardized scaled scores range from -3 to 3. Use of the *Learning Mathematics for Teaching* instrument requires an agreement not to release items used in the measure; therefore, appendix A includes released items representative of the items used on the scale.

Table 3 presents the sample specific reliability coefficients for the instruments used in my study to measure teacher traits. A traditional indication of internal reliability, the Cronbach's alpha coefficient of reliability (.78) for the form that I selected to measure mathematical knowledge for teaching is considered moderately high. Additionally, I included the Rasch, or 1-PL reliability of person separation statistic (.80), also considered a good measure of internal reliability. Both values, Cronbach's alpha and the Rasch reliability of person separation, indicate that the measure is consistent and captures a unidimensional latent construct, respectively.

Process variable: Teaching practices. Because of the quantitative nature of this study, a survey was used to estimate teachers' implementation of standards based reform in mathematics. Ross, McDougall, and Hogabaum-Gray (2002) developed a survey instrument made up of 20 items using a 6-point Likert scale from strongly disagree (1) to strongly agree (6). The survey measures dimensions of the implementation of standards based reform methodologies. This measure was selected to directly address the need for future research identified in Hill, Rowan, and Ball (2005) connecting mathematical knowledge for teaching to specific teaching practices in mathematics such as representation, explanation, and interaction with students. The dimensions include teachers beliefs related to (a) program scope, (b) student tasks, (c) discovery, (d) teacher's role, (e) manipulatives and tools, (f) student-student interaction, (g) student assessment, (h) teacher's perception of math as a discipline, and (i) student confidence. The authors report the reliability coefficient (alpha) for the survey instrument at .81 with a large sample of over 1000 teachers.

Table 3 presents the sample-specific reliability coefficients for the two teacher-level instruments that I used in my study. The measure that I used to capture teaching practices has a Cronbach's alpha coefficient of .78. This indication of internal consistency is somewhat lower for my sample than it was for the original sample wherein this measure was developed. Similar to the reliability examination that I conducted for the measure of mathematical knowledge for teaching, I also used Rasch measurement theory to determine the reliability of item separation (.93). I used the Rasch rating scale model and the results indicate that the measure possesses a large difficulty range, thus resulting in such a high reliability of item separation. Also, item separation is an indicator that my sample is large enough for this measure to precisely locate people on the latent variable, use of standards-based, reform-oriented teaching practices.

Additionally, the authors conducted three validity studies, each specific to predictive, concurrent, and construct validity concerns. Each study resulted in evidence supporting the validity of the instrument. The authors of this measure recommend that raw data from this survey across all teacher participants be partitioned into quartiles.

This delineation results in four categories of implementation of standards-based reform methodologies. The first and the fourth quartile are identified as low reform or high reform respectively. The higher the score on the measure, the more likely that the teacher implements standards-based, reform-oriented practices with their students. For the purposes of this study, the total raw score is used as a continuous variable as opposed to the transformation of the raw score into quartiles, or categorical data. Because of the teacher sample size proposed in this study, the utilization of the survey allows a glimpse into classroom practice without the time or expense that would have been incurred during a large-scale observational study. Appendix B contains the measure of teaching practices used in this study.

Context variable: Student race/ethnicity, class status, and prior achievement. The measures of student race/ethnicity and class status were collected from the district's data management system. Students' race/ethnicity can be one of the following categories: a) Hispanic, b) African- American, c) Caucasian, d) Asian, or e) Multi-racial. I am specifically interested in the likelihood of African American children being taught by teachers who possess mathematical knowledge for teaching and implement standards-based, reform-oriented practices. For this purpose, I created a dichotomous dummy variable wherein African American students are indicated with a value of 1.

Student class status was collected from the district's data management system. Class status is estimated using a student's eligibility for the free or reduced lunch program. Class status is, therefore, a dichotomous variable wherein the two possible values are students who are eligible for free or reduced lunch and students who are not eligible for free or reduced lunch.

Context variables play a critical role in this study. However, they are not included when examining the teacher effect on student learning. My examination of teacher effects, specifically mathematical knowledge for teaching and standards-based, reform-oriented teaching practices on student learning does not include covariates such as race or class, often used in an attempt to control for non-schooling factors that might bias the teacher effect estimates. McCafferey, Lockwood, Koretz, and Hamilton (2003) indicate that the teacher effect bias is small at the student-level, but sometimes larger at the aggregate levels of school or classroom. Given my limited sample size and resulting loss in statistical power, I have elected to only use student prior learning as a covariate in my teacher effects model for research question two. The inclusion of prior learning is necessary given the use of covariate adjustment. As a result, teacher effects may be misestimated given that students in my sample are stratified by the student-level covariates race and class based on the housing patterns in the communities that the schools serve.

Teacher race is also not included in this study. The variation in teacher race in my teacher sample is limited given that most of the teachers are African American. Additionally, by using Dunkin and Biddle’s (1964) model, I am focusing on the context variables of the students. In other words, I remain focused on how the student race, class status, and prior learning in mathematics might affect the staffing within schools and classrooms.

Product variable: Student learning. The measures of student learning that I used in this study was drawn from the state’s mandated criterion-referenced competency assessment administered annually. Student scale scores from their grade two assessment, administered in the spring of 2008, and their grade three assessment, administered in spring 2009 are be included in the model. Each assessment is aligned to the state mandated mathematics curriculum standards in the given grade level. This assessment is a high-stakes assessment in grade three. Students who do not meet the proficiency score set by the state can not be promoted to grade four.

Both the grade two, spring 2008 administration and the grade three, spring 2009 administration have scale scores ranging from the lowest obtainable scale score of 650 to the highest obtainable scale score of 930 in grade two and 990 in grade three. The performance expectations have been set by the state board of education such that a score below 800 is deemed *Below Expectations,* a score between 800 and 849 is deemed *Meets Expectations*, and a score at or above 850 is deemed *Exceeds Expectations*. The reliability indices reported for the state assessments are Cronbach's alpha reliability coefficient and the standard error of measurement (SEM). The state reports reliability indices for the spring 2008 administration of grade two mathematics with a Cronbach's alpha of .91 and a SEM = 2.98. The grade three, spring 2009 administration of the CRCT in mathematics is reported to have a Cronbach's alpha of .92 and a SEM = 3.07. The se data are displayed in Table 3 along with the reliability information for the teacher trait measures. An important difference in the administration of these two assessments to note is that the grade 2 assessment is read to students, verbatim, by the classroom teacher while in grade three, students are expected to read the assessment independently.

For research question one, a correlational analysis was employed to better understand the relationship among aggregate levels of student race, class, or prior learning and teacher quality. The mean scale score (CRT08) by classroom is used as a correlate with teacher-level characteristics. The mean scale score (CRT08) by school is used as a correlate with school-level means of teacher characteristic variables.

For the purposes of the second research question in this study, I use the second administration of the assessment (grade three, spring 2009) as my outcome variable and the first administration (grade 2, spring 2008) as a covariate to create a covariate adjustment model. This model allows for an understanding of how students who begin at the same level of prior knowledge might differ in their achievement in mathematics due to teacher-level variables. Alignment between the teacher knowledge being assessed, the written curriculum, and the student knowledge being assessed was cited as a limitation of Hill's work connecting teacher knowledge to student achievement (Hill, Rowan, & Ball, 2005). This alignment is addressed in the design of this study.

Data Collection and Analyses

Student level achievement data were collected at two time points, the first being the spring 2008, grade two administration of the state criterion-referenced assessment and the second being the spring 2009, grade three administration. Student level data were matched with the teacher participants in the study. Students were nested within teachers. Only students had two student learning data points are included in the study. Originally, data were collected from 608 students. Seventy-seven of those students did not have a Spring 2008 score and were subsequently removed from the student sample. The removal resulted in each student having two scale scores reported from the district, a spring 2008 (CRT08) and a spring 2009 (CRT09). The *presage* variable mathematical knowledge for teaching (MKT) and the *process* variable teaching practices (SB-RO) was collected using the aforementioned instruments in the months of October 2008 through

March 2009 for my empirical study, prior to the spring 2009, grade three administration of the state criterion-referenced test. The district queried the student data management system for the student scale score for both years of the measure along with the student *context* variables race/ethnicity and class status.

For research question one, a correlational analysis was performed. Theoretical arguments have been made that a relationship exists between the assignment of an effective teacher and aggregates levels of student race, class, and prior learning. In an effort to explore the relationships among these student-level contextual variables and the teacher-level presage and process variables, I employed a series of correlational analyses. As classroom aggregates of the concentration of African-American students increase, what is the association with the staffing of that class with a teacher possessing high levels of mathematical knowledge for teaching? What about the association between the aggregates of student race, class, and prior learning and the teacher's use of standards-based, reform-oriented teaching practices? Additional correlational analyses will be run examining the relationship between the proportion of students eligible for free or reduced lunch and each of the two teacher-level variables, and between the mean prior year CRCT score and each of the two teacher-level variables.

Pearson product-moment correlation coefficient (*r*) is calculated for 12 potential relationships. Six calculations of Pearson's *r* are employed at the classroom level using class aggregates of student variables (the proportion of African American students, the proportion of students eligible for free or reduced lunch, and the mean prior performance in mathematics) with teacher-level scores on the two teacher-level variables. Six additional correlation coefficients of Pearson's *r* are calculated at the school level.

Student-level variables are aggregated across the classrooms within each school to produce school-level aggregates of the three student variables. Additionally, school-level aggregates of the two teacher-level variables are calculated within each school in the sample. Pearson *r* is calculated for the six potential school-level relationships.

For research question two, a two-level hierarchical linear model was used as students were nested in teachers. As a first step, the variance in the outcome variable *CRT2* (2009 scale score) was estimated at the student and teacher levels. A similar variance decomposition model was used to estimate the variance at two levels for the 2008 (*CRT1*) student measure. Finally, a two-level hierarchical linear model was employed. The 2009 student assessment scale score (*CRT2*) was the dependent variable and the 2008 scale score (*CRT1*) was included as a predictor at level 1. At the teacher level, *MKT* and the use of standards-based, reform-oriented teaching practices (*SB-RO*) were included as predictors. Grand mean centering was employed for both student-level variables and for the teacher-level variable standards-based, reform-oriented (SB-RO) teaching practices wherein a value of zero is indicative of the proficiency of an average student (*CRT1* = 0) or an average teacher (*SB-RO* = 0). Since mathematical knowledge for teaching scores possessed a meaningful zero, a score indicative of the knowledge of an average mathematics teacher, MKT was not grand-mean centered. The following model was used for research question two:

Level 1 model Student Level (within teacher):

$$CRT2_{ij} = \beta_{0j} + \beta_{1j}\,(CRT1) + e_{ij} \qquad (1)$$

Level 2 model Teacher Level (between teachers):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(MKT) + \gamma_{02}(SB\text{-}RO) + r_{0j}$$

(2)

$$\beta_{1j} = \gamma_{10} \qquad (3)$$

Where $CRT2_{ij}$ is the 2009, grade three criterion-referenced test scale score for student *i* of teacher *j* and β_{0j} is the adjusted 2009 score for students with teacher j. γ_{01} and γ_{02} are the partial effects of MKT and SB-RO, respectively on the adjusted spring score.

Limitations

The potential confounding factors include equitable access to resources at the school and classroom level, level of implementation of the curriculum across classrooms, teacher preparation, and student attributes that have been shown to influence student learning. Additionally, honest responses from teachers around their implementation of a standards-based methodology are not guaranteed.

The external validity threats include that this study is conducted within the boundaries of a state and that the assessment of student learning is based on a measure designed specifically to align with the state-mandated curriculum standards. The results related to student learning may not be generalizeable on a national assessment or in states with different content standards.

The limitations of instrumentation for teachers include the inability to directly match the content assessed for teachers and students for the LMT and the state criterion-referenced test. This limitation was also cited in earlier work conducted by Hill, Rowan, and Ball (2005), so I used a crude framework of curricular alignment attempting to select an instrument of mathematical knowledge for teaching that most closely aligned with the grade three standards in the state in which my study was conducted. A major limitation of the survey for measuring implementation of standards-based reform in that the response of a teacher may not be directly linked to the practices of the teacher in the classroom in real time. In other words, a teacher may report the use of standards-based, reform-oriented practices while not actually using them in the classroom. A qualitative approach to measuring classroom practices would serve to provide a more authentic capture of teacher practice through observation and interview protocols for a randomly selected subset of the teacher participants.

In January 2010, the Governor's Office of Student Achievement (GOSA) released a report wherein erasures were examined across the state. Specifically, the number of student erasures and the frequency of the erasures wherein an answer choice was changed from wrong to right was compared the state frequency. The report placed schools on a "severe concern" list when the frequency of wrong to right erasures was more than four standard deviations above the state mean wrong to right erasures. Nine of the 17 schools in my study (n = 9) found themselves on the severe concern list for the Spring 2009 administration of the state's CRCT. If, in fact, the erasures are indicative of cheating, the Spring 2009 student-level data may be elevated. However, no evidence has been produced substantiating cheating in the flagged schools.

Sample size is also a limitation of this study. In addition, the sampling of teachers within one school district may bias the results. The sample size may limit the statistical power of the analyses and the potential sampling bias resulting from the convenience sampling method must be considered when interpreting the results of the study. Because of the sample size conundrum, future analyses may include a Bayesian analysis, more appropriate when the sample size is small. Bayesian analysis is controversial in that it allows information related to the distribution of variables found in the population to be included when examining the distribution of the same variable in the sample. Although the Bayesian analysis' inclusion of prior evidence is beneficial when sample sizes are small, the inclusion of information not evidenced by the researcher is controversial. The lack of variation in the student contextual factors, specifically student race, may also be a limitation.

Chapter Four: Results

Table 4 provides a detailed view of the 17 schools included in the sample. School-wide data such as the proportion of African-American students, the proportion of students eligible for free or reduced lunch, and the mean CRT 2008 scores were collected from the publicly available school report cards. The number of teachers in the sample from each school ranges from one to four. The n-count of students included in the sample from each of the schools ranges from nine to 65 with a total of 531 students included in the sample. The proportion of African-American students within each school ranged from 13% to 100%, with 12 of the schools serving a student body wherein more than 80% of the students were African American. The proportion of students who were eligible for free or reduced lunch ranged from 13% to 100% with 13 schools serving a student population wherein more than 80% of students were eligible for free or reduced lunch. Three schools in the sample have mean CRT 2008 scores less than the scale score cut (800) for proficiency on the state mandated assessment. Mean CRT 2008 scale scores range from 783.81 to 858.29 ($SD = 22.83$, $SD = 27.13$ respectively) across the 17 schools. Mean MKT and SB-RO practices scores were calculated from the sample. Mean MKT scores range from -1.96 to 0.24 with only one mean MKT score greater than zero. The mean SB-RO teaching practices score ranged from 71 to 111.

Table 5 provides detailed summary information for each of the original 37 teachers and their classroom demographics. After linking students to teachers using the district data management system, it became necessary to remove two teachers from the sample. Teacher 12 only had one student included in the data management system and that student was missing his or her CRT09 score, so teacher 12 was removed from the sample. Teacher 22 had no students included in the data management system and was therefore removed from the sample. Both teachers are included in Table 5 without student information. The student n-count in this analysis ranged from nine for teacher 35 to 54 for teacher 22. We must assume that teacher 22 was responsible for teaching multiple classes of grade three mathematics, perhaps in a departmentalized model, thus resulting in the high number of students assigned to teacher 22 in the data management system. Students who did not have both time points, 2008 and 2009, were removed from the analysis. A total of 531 students, across 35 teachers and their classrooms, in 17 schools remained.

Table 6 shows sample means and standard deviations for the variables included in this analysis. The mean scale score for both the CRT08 (M = 828.73, SD = 31.10) and the CRT09 (MEAN = 833.16, SD = 49.31) should be considered within the context of the performance categories and their related scale cut scores. For both years, a scale score less than 800 is indicative of a student who doe not meet the standard. A scale score between 800 and 849 is indicative of a student who does meet the standard. Finally, a score greater than or equal to 850 is indicative of a student who exceeds the standard. Across all 531 students included in the sample, 79.3% are African American and 75.5% are eligible for free or reduced lunch.

Teacher level descriptive statistics require substantive interpretations. The mean mathematical knowledge for teaching (MKT) for the sample is -1.00, in other words, the teachers in this sample scored approximately one standard deviation below the average teacher in the national sample used to scale the measure. The average score on the teaching practices measure was 90.26. As reported previously, Table 3 displays sample specific reliability coefficients for the two teacher measures used herein. Both measures reported moderately high internal reliability whether using the traditional Cronbach's alpha or the Rasch reliability of item separation.

Figures 3 and 4 show the distribution of student-level CRT scores in 2008 and 2009, respectively. Important to note is that the 2008 scores are indicative of achievement status of students prior to entering a teachers' grade three classroom while 2009 scores are after the hypothesized teacher effect has occurred. Student CRT 2008 and CRT 2009 scale scores are significantly correlated (r (529) = .690, $p < .01$) such that as students' CRT 2008 scale score increases, so does their CRT 2009 scale score.

Table 7 provides a clear illustration of how the sample's level of mathematical knowledge for teaching compares to the average teacher's mathematical knowledge for teaching based on Hill, Schilling, & Ball's (2004) national sample. I used a Rasch measurement model to calibrate the teachers' scores on the LMT measure in an effort to better understand the distribution of mathematical knowledge for teaching of my sample. Please note that the column labeled measure is the Rasch measure of where the participants fall on the ruler, if you will, measuring the latent variable mathematical knowledge for teaching. The range of the Rasch measure, or theta, is from -3 to 3. Only five teachers in the sample possessed the level of mathematical knowledge for teaching above the national average. Therefore, 30 teachers, or 86% of the sample possessed levels of mathematical knowledge for teaching below the national average.

Figure 5 illustrates the relationship between teacher MKT and SB-RO using a scatter plot and including the line of best fit. Teacher MKT and SB-RO are positively and significantly correlated (r (33) = .298, $p < .01$). However the relationship does appear to be weak. Only 8.9% of the variance in SB-RO can be attributed to a teacher's MKT ($R^2 = .089$).

Research question one (Are there any associations between the ethnic, SES, and ability compositions of classrooms or schools and teachers' level of mathematical knowledge for teaching and use of standards-based, reform-oriented teaching practices?), required a correlational analysis consisting of a total of twelve calculations across two levels of aggregates, classroom and school. In order to address the classroom-level component of the question, aggregates of the proportion of African-American students, the proportion of students eligible for free or reduced lunch, and the mean prior year's CRT score were calculated for each of the 35 participants.

Table 8 details the correlational analysis completed for research question one. MKT appears to decrease as the mean prior year's CRT score increases within a classroom ($r(33) = -.12$, ns). Similarly, it appears that as the proportion of African American students increases at the classroom level, MKT decreases ($r(33) = -.03$, ns). Finally, as the proportion of students eligible for free or reduced lunch increases, so does MKT. SB-RO appears to increase as the mean prior year's CRT score increases ($r(33) = .23$, ns) and as the proportion of African American students increases ($r(33) = .14$, ns). Also opposite of what occurred with MKT, as the proportion of students eligible for free or reduced lunch increases, SB-RO teaching practices decrease ($r(33) = -.02$, ns). None of these correlations were significant, meaning that they could have occurred by chance.

The second part of research question one deals with the school level aggregates of the proportion of African American students, the proportion of students eligible for free or reduced lunch, and the mean prior year's grade two CRT scale score. These aggregations were gathered from the school-based report cards for the 2007-2008 academic year (www.gadoe.org). Given the limited number of teachers within each school (mean = 2 per school) I chose not to aggregate using my data set as this would only provide a limited view of the school-wide demographics. I chose, instead, to use the school-wide data that were reported for Adequate Yearly Progress purposes during the 2008-2009 academic year. As the mean CRT scale score for second graders in 2008 decreased, both MKT and SB-RO teaching practices decreased (r (15) = -.07, ns and r (15) = -.01, ns. respectively). As the proportion of both African American students and students eligible for free or reduced lunch increased at the school level, the MKT increased for the teachers in my sample (r (15) = .12, ns and r (15) = .08, ns, respectively). The same directionality was evident at the school level for SB-RO practices, wherein the use of SB-RO teaching practices increased as the proportion of African American and students eligible for free or reduced lunch increased (r (15) = .18, ns and r (15) = .07, ns, respectively).

Research question two (What is the relationship among mathematical knowledge for teaching, the use of standards-based, reform-oriented teaching practices, and student learning outcomes in mathematics?) was approached using hierarchical linear modeling. Table 9 presents the results of unconditional models that decomposed the variance in student CRT08 scores and student CRT09 scores into that residing among teachers and among students within teachers. For student CRT08 scores, the largest amount of variance, 79% resided within teacher, or between students. The remaining variance, 21%, resided between teachers. For student CRT09 scores, the largest amount of variance, 65% resided within teacher, or between students. The remaining variance, 35%, resided between teachers. The percentage of variance found between teachers increased in CRT09 scores when compared to CRT08, from 21% to 35%, respectively. This finding illustrates that teachers could have a larger influence on the variability of the CRT09 scores than on the CRT08 scores. Additionally, the within teacher variance is larger than what was found by Hill, Rowan and Ball (2005) to be only 2% in grade three, but they used a three-level model wherein students were nested within classrooms that were nested within schools. The amount of variance found at the teacher-level in this study's model supports the use of hierarchical linear modeling.

For research question two, I ran multiple models. Initially, I ran two two-level models, one for each teacher-level variable (MKT and SB-RO) included as the sole predictor variable at the teacher level:

Level 1 Student Level (within teacher):

$$CRT2_{ij} = \beta_{0j} + \beta_{1j}(CRT1) + e_{ij} \quad (4)$$

Level 2 Teacher Level (between teachers, within schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(teacher\_variable) + r_{0j}$$

(5)

$$\beta_{1j} = \gamma_{10} \quad (6)$$

where β_{0j} is the mean CRT2 score of the students of teacher *j* adjusted for prior learning (CRT1) and γ_{00} is the overall adjusted CRT2 score when the teacher-level variable is equal to zero. Therefore γ_{01} is my coefficient of interest in that it represents the change in the adjusted CRT2 score for every one unit increase in the selected teacher variable. Recall that prior to entering my data into HLM for analysis, I grand-mean centered the teacher-level variable SB-RO teaching practices such that zero is meaningful. In the case of MKT, a value of zero is representative of the MKT of an average teacher in the national sample. In the case of SB – RO teaching practices, zero is equivalent to the average reported use of such practices for the teachers in my sample. In this case, centering provides a concise interpretation of γ_{00} in each model such that γ_{00} is the adjusted CRT2 score for an average teacher (teacher_variable = 0). I tested the hypotheses that each of the teacher-level variables (MKT and SB-RO) was unrelated to student learning as defined by the CRT2 score adjusted for prior learning (CRT1).

Table 10 displays the model coefficients, standard errors, and degrees of freedom for two models. Each model includes only one teacher-level variable at the second level of the hierarchical linear model as described previously. When a teachers' mathematical knowledge for teaching is equivalent to that of an average teacher, the adjusted CRT2 score is 832.72. For every one-unit increase in a teacher's mathematical knowledge for teaching, there is a 1.25 point decrease in the student's adjusted CRT2 score. This relationship is not statistically significant however, and could have occurred by chance. When a teachers' reported use of standards-based, reform-oriented teaching practices is equivalent to that of an average teacher in the sample, the adjusted CRT2 score is 833.97. For every one-unit increase in a teacher's reported use of standards-based, reform-oriented practices, there is a .06 increase in the student's adjusted CRT2 score. This relationship is not statistically significant however, and could have occurred by chance. Evidence is not sufficient to reject the null that either MKT or SB – RO teaching practices are unrelated to student learning.

I also ran a model including MKT and SB-RO teaching practices as predictors at the teacher level simultaneously. The model was as follows:

Level 1 Model Student Level (within teacher):

$$CRT2_{ij} = \beta_{0j} + \beta_{1j}\,(CRT1) + e_{ij} \tag{7}$$

Level 2 Model Teacher Level (between teachers):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(MKT) + \gamma_{02}(SB\text{-}RO) + r_{0k} \tag{8}$$

$$\beta_{1j} = \gamma_{10} \tag{9}$$

where β_{0j} is the mean CRT2 score of the students of teacher *j* adjusted for prior learning (CRT1) and γ_{00} is the grand mean CRT2 score when both MKT and SB-RO practices are zero, or when the teacher is average in both MKT and SB-RO practices given the sample statistics, γ_{01} is the partial effect of MKT on the adjusted CRT2 score when controlling for SB-RO practices, and γ_{02} is the partial effect of SB-RO practices on the adjusted CRT2 score when controlling for MKT. Table 11 presents the model coefficients, standard error, and the degrees of freedom for the model. As seen in the previous model, neither teacher-level variable had a statistically significant effect (γ_{01} = -1.59, p = .68; γ_{02} = 0.11, p = .75), even when controlling for the other teacher-level variable. Enough evidence did not exist to reject the null that either variable was unrelated to student learning.

Chapter Five: Discussion

My sample consisted of 17 schools housed in an urban district in the southeastern United States. As previously reported, Table 4 details the characteristics of school-level data. Restriction of range appeared to be an issue in the sample at the school level. Although ranges of values appeared variable, the proportion of African American students within a school ranged from 23% to 100% for example, the sheer number of schools with high proportions of African American students overshadowed schools with more diversity. Only four schools serve a student body that is less than 80% African American. This restriction of range limits the statistical ability to tease out potential relationships or effect sizes of teacher-level variables.

Restriction of range at the classroom level was also evident. Classroom demographics, in general, appeared to mirror the school-level demographics, resulting in a similar lack of variance in regard to student race and class status. Student performance, when comparing the CRT08 scale score to the CRT09 scale score at the teacher level resulted in 19 teachers, or 54% wherein the mean CRT score increased. Simultaneously, the standard deviation increased for 30 of the teachers, or 86%. So, although a slight majority of classroom teachers in this study saw gains in their mean scale score, a majority of teachers also experienced an increase in the variance of their student scores within their classroom. I would have expected the spread of the data to decrease after one year of instructional time with a mutual teacher.

The mean mathematical knowledge for teaching score for the sample in this study is one standard deviation below what Hill, Rowan, & Ball (2005) found in their nationally representative sample of 365 grade three teachers from 42 districts in 15 states.

Hill, Rowan, & Ball (2005) included teachers from suburban, urban, and urban fringe communities, deliberately oversampling teachers working in high poverty schools. High poverty was defined by the researchers as schools with larger than the average proportion of high poverty students (13%) citing findings from another study indicating that the average school included approximately 13% of students in high poverty (Benson, 2002). My sample is not equivalent to the former study in that my teacher sample is smaller and responsible for teaching only in an urban district in a southern state. Additionally, the definition of high poverty in the former study, greater than 13%, is not equivalent to the poverty level of students in my sample where 71% of the students participate in the free lunch program. As evidenced by the distribution of the levels of mathematical knowledge for teaching when using the Rasch measurement model, 86% of the teachers serving this high minority, high poverty population of students possessed levels of mathematical knowledge for teaching below the national average. Increasing the mathematical knowledge for teaching of my teacher sample by one standard deviation puts this sample of teachers at what is considered average grade three teacher knowledge. This significant finding supports previous research finding that students in high poverty, high minority schools may receive instruction from less knowledgeable teachers (Haycock, 1998; Hill, Rowan, & Ball, 2005).

Allegations have been made against 14 of the schools included in this sample of having irregular erasure patterns by the Governor's Office of Student Achievement (GOSA), specifically a larger number of wrong to right erasures on the 2009 CRT student data. These allegations have lead to an investigation wherein five of the 17 schools have been identified as having testing irregularities beyond the irregular erasure patterns.

Table 12 details each sample school's classification in both the GOSA investigation and the second Blue Ribbon Commission (BRC) investigation. As a result, the CRT 2009 scores for these students may be inflated. If allegations of test tampering are substantiated, the tampering may have also been present during the 2008 administration of the CRT, also resulting in inflated student scores.

I hypothesized for research question one (Are there any associations between the ethnic, SES, and ability compositions of classrooms or schools and teachers' level of mathematical knowledge for teaching and use of standards-based, reform-oriented teaching practices?) that as the proportion of African American students, students eligible for free or reduced lunch, or students who are less mathematically proficient increases that the teachers staffed in the classrooms and within the schools serving these students will less likely be in possession of mathematical knowledge for teaching and be less likely to implement standards-based, reform-oriented practices. The directionality of the correlations may be of interest; however, given that none of the relationships were found to be statistically significant, they could have occurred by chance. Only three of the 12 correlation coefficients that were calculated in this study illustrated the directionality that I expected. All three occurred at the teacher, or classroom level. As the proportion of African-American students increased, a very small (r (33) = -.03, ns), negative association was found with MKT. Similarly, as the proportion of students who are eligible for free or reduced lunch increased at the classroom level, a teacher's implementation of SB-RO teaching practices also decreased, again with a very weak association (r (33) = -.02, ns). Finally, the largest, but still not statistically significant, association was found at the classroom level between prior learning and a teachers use of

SB-RO teaching practices (r (33) = .23, ns).

Although my findings for research question one as it related to the aggregate levels of student race, social class, and prior learning in mathematics were not statitistically significant, the overall low levels of MKT in my sample is important to note. The teachers in the sample served a population of students that were predominately African American (79%) and predominately students eligible for free or reduced lunch (71%). Eighty-six percent (86%) of the teachers' who were responsible for the mathematics' teaching of this population of students possessed levels of MKT below the national average. Given the correlation between MKT and a teachers' use of SB-RO teaching practices, it is reasonable to argue that teachers with lower levels of MKT report less use of such practices.

I hypothesized for research question two (What is the relationship among mathematical knowledge for teaching, the use of standards-based, reform-oriented teaching practices, and student learning outcomes in mathematics?), given the previous work done by Hill, Rowan, and Ball (2005) in relating a teacher's mathematical knowledge for teaching to student learning outcomes in mathematics, that MKT would have a significant partial effect on student learning in mathematics. Based on my own work relating teacher knowledge and teaching practices (Sullivan, 2009), I expected SB-RO teaching practices to have a significant partial effect on student learning in mathematics. Given the increased understanding of the psychometric properties of the student measures to be used in this study, I also expected estimates of effect sizes to be more precise than in my previous work using raw scores on a district-level assessment (Sullivan, 2009).

My findings for research question two indicated that neither mathematical knowledge for teaching nor the reported use of standards-based, reform-oriented practices significantly predicted student learning outcomes as measured by the state's assessment. In fact, the direction of the relationship between mathematical knowledge for teaching and student learning was the opposite of what my own previous work as well as that of Hill, Rowan, and Ball (2005) had found. As a teacher's mathematical knowledge for teaching increased, the adjusted student learning outcome decreased. Although the finding is not statistically significant, and therefore may have occurred by chance, the directionality of the relationship remains surprising.

One potential argument supporting the lack of statistical significance of either teacher-level variable on student learning outcomes might be because of the difference in the student assessment administrations at grade levels that I elected to study. The grade two state assessment is read aloud to students while the grade three assessment is not read aloud. Additionally, the grade two assessment only includes three answer choices from which students must select the correct answer choice. The grade three assessment includes four answer choices from which the students must select the correct answer choice. The difference between the administration and the design of these two assessments may have resulted in a fatal flaw for this study. Additionally, the controversy surrounding the CRT09 scores causes one to question the validity of the student-level results, even with the increased psychometric information available for the measure, nullifying much of why the assessment was selected for inclusion in the study as the most authentic representation of student learning available.

Two important findings emerged from this study. The first was related to the low levels of MKT possessed by teachers who served a population with a high minority, high poverty students in the southeastern United States. Although the classroom or school aggregated were not related to the levels of MKT based on these data, the theory may still be evidenced when examining MKT levels situated within the national levels. Given the correlation of MKT with SB-RO teaching practices, low levels of MKT could result in lower levels of use of these practices. The second important finding related to the variance between student learning outcomes at the teacher level. Thirty-five percent (35%) in student learning resided at the teacher level for the CRT09 scores. Although these data did not support that either MKT or SB-RO teaching practices contributed to the variance, evidence supports that teachers contributed to the variance in student learning. Teachers do matter and additional research should be conducted to determine the teacher knowledge, practices, or other characteristics that make-up that contribution.

Implications

The implications of this study reach into three distinct fields. Educational practice, policy, and research may find important next steps as a result of this study. Components of the study that could affect future work are based in what was found as well as what was not found alike.

Although neither student race, class status, nor prior learning in mathematics were significantly related to the teacher characteristics of knowledge and practices in this study, the line of inquiry is certainly not exhausted. The sample size of teachers, the lack of variability across schools with regard to student demographics, and perhaps the differences between the student learning measures may have clouded these findings. An alternative explanation might also be that the district has intentionally focused on the equitable distribution of teacher quality across schools and classrooms serving variable student bodies.

Any future research should address the measurement concerns associated with the student-level assessment measures. The entire premise of this study relies heavily on the ability of the measures to accurately capture teacher- and student-level attributes. Future work should include student measures that are more closely related in the ways in which they are administered. Perhaps looking at students and teachers in grade four such that the grade three assessment can be used to control for student prior learning. Additionally, in attempting to address a student measurement issue in my empirical study, I may have muddied the waters through the inclusion of a high stakes assessment in a district with high levels of accountability.

The use of hierarchical linear modeling when attempting to better understand and define teacher effect on student learning was supported in this study. Currently, the federal government, state education agencies, and local education agencies alike are struggling to design models that simultaneously describe student growth and estimate the effect of teachers on student learning. The ability to isolate the variability at multiple levels makes the use of hierarchical linear modeling ideal. Without employing multi-level models, teacher effect can be confounded with other influences, such as school effect.

Policy makers concerned with teacher quality, specifically the defining of a highly qualified teacher, might want to reconsider the characteristics currently employed by state agencies in the designation of teachers as highly qualified. Considering both my empirical study and my dissertation study jointly, it appears as if traditional proxy measures typically employed at the state level for teacher quality, such as years of experience, are no longer adequate. Content-specific indicators, such as mathematical knowledge for teaching and the used of standards-based, reform-oriented practices, while not significantly affecting student learning in this study, should continue to be empirically examined. In both of my studies, the empirical and the dissertation, knowledge and practice were correlated, and when using a student assessment instrument designed by the district as a benchmark, teaching practices had a significant effect on student learning. These findings, conjointly, should not be dismissed. The line of inquiry should, in fact, be extended in a continued effort to unpack the ways in which the professional knowledge of teachers is translated into practice, specifically within urban contexts.

Conclusion

Teachers contributed to the variance in student learning and therefore were an important factor in student achievement. The ways in which teachers contributed to variation between student learning outcomes should continue to be unpacked. Federal, state, and district level policy related to teacher quality should keep student learning in the forefront of policy initiatives as opposed to assuming that we know what teacher characteristics and practices matter the most. Advancements in statistical methodologies, a better understanding of the types of teacher knowledge and teaching practices that relate to student learning, and the intentional inclusion of student contextual factors must be considered by policy makers, researchers, and practitioners alike.

Educational researchers have a moral imperative to consider context variables in studies exploring the effects of teachers on student learning outcomes. In mathematics education, as argued by Tate (1997), the research focuses too much on students in general and not enough on students situated in urban learning contexts. Without an intentional focus on teacher quality for African American students, students eligible for free or reduced lunch, or students with limited mathematical proficiency, how can we truly engage in research meaningful to the districts responsible for the schooling of students of color and students in poverty? Policy implications related to ensuring all students have access to high quality teachers and the indicators of teacher quality should be considered within the context of the communities in which schooling takes place rather than only considering students in general.

References

Ball, D. L. and H. Bass (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.

Banks, C.A., and Banks, J.A. (1997). Reforming schools in a democratic pluralistic society. *Educational Policy*. 11(2): 183-193.

Benson, G. (2002). *Study of Instructional Improvement school sample design.* Ann Arbor: University of Michigan, Institute for Social Research.

Brophy, J.E., and Good, T. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), *Handbook of research on teaching*, 3rd ed. N.Y.: Macmillan.

Carnervale, A.P. (1999). *Education = success: Empowering Hispanic youth and adults.* Princeton, NJ: Educational Testing Services.

Cohen, D. K. and D. L. Ball (1999). Instruction, capacity, and improvement. *Consortium for Policy Research in Education*. Philadelphia, Pennsylvania, University of Pennsylvania.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., et al. (1966). *Equality of Educational Opportunity.* Washington, D.C.: U.S. Department of Education.

Congress, U. S. (2001). *No Child Left Behind*. Washington, D.C.: U.S. Department of Education. 107-110.

Cooney, T.J. (1980). Research on teaching and teacher education. *Research in Mathematics Education: Professional Reference Series*. Reston, VA.: NCTM.

CTB-McGraw-Hill. (2004). *Terra Nova CTBS.* Monterey, CA: McGraw-Hill Education.

Darling-Hammond, L. and G. Sykes (2003). Wanted: A National Teacher Supply Policy for Education: The Right Way to Meet The "Highly Qualified Teacher" Challenge. *Education and Policy Analysis Archives* 11(33).

Dunkin, M. and Biddle, B. (1974). *The Study of Teaching.* New York: Holt, Reinhart & Winston.

Edsel, Thomas. (2009, April 2). Michelle Rhee threatens end-run around teachers union. *The Huffington Post.* Retrieved April 27, 2009 from The Huffington Post. Web site: http://www.huffingtonpost.com.

Education Commission of the States. (2004). Highly Qualified Teacher Policy, Multiple State Reports and On-line databases. Retrieved October, 2006 from the Education Commission of the States. Web site: *http://www.ecs.org/*.

Elbaz, F. (1993). *Teacher thinking: A study of practical knowledge.* New York: Nichols Publishing.

Engelhard Jr., G. and Sullivan, R. (2007). Re-conceptualizing validity within the context of a new measure of mathematical knowledge for teaching. *Measurement: Interdisciplinary Research & Perspectives. 5*(2), 142-156.

Entwisle, D. and Alexander, K. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review.* Vol. 57(February): 72-84.

Fennema, E. and Loef Frank, M. (1992). Teachers' knowledge and its impact. In Grouws, D.A. (Ed.) *Handbook of Research on Mathematics Teaching and Learning.* Arlington, VA: NCTM.

Ginsburg, H.P. and Russell, R.L. (1981). Social class and racial influences on early mathematical thinking. *Monographs of the Society for Research in Child Development.* 46(6): 193.

Goldhaber, D.D., and Brewer, D.J. (1997). Evaluating the effect of teacher degree level on educational performance. *Developments in School Finance*, 197-210.

Goldhaber, D.D., and Brewer, D.J. (2000). Does teacher certification matter? High school certification status and student achievement. *Educational Evaluation and Policy Analysis*, *22*, 129-146.

Green, P.J., Dugoni, B.L., Ingels, S.J., and Camburn, E. (1995). *A profile of the American high school senior in 1992.* Washington, DC: U.S. Department of Education.

Gutierrez, R. (2008). A 'gap gazing' fetish in mathematics education? Problematizing research on the achievement gap (Research commentary). *Journal for Research in Mathematics Education.* 39(4): 357-364.

Haycock, K. (1998). *Good Teaching Matters: How Well-Qualified Teachers Can Close the Achievement Gap.* Washington, D.C.: The Education Trust.

Harbison, R.W., and Hanushek, E.A. (1992). *Educational performance of the poor: Lessons from rural northeast Brazil.* (pp. 81-177). Washington, D.C.:World Bank.

Harris, D.N. and Sass, T.R. (2007). *Teacher training, teacher quality and student achievement.* (National Center for Analysis of Longitudinal Data in Education Research: working paper #3). Washington, DC: Urban Institute.

Hedges, L. and Nowell, A. (1999). Changes in the Black-White gap in achivement test scores. *Sociology of Education.* Vol. 72 (April): 111-135.

Henningsen, M. and Stein, M.K. (1997). Mathematical tasks and student cognition: Classroom based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524-549.

Hierbert, J.C., Carpenter, T.P., Fennema, E., Fuson, K.C., Human, P.G., Murray, H.G., Olivier, A.I., and Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher. 25*(4), 12-21.

Hill, H. C., Rowan, B., and Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal* 42(2): 371-406.

Hill, H. C., Schilling, S.G., and Ball, D. L. (2004). Developing measures of teachers' Mathematics knowledge for teaching. *The Elementary School Journal.* 105(1): 11-30.

Institute of Educational Sciences (2002). *What Works Clearinghouse*. Retrieved April 16, 2009 from the Institute of Educational Sciences. Web site: http://ies.ed.gov/ncee/wwc/.

Irvine, J. (1990). *Black students and school failure.* Westport, CT: Greenwood.

James, W. (1962). *Talks to teachers on psychology and to students on some of life's ideals*. Mineola, NY. Dover Publications, Inc.

Jencks, C. (1992). *Rethinking social policy: Race, poverty, and the underclass.* Cambridge, MA: Harvard University Press.

King Rice, J. (2003). The incidence and impact of teacher professional development: Implications for education productivity. In M.L. Plecki, & D.H. Monk (Eds.), *School finance and teacher quality: Exploring the connections* (pp. 111-136). Larchmont, NY: Eye on Education, Inc.

Ladson-Billings, G. (1990). Like lightening in a bottle: Attempting to capture the pedagogical excellence of successful teachers of Black students. *Qualitative Studies in Education.* 3: 335-344.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal.* 32:465-491.

Ladson-Billings, G. and Tate, W. F. (1995). Toward a critical race theory of education. *Teachers College Record.* 97: 47-68.

Lubienski, S. (2008). On ‘gap gazing’ in mathematics education: The need for Gaps Analyses (Research commentary). *Journal for Research in Mathematics Education.* 39(4): 350-356.

Ma, L. (1999). *Knowing and Teaching Elementary Mathematics*. Mahwah, New Jersey, Lawrence Earlbaum Associates.

Mandeville, G.K., and Liu, Q.D. (1997). The effect of teacher certification and task level on mathematics achievement. *Teaching and Teacher Education, 13*, 397-407.

McCaffrey, D., Lockwood, J. R., Koretz, D., and Hamilton, L. (2003). *Evaluating Value Added Models for Teacher Accountability.* Santa Monica, CA: RAND.

Monk, D.H. and King, J.A. (1994). Multilevel teacher resource effects on pupil performance in secondary mathematics and science: The case of teacher subject matter preparation. In R.G. Ehrenberg (Ed.), *Choices and consequences:*

Contemporary policy issues in education (pp. 29-58). Ithaca, NY: ILR Press.

Mullens, J.E., Murnane, R.J., and Willett, J. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review*, 40, 139-157.

Murnane, R. and Levy, R.J. (1996). *Teaching the new basic skills.* New York: The Free Press.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: NCTM.

National Council of Teachers of Mathematics (2001). *Principles and standards for school mathematics.* Reston, VA: NCTM.

National Mathematics Advisory Panel. (2008). Report of the Task Group on Teachers and Teachers Education. *Foundations for Success: The Final Report of the National Mathematics Advisory Panel.* Retrieved November 12, 2008 from the Department of Education. Web site: www.ed.gov/MathPanel.

Nye, B., Konstantopoulos, S., and Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 337-257.

Ogbu, J.U. (1994). Racial stratification and education and education in the United States: Why inequality persists. *Teachers College Record.* 96, 264-298.

Rasinski, K.A., Ingels, S.J., Rock, D.A., and Pollack, J.M. (1993). *America's high school sophomores: A ten year comparison.* (NCES 93-087). Washington, DC: National Center of Education Statistics.

Raudenbush, S. W. (2004). Schooling, statistics, and poverty: Can we measure school improvement? William H. Angoff Memorial Lecture Series. Princeton, NJ.

Raudenbush, S. W., Bryk, Cheong, and Congdon. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL, Scientific Software International.

Rogosa, D. (1995). Myths and methods: Myths about longitudinal research plus supplemental questions. In, Gottman, J.M. (Ed.), *The analysis of change.* Mahwah, NJ: Lawrence Earlbaum Associates.

Ross, J. A., D. McDougall, and Hogobaum-Grey (2002). A survey measuring elementary teacher' implementation of standards-based mathematics teaching. *Journal of Research in Mathematics Education.* 34(4): 344-363.

Rowan, B., Chiang, F., and Miller, R.J. (1997). Using research on employees' Performance to study the effects of teachers on students' achievement. *Sociology of Education*, 70, 256-284.

Rowan, B. (1999). The task characteristics of teaching: Implications for the organizational design of schools. In R. Bernhardt, et al., *Curriculum leadership for the 21st century.* Cresskill, NY: Hampton Press.

Rowan, B., Correnti, R., and Miller, R.J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the *prospects* study of elementary schools. *Teachers College Record, 104,* 1525-1567.

Sanders, W. and Rivers, J. (1996) *Cumulative and residual effects of teachers on future student academic achievement.* Knoxville, TN: University of Tennessee Value Added Research and Assessment Center.

Saxe, G.B., Gearhart, M., and Seltzer, M. (1999). Relations between classroom practices and student learning in the domain of fractions. *Cognition and Instruction, 17*(1), 1-24.

Schaeffer, R. (chair), Working group on statistics in mathematics education research. (2007). *Using statistics effectively in mathematics education research: A report from a series of workshops organized by the American Statistical Association with funding from the National Science Foundation*. The American Statistical Association.

Secada, W.G. (1992). Race, ethnicity, social class, language, and achievement in mathematics. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 623-660). New York: MacMillan.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*. 15(2): 4-14.

Sullivan, R. K. (2009). *The effects of teacher knowledge and teaching practices on student learning in mathematics: A multilevel exploration.* Unpublished empirical paper, Emory University, Atlanta.

Tate, W. F. (1997). Race-Ethnicity, SES, gender, and language proficiency trends in mathematics achievement: An update. *Journal for Research in Mathematics Education*. 28(6): 652-679.

Figure 1. Conceptual map of mathematical knowledge for teaching (Engelhard & Sullivan, 2007).

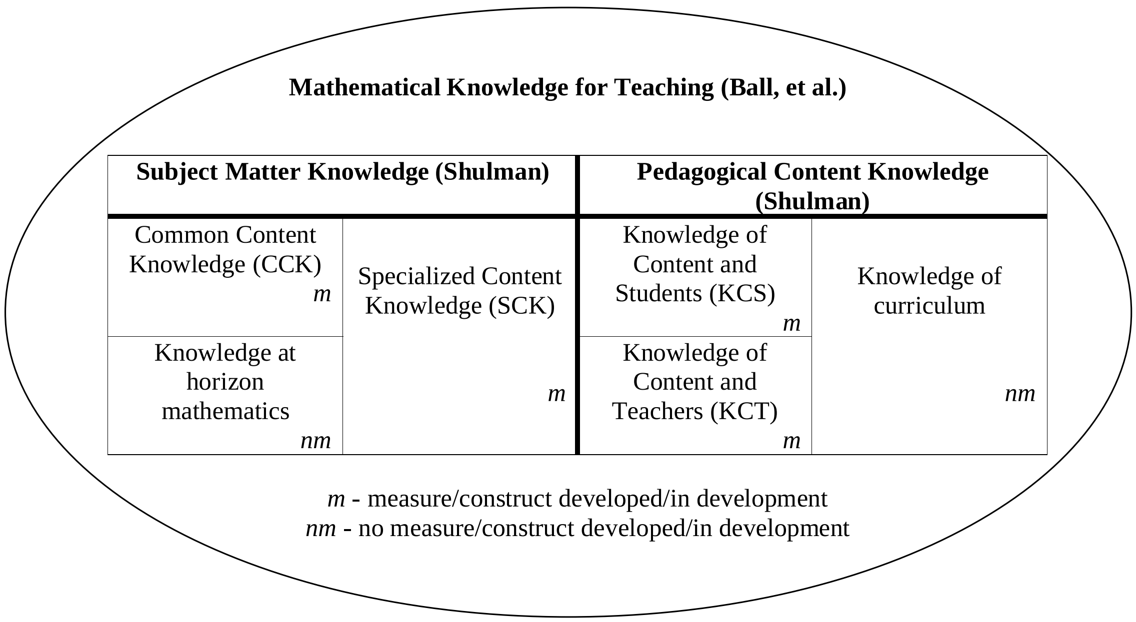
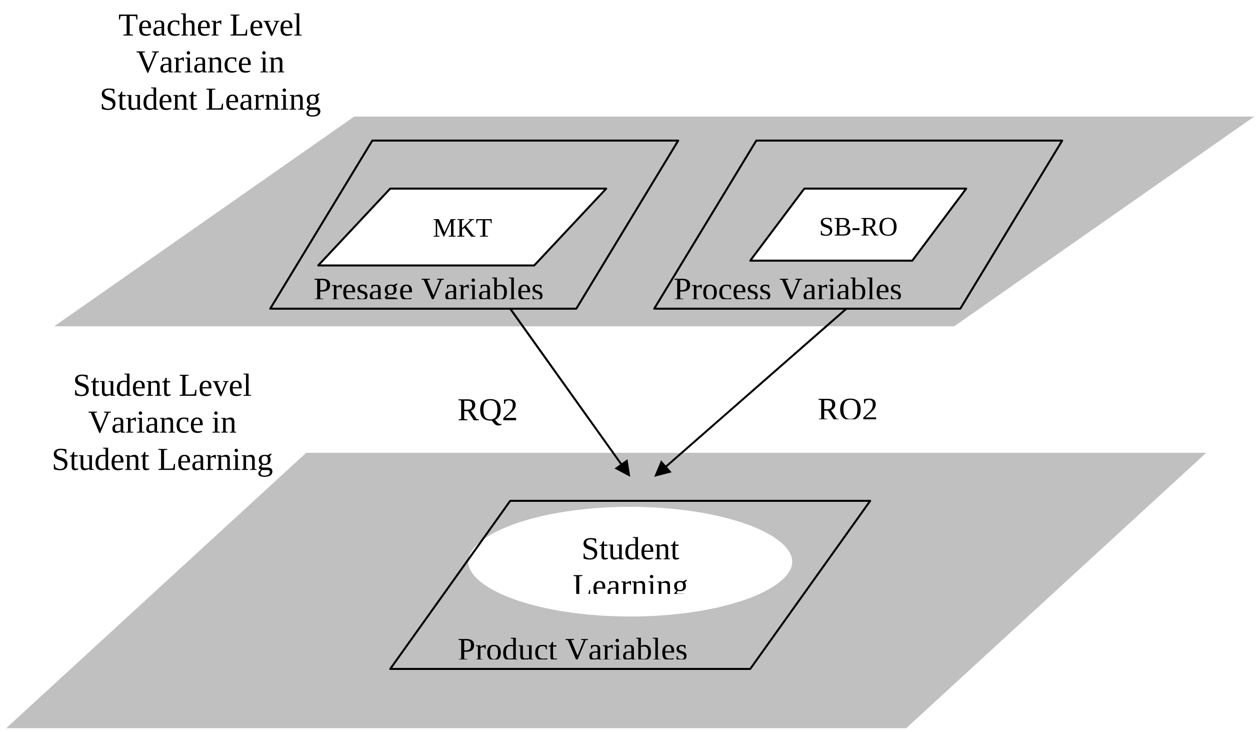


Figure 2. Two-level model of teacher effects on student learning in mathematics: Presage, process, and product variables



MKT – Teacher knowledge (mathematical knowledge for teaching)

SB-RO – Standards-Based, Reform Oriented Teaching Practices

RQ2 – Research question 2

Table 1. Literature Reviewed: Variable level, type, and name and related literature.

Level	Variable Type	Variable Name	Related Literature
Teacher	Presage	Knowledge	Monk & King (1994)
			Mullens, Murnane, & Willett (1996)
			Rowan, Chiang & Miller (1997)
			Ma (1999)
			Goldhaber & Brewer (2000)
			Harbison & Hauchek (2002)
			Rowan, Correnti & Miller (2002)
			King Rice (2003)
			Hill, Rowan & Ball (2005)
			Harris & Sass (2007)
Teacher	Process	Teaching	Henningsen & Stein (1997)
		Practices	Cohen & Ball (1999)
			Saxe, Gearhart, & Seltzer (1999)
			NCTM Principles and Standards (1989, 2000)
			Ross, McDougall, & Hogabaum-Gray (2002)
			Hierbert (2003)

Table 1. Literature Reviewed: Variable level, type, and name and related literature (cont.).

Level	Variable Type	Variable Name	Related Literature
Student	Context	Race/ Ethnicity & Class Status	Ginsburg & Rusell (1981) Entwisle &Alexander (1992) Secada (1992) Green, et al. (1995) Tate (1997) Hedges & Nowell (1999) Gutierrez (2008) Lubienski (2008)
Student	Product	Learning	Rogosa (1995) Cohen & Ball (1999) Rowan, Correnti, & Miller (2002) Nye, Konstantopolous & Hedges (2004) Raudenbush (2004) Scheaffer, et al (2007)

Table 2. Nine dimensions of standards-based, reform-oriented instructional practices (Ross, McDougall, & Hogabaum-Gray, 2002).

Dimension	Description
Program Scope	Characterized by a broad view of mathematics with all students having access to all types of mathematics.
Student Tasks	Characterized by complex, open-ended problems embedded in a real-life context.
Discovery	Characterized by a focus on the discovery of mathematical ideas led by student discovery.
Teacher's Role	Characterized as that of a co-learner and creator of a mathematical community within the classroom.
Manipulatives & Tools	Characterized by ready access to and use of manipulatives and other mathematical tools such as calculators.
Student – Student Interaction	Characterized by the promotion of student – to – student interaction.
Student Assessment	Characterized as authentic and embedded in everyday instruction.
Teachers' Perception of Math as a Discipline	Characterized by an understanding of mathematics as a dynamic subject as opposed to a fixed body of knowledge.
Student Confidence	Characterized by a teacher's intent to boost mathematical confidence in students.

Table 3. Measures.

Variable	Measure	Characteristics	Sample-Specific Reliability Cronbach's Alpha	Sample-Specific Reliability Rasch
Mathematical knowledge for teaching	Elementary Number Concepts and Operations, Specialized Content Knowledge, Learning Mathematics for Teaching	24 multiple choice items, standardized scale score ranging from -3 to 3 wherein a score of 0 is representative of average teacher knowledge	.78	.80
Teaching practices	Implementation of Standards-Based Mathematics Teaching	20 items using a 6-point Likert scale ranging from strongly disagree (1) to strongly agree (6), score ranges from 20 to 120 wherein a higher score is indicative of higher use of said practices	.78	.93
Student learning	Spring, grade two state criterion reference test	Multiple choice	.91	NA*
	Spring, grade three state criterion reference test	Multiple choice; required for promotion	.92	NA*

*Not Available

Table 4. Detailed description of school sample.

School Code	Sample Teacher N	Sample Student N	Proportion African American	Proportion Free or Reduced Lunch	Mean CRT08 Scale Score	Sample Mean MKT*	SD	Sample Mean SB-RO** Practices	SD
1	4	53	0.99	0.99	818.44	-1.22	0.70	92.50	10.34
2	4	65	0.99	0.91	825.28	-1.02	0.64	95.75	6.40
3	3	38	0.43	0.43	831.10	-0.59	1.39	85.75	7.81
4	3	36	0.64	0.98	804.64	-1.45	1.18	90.33	4.04
5	3	40	1.00	0.50	832.74	-0.92	1.69	95.67	4.73
6	1	21	0.94	0.91	823.25	-1.96	-	74.00	-
7	2	27	0.98	0.95	858.29	-1.57	0.56	86.00	5.66
8	1	54	0.99	0.92	804.24	-1.57	-	92.00	-
9	1	12	0.82	0.95	826.04	-0.77	-	71.00	-
10	1	14	0.99	0.96	789.00	-0.36	-	111.00	-
11	1	16	0.99	0.89	803.14	-1.37	-	89.00	-
12	2	33	0.99	0.94	842.09	.24	1.99	103.00	8.49

*Mathematical Knowledge for Teaching; **Standards-Based, Reform-Oriented

School Code	Sample Teacher N	Sample Student N	Proportion African American	Proportion Free or Reduced Lunch	Mean CRT08 Scale Score	Sample Mean MKT	SD	Sample Mean SB-RO Practices	SD
13	3	37	0.97	0.96	796.59	-0.86	1.23	79.00	13.53
14	3	45	0.13	0.13	845.07	-1.64	0.65	86.00	15.72
15	1	14	0.80	1.00	783.81	-0.36	-	97.00	-
16	1	9	0.99	0.88	817.47	-1.17	-	85.00	-
17	1	17	0.23	0.13	850.10	-1.57	-	94.00	-

*Mathematical Knowledge for Teaching; **Standards-Based, Reform-Oriented

Table 5. Detailed description of teacher sample.

Teacher	School	Student N	Proportion African American	Proportion Free or Reduced Lunch	MKT	SBRO	CRT08 mean	SD	CRT09 mean	SD
1	1	14	1	.79	-1.96	89	826.93	22.66	812.57	36.63
2	1	13	1	1	-1.57	97	813.15	18.54	796.38	32.56
3	1	13	1	.85	-0.97	80	814.31	24.21	811.77	26.27
4	1	13	1	1	-0.36	104	811.46	16.08	811.92	25.74
5	2	15	1	.80	-0.77	87	810.60	22.08	788.13	33.05
6	2	17	.94	.82	-0.57	95	811.82	32.97	796.65	36.53
7	2	17	1	.77	-1.96	101	827.00	21.94	824.47	38.24
8	2	16	.94	.75	-0.77	100	827.12	35.33	819.44	35.17
9	3	14	.21	.36	0.07	97	837.43	23.71	841.07	8.73
10	3	11	.27	.82	-0.36	80	821.09	22.22	803.09	30.92
11	3	13	.08	.31	0.54	85	843.15	30.14	858.15	48.73
12	3	0	*	*	-2.59	81	*	*	*	*
13	4	11	.54	.92	-0.77	95	808.23	26.89	856.46	44.97

Teacher	School	Student N	Proportion AA	Proportion FRL	MKT	SB-RO	CRT08 Mean	SD	CRT09 Mean	SD
14	4	11	.46	.91	-0.77	88	819.27	16.44	825.09	48.68
15	4	12	.67	.92	-2.81	88	826.58	18.01	873.50	19.73
16	5	13	1	.54	0.79	92	856.54	35.49	897.85	75.07
17	5	15	1	.73	-2.59	94	862.00	30.57	874.80	37.20
18	5	12	1	.50	-0.97	101	855.08	27.15	923.67	64.10
19	6	21	1	.91	-1.96	74	827.48	31.85	805.76	35.18
20	7	14	1	1	-1.17	90	829.43	33.47	826.86	32.05
21	7	13	1	.92	-1.96	82	833.23	19.31	847.69	37.31
22	8	54	1	.96	-1.57	92	824.72	23.59	827.19	31.26
23	9	12	.92	.92	-0.77	71	816.00	30.92	824.33	57.09
24	10	14	1	.86	-0.36	111	848.43	46.15	820.79	41.07
25	11	16	1	.88	-1.37	89	815.88	23.13	837.81	48.97
26	12	17	1	.88	1.65	109	837.82	35.58	836.53	62.78

Teacher	School	Student N	Proportion AA	Proportion FRL	MKT	SB-RO	CRT08 Mean	SD	CRT09 Mean	SD
27	12	16	1	.94	-1.17	97	811.75	32.87	800.31	35.29
28	13	13	1	1	0.54	80	813.31	22.59	815.00	38.51
29	13	13	1	.77	-1.37	92	823.69	23.30	825.85	39.11
30	13	11	.91	.91	-1.76	65	807.45	46.06	802.91	40.57
31	14	13	.08	.23	-2.38	75	855.31	26.39	861.85	33.29
32	14	0	*	*	*	*	*	*	*	*
33	14	14	0	.07	-1.17	79	846.36	25.62	874.21	30.72
34	15	14	.71	1	-0.36	97	812.14	22.83	804.57	40.81
35	16	9	1	.89	-1.17	85	829.22	17.54	824.11	38.18
36	14	18	.06	.11	-1.37	104	855.56	27.67	868.44	31.16
37	17	17	.29	.12	-1.57	94	844.94	33.31	872.18	54.85

Note:
MKT: Score ranging from -3 to 3 where 0 can be interpreted as the MKT of an average teacher based on a nationally representative sample; SB-RO: raw score from 20 to 120
*Missing Data

Table 6. Sample proportions, means and standard deviations.

Measure	Prop	Mean	SD	*n*
Student				
2008 State Criterion Referenced Test (2008 CRT)		828.73	31.10	531
2009 State Criterion Referenced Test (2009 CRT)		833.16	49.31	531
Proportion African American (AA)	.79			531
Proportion in Free or Reduced Lunch (FRL)	.71			531
Teacher				
Mathematical knowledge for teaching (MKT)		-1.00	.97	35
Teacher practices (SB-RO)		90.26	10.63	35

Note: A score of 800 on the CRT is indicative of a student meeting the grade level expectation. MKT is interpreted such that a mean of 0 is indicative of the mathematical knowledge for teaching of an average teacher based on a nationally representative sample. SB-RO is a raw score ranging from 20 to 120, so a mean of 90 is indicative of a teacher reporting moderately high use of standards-based, reform oriented teaching practices.

Table 7. Rasch variable map illustrating the level of mathematical knowledge for teaching.

```
+-----------------------------------
|Measr   |+Teacher             |
|--------+---------------------+
|  4     +                     +
|        | 26                  |
|        |                     |
|        |                     |
|        |                     |
|        |                     |
|        |                     |
|  3     +                     +
|        |                     |
|        |                     |
|        |                     |
|        |                     |
|        |                     |
|        |                     |
|  2     + 16                  +
|        |                     |
|        |                     |
|        | 11 28               |
|        |                     |
|        |                     |
|        | 9                   |
|  1     +                     +
|        |                     |
|        |                     |
|        | 4  10 24 34         |
|        | 6           | 3     |
|        | 5  8  13 14 23      |
|        |                     |
*   0    * 3  18               *
|        | 20 27 33 35         |
|        |                     |
|        | 25 29 36            |
|        | 2  22               |
|        |                     |
|        | 30                  |
|  -1    + 1  7  19 21         +
|        | 31                  |
|        |                     |
|        |                     |
|        |                     |
|        |                     |
|        | 12 17               |
|  -2    +                     +
|        |                     |
|        | 15                  |
|        |                     |
|        |                     |
|        |                     |
|        |                     |
|  -3    +                     +
|--------+---------------------+
|Measr   |+Teacher             |
+-----------------------------------
```

Average Teacher's
Mathematical
Knowledge for Teaching

Figure 3. Distribution of student CRT08 scale scores by teacher.

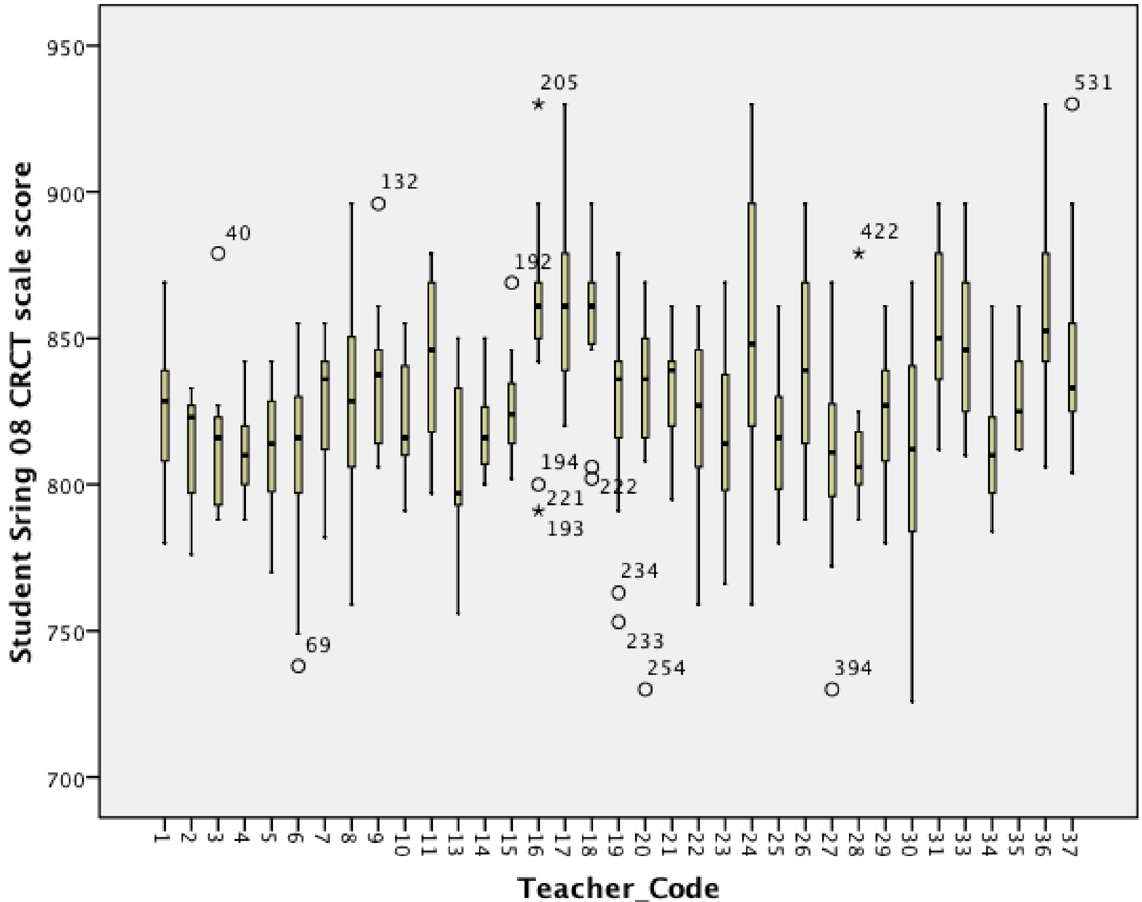


Figure 4. Distribution of student CRT09 scale scores by teacher.

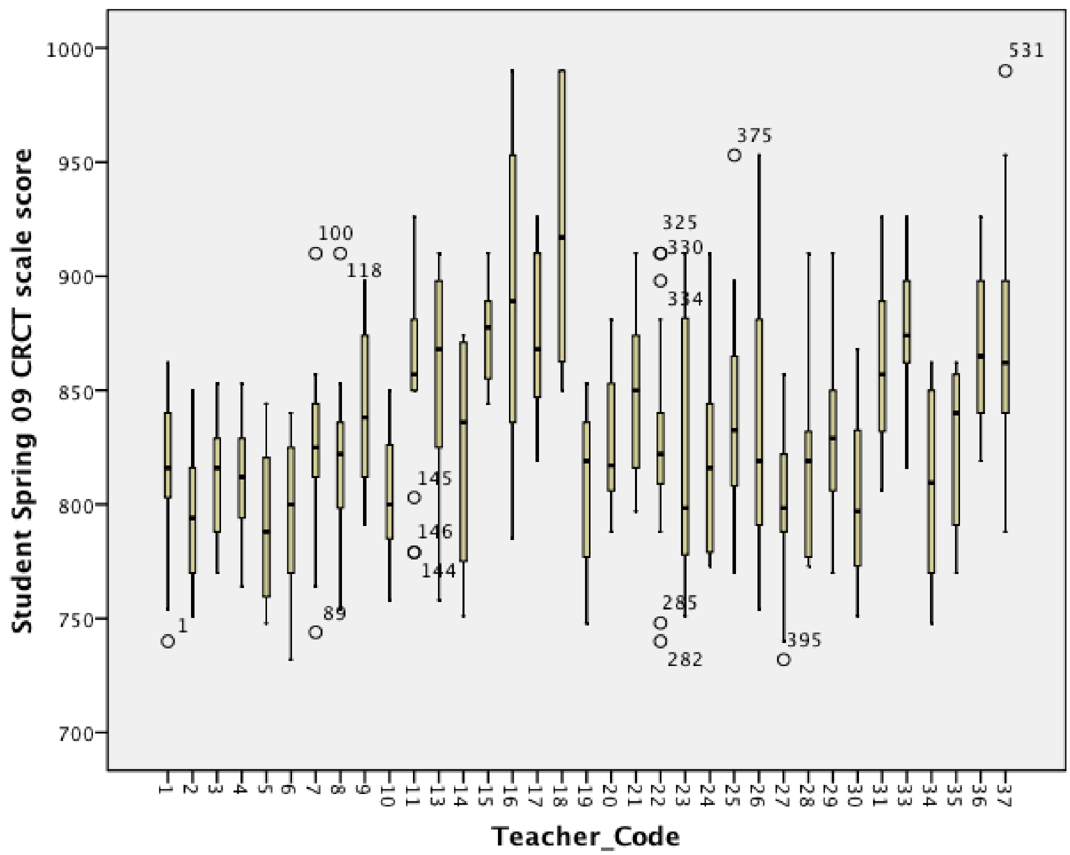


Figure 5. Teacher mathematical knowledge for teaching and use of standards-based, reform-oriented teaching practices scatterplot with line of best fit.

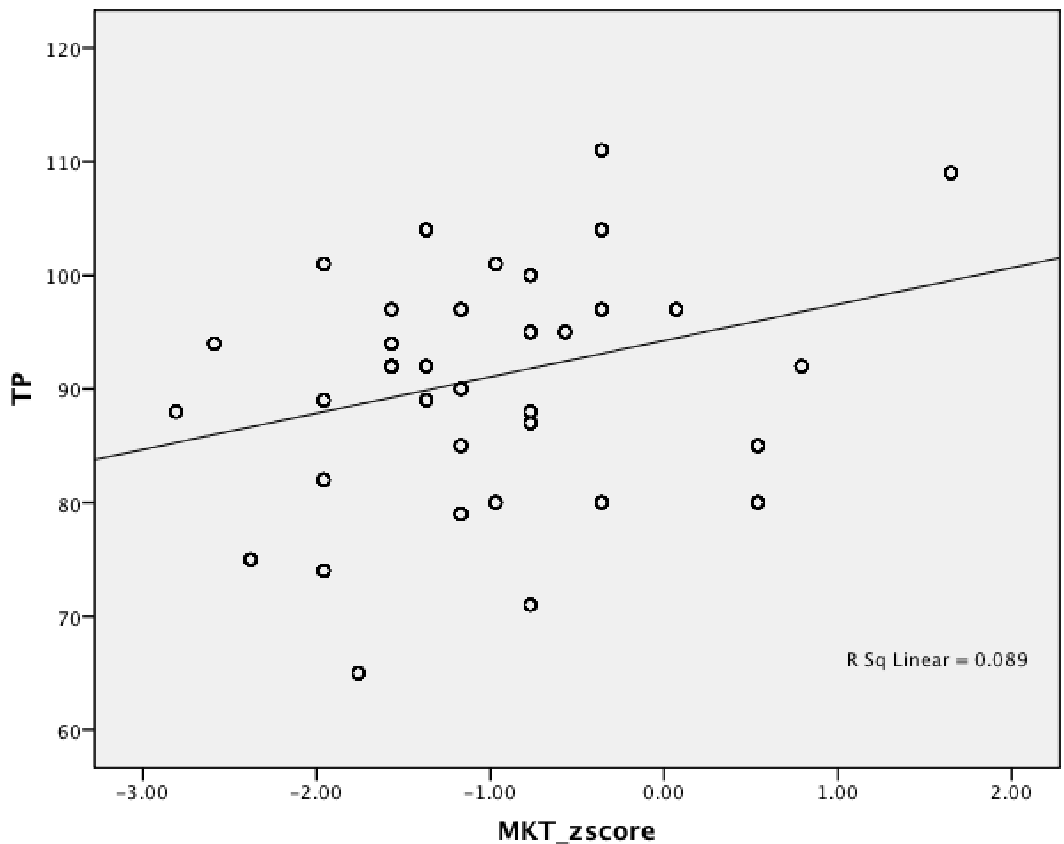


Table 8. Correlation coefficients for research question one.

	Proportion of African-American Students	Proportion of Students eligible for Free or Reduced Lunch	Mean prior year CRCT (Prior Learning)
MKT, teacher	-.03	.01	-.12
SB-RO, teacher	.14	-.02	.23
MKT, sample mean within school	.12	.08	-.07
SB-RO, sample mean within school	.18	.07	-.01

Note: Student-level variables (race, class status, and prior learning) were aggregated at the classroom level. For the school level, data were gathered from the school-based report card available on the GA Department of Education's website.

Table 9. Variance components of 2008 CRT and 2009 CRT outcomes.

Component	2008 CRT		2009 CRT	
Students (σ^2)	779.39	79%	1652.83	65%
Teachers (τ)	205.63	21%	875.15	35%

Note: Variance components are calculated using values from the HLM output. σ^2 and τ report the variance at the student and teacher-levels respectively. The percent variance is then calculated by dividing the variance at each level by the sum of the variance at both levels.

Table 10. Two-level analyses of teacher effects on student adjusted 2009 CRT score (each teacher-level variable run independently, resulting in two models).

Model	Measure	Coefficient	SE	df
1	Intercept (γ_{00})	832.72*	5.39	33
	MKT (γ_{01})	-1.25	3.89	33
2	Intercept (γ_{00})	833.97*	3.59	33
	SB-RO (γ_{01})	0.06	0.32	33

* $p < .05$

Table 11. Two-level analysis of the effects of MKT and SB-RO on student adjusted 2009 CRT score.

Measure	Model Coefficients	SE	*df*
Intercept (γ_{00})	832.37*	5.01	32
Mathematical knowledge for teaching (MKT) (γ_{01})	-1.59	3.76	32
Teaching practices (SB-RO) (γ_{02})	0.11	0.33	32

*$p < .05$

Table 12. Schools flagged for potential testing irregularities in the Spring 2009 administration of the CRT.

School Code	GOSA*	BRC**
1	Moderate	3
2	Severe	3
3	Not Flagged	Not Flagged
4	Severe	2
5	Severe	2
6	Moderate	3
7	Severe	3
8	Severe	2
9	Severe	2
10	Severe	2
11	Moderate	3
12	Moderate	3
13	Severe	3
14	Not Flagged	Not Flagged
15	Severe	3
16	Minimum	Not Flagged
17	Not Flagged	Not Flagged

*Governor's Office of Student Achievement (GOSA) utilized the following categories to classify schools with abnormal erasure patterns: Severe – 25% or more classrooms flagged; Moderate – 11 to 24% classrooms flagged; Minimum – 6 to 10% of classrooms flagged.

**Blue Ribbon Commission (BRC) utilized the following categories to classify schools based on erasure analysis and interviews: 1 – strong circumstantial evidence ot testing irregularities; 2 – anomalous data indicative of irregularities; 3 – little to no indication of irregularities.

Appendix A

Learning Mathematics for Teaching Released Items

1. Ms. Dominguez was working with a new textbook and she noticed that it gave more attention to the number 0 than her old book. She came across a page that asked students to determine if a few statements about 0 were true or false. Intrigued, she showed them to her sister who is also a teacher, and asked her what she thought.

Which statement(s) should the sisters select as being true? (Mark YES, NO, or I'M NOT SURE for each item below.)

	Yes	No	I'm not sure
a) 0 is an even number.	1	2	3
b) 0 is not really a number. It is a placeholder in writing big numbers.	1	2	3
c) The number 8 can be written as 008.	1	2	3

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 in the following ways:

Student A	Student B	Student C
35 x 25 125 +75 875	35 x25 175 +700 875	35 x25 25 150 100 +600 875

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 whole numbers	I'm not sure
a) Method A	1	2	3
b) Method B	1	2	3
c) Method C	1	2	3

3. Ms. Harris was working with her class on divisibility rules. She told her class that a number is divisible by 4 if and only if the last two digits of the number are divisible by 4. One of her students asked her why the rule for 4 worked. She asked the other students if they could come up with a reason, and several possible reasons were proposed. Which of the following statements comes closest to explaining the reason for the divisibility rule for 4? (Mark ONE answer.)

a) Four is an even number, and odd numbers are not divisible by even numbers.

b) The number 100 is divisible by 4 (and also 1000, 10,000, etc.).

c) Every other even number is divisible by 4, for example, 24 and 28 but not 26.

d) It only works when the sum of the last two digits is an even number.

4. Ms. Chambreaux's students are working on the following problem:

Is 371 a prime number?

As she walks around the room looking at their papers, she sees many different ways to solve this problem. Which solution method is correct? (Mark ONE answer.)

a) Check to see whether 371 is divisible by 2, 3, 4, 5, 6, 7, 8, or 9.

b) Break 371 into 3 and 71; they are both prime, so 371 must also be prime.

c) Check to see whether 371 is divisible by any prime number less than 20.

d) Break 371 into 37 and 1; they are both prime, so 371 must also be prime.

Appendix B

Teaching Practices Instrument

Teaching Practices

Using the 1 to 6 point scale, indicate the extent to which you disagree or agree with each statement by circling the appropriate number.

1=Strongly Disagree 2=Disagree 3=Mildly Disagree 4=Mildly Agree 5=Agree 6=Strongly Agree

1. I like to use math problems that can be solved in many different ways.

1 2 3 4 5 6

2. I regularly have my students work through real-life math problems that are of interest to them.

1 2 3 4 5 6

3. When two students solve the same math problem correctly using two different strategies I have them share the steps they went through with the class.

1 2 3 4 5 6

4. I tend to integrate multiple topics of mathematics within a single unit (i.e. geometric and algebraic concepts together).

1 2 3 4 5 6

5. I often learn from my students during math class because my students come up with ingenious ways of solving problems that I have never thought of.

1 2 3 4 5 6

6. It is not very productive for students to work together during math class.

1 2 3 4 5 6

7. Every student in my class should feel that mathematics is something he/she can do.

1 2 3 4 5 6

8. I integrate math assessment into most math activities.

1 2 3 4 5 6

9. In my classes, students learn math best when they can work together to discover mathematical ideas.

1 2 3 4 5 6

10. I encourage students to use manipulatives or technology to explain their mathematical ideas to other students.

1 2 3 4 5 6

11. When students are working on math problems, I put more emphasis on getting the correct answer rather than on the process followed.

1 2 3 4 5 6

12.Creating rubrics for math is a worthwhile assessment strategy.

1 2 3 4 5 6

13.In high school it is just as important for students to learn geometry and statistics as it is to learn algebra.

1 2 3 4 5 6

14.I don't necessarily answer students' math questions but rather let them puzzle things out for themselves.

1 2 3 4 5 6

15.A lot of things in math must simply be accepted as true and remembered.

1 2 3 4 5 6

16. I like my students to master basic mathematical procedures before they tackle complex problems.

1 2 3 4 5 6

17.I teach students how to explain and defend their mathematical ideas.

1 2 3 4 5 6

18.Using computers to solve math problems distracts students from learning basic algebraic and procedural skills.

1 2 3 4 5 6

19. If students use calculators they won't master the basic algebraic and procedural skills they need to know.

1 2 3 4 5 6

20. You have to study math for a long time before you see how useful it is.

1 2 3 4 5 6

Appendix C

School District Study Approval Letter

ATLANTA PUBLIC SCHOOLS

Our Focus...Student Success

August 12, 2008

RESEARCH, PLANNING, AND ACCOUNTABILITY
130 TRINITY AVENUE, S.W.
ATLANTA, GEORGIA 30303-3624
404-802-2780

Ms. Rubye Sullivan
309 Elmira Place
Atlanta, GA 30307

Dear Ms. Sullivan:

Your request to conduct research within the Atlanta Public Schools (APS) was reviewed by the Research Screening Committee in accordance with the guidelines. The research study entitled "**Mathematical Knowledge for Teaching, Instructional Practice, and Students' Performance on a Standardized Achievement Test**" was approved under the following conditions:

1. Your study is APS district-wide. You must obtain the approval of the principals prior to beginning your research study. The principal has the final approval on whether research studies are conducted in the school. If a principal of a school does not approve of your study or does not believe that it is in the best interest of the school to participate, you may select a comparable APS school as replacement with the principal's approval.
2. Your study involves 1,500 third grade students and you are requesting student data, including test results of Grade 3 Benchmark assessment for Fall and Spring. **Student data identifiable by individual students cannot be released. You may be provided with student data only in blind, aggregate, or encrypted format.**
3. Your study involves 100 randomly selected third grade teachers in the district and you intend to survey those teachers.
4. **Students, teachers, principals and other APS employees can participate in or assist with your research study only on a voluntary basis.**
5. The confidentiality of students, teachers, other APS staff members, the schools, and the school system must be ensured. Pseudonyms for people and the schools, as well as references to APS as "a large urban school system," are required in the title and text of your final report before publication or presentation outside of APS.
6. The data collection phase of your research study must be completed by the end of the 2008 calendar year. Longitudinal studies cannot be approved, except on an annual basis. At the end of the 2008 calendar year, you may formally request a continuation for another year. Approval for one year does not guarantee approval for subsequent years.
7. If changes are made in the research design or in the instruments used, you must notify the Department of Research, Planning, and Accountability prior to beginning your study.

This letter serves as official notification of the approval for your proposed research study, pending the above conditions. Remember that a copy of the results of your completed study must be submitted to the Department of Research, Planning, and Accountability. Please contact me at (404) 802-2710 or kmaddula@atlanta.k12.ga.us if I can be of further assistance.

Sincerely,

Kavi Maddula
Research Associate

KM:dd - #12

xc: Mr. Lester McKee
Executive Directors (SRT1-4)
Elementary Schools Principals

Appendix D

Informed Participant Consent

Emory University Division of Educational Studies
Participant Consent

Title: Mathematical Knowledge for Teaching, Instructional Practice, and Students' Performance on a Standardized Achievement Test

Principal Investigator: Rubye Sullivan

Sponsor: None

Introduction and Purpose:
You are being asked to take part in a research study. The purpose of the study is to explore the relationship among mathematical knowledge for teaching, the instructional practices of teachers, and student learning in mathematics

Procedures:
You and the other participants will complete a survey asking questions related to content knowledge and your instructional practices. In addition, questions will be asked related to your experience and training as a mathematics teacher. Student data will also be used in this study in an attempt to better understand the effect of content knowledge and instructional practices on student performance in mathematics. The survey should take approximately one hour to complete.

Risks, Discomforts, and Inconveniences:
Although highly unlikely, a breech of confidentiality is possible. Only one form will link participants to their identification number and then to their survey responses. This form will be kept in a locked file cabinet on Emory University's campus.

Benefits:
Although there are no direct benefits, this work will shed light on teacher knowledge and practices that benefit students directly in the area of mathematics. This level of insight may inform future professional development and teacher preparation programs.

Confidentiality:
Your participation is completely voluntary, and will be kept strictly confidential. However, agencies and Emory departments that make the rules and policies about how research is done have the right to review study records. Your name will not appear on any reports or publications that may be written. Although no private matters of any kind will be discussed, the surveys will be kept in a secure office at Emory University. Your employment status will not be affected due to participation in this study.

Contact Persons:
If you have any questions about this study you may call Rubye Sullivan, the Principal Investigator:

Rubye Sullivan 404.223.5849
Rsulli4@emory.edu

If you have any questions about your rights as a participant in this research study, you may contact Emory University's Institutional Review Board:

Call toll-free at 1-877-503-9797 or (404) 712-0720; email irb@emory.edu; or write to the office at 1599 Clifton Road, Atlanta GA 30322.

It's Your Choice:
You are free to choose whether or not you want to take part in this study. You can change your mind and stop at any time without penalty. This decision will not adversely affect your relationship with the researchers or Emory. It will not affect any benefits you may receive outside of the research. It's your choice.

Withdrawal:
The lead researcher and/or sponsor may withdraw you from the study if they decide that it is in your best interest.

If you are willing to volunteer for this research, please sign below. You do not give up any rights by signing this form. You will be given a copy of this form for your records.

____ __________ _________
Participant's name /Signature Date Time

____ __________ _________
Person Obtaining Consent Date Time