University of Nevada Reno

The Relationship of Teacher Attitude and Self-Reported Behavior Associated with Measures of Academic Progress Assessment and Student Academic Growth

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Educational Leadership

by

Jennifer Ward DeJoseph

Dr. Bill Thornton, Chair/Dissertation Advisor

May, 2012

University of Nevada Reno

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

JENNIFER ANN WARD DEJOSEPH

entitled

The Relationship of Teacher Attitude And Self-Reported Behavior Associated With Measures of Academic Progress Assessment And Student Academic Growth

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Bill Thornton, Ph. D., Advisor

Gus Hill, Ph. D., Committee Member

Janet Usinger, Ph. D., Committee Member

John Cannon, Ph. D., Committee Member

Michele Pelter, Ph. D., Graduate School Representative

Marsha H. Read, Ph. D., Dean, Graduate School

May, 2012

Abstract

The overall intent of No Child Left Behind is to ensure the success of all students; however, within many schools, the components of the law have become the focus of leadership. For example, assessment literacy, measures of achievement, teacher behaviors and attitudes, and student achievement have become high priority issues. The results of student achievement data have been utilized to plan, implement, evaluate, and improve school reform. As a result of increased use of data in schools, need for assessment literacy of all stakeholders has increased. Teacher influence on student achievement and the importance of assessment literacy of teachers have been well documented; however, teacher attitude toward assessments and relationships to student achievement are not well understood nor well documented.

This study focused on relationships among student achievement, teacher attitude about assessments, and teacher application of assessment data. Selected teacher demographic variables were used to define groups including grade level taught, length of service and education level. Data was analyzed using quantitative methods.

Two major findings were identified. First, teacher self-reported behavior concerning assessments can significantly predict math achievement at the classroom level. Second, differences in attitude toward assessment existed between primary and intermediate teachers.

Dedication

I dedicate my dissertation work to my family and many friends. I am especially grateful to my husband, Kinkade, my two wonderful step-children, Jayden and Shea, and my sister, Shannon. Your support, words of encouragement, and belief in my abilities inspired me to finish this journey. You made this endeavor feel not only possible but always within my reach.

A special dedication is to my mom, Judy Heckman, who without, I would be lost. Your encouragement was only matched by your countless hours of tireless proofreading and revision sessions. I cannot imagine this happening without your support. You have always been my biggest cheerleader and I am eternally grateful to you. You have inspired me in more ways than you will ever know. Thank you for always being in my corner.

I am blessed to have such a wealth of great friends who have supported me in this undertaking with enthusiasm and encouagement. They were never at a loss of kind words and support. You are blessing to the students you teach and to the colleagues around you. Thank you to LeAnn, Linda, Terry, Lisa, Laurel, Johanna, Gayle, Christina, Judy, Carol, Marcy, Jason, and Kari. You are a rare breed and I am honored to call you colleagues as well as friends. A special thanks to Amanda Cox and Susan Squires without whom I would have lost a lot sleep, a bit of my sanity, and any humor I had left.

Acknowledgements

I wish to thank my committee members who were more than generous with their time and expertise. A special thanks to Dr. Bill Thornton, my committee chairman for his countless hours of reflecting, reading, encouraging, and most of all patience throughout the entire process. Thank you to Dr. Janet Usinger who played a role in my decision to enter into the program years before. Her professionalism and dedication to student excellence is admirable. Thank you to Dr. Gus Hill who not only provided feedback for my study but for leading hours of course discussions that led me to a better understanding along the way. Thank you to Dr. John Cannon who provided unique perspective as a teacher of teachers. I enjoyed being in his undergraduate class twenty years ago and continue to admire his enthusiasm for his teaching as well as his content. I would also like to express gratitude to Dr. Michele Pelter who accepted my request to serve on my committee as my out of college representative. Her fresh perspective and questions during this process were appreciated.

I would like to acknowledge the personnel of my school district who made this research possible. The staff of my school district is a group of learner-centered professionals who I am blessed to have as friends. A special thanks to Dr. Steven Pradere and Dr. Ricky Medina. Their technical support was only outmatched by their overwhelming words of encouragement and support. They never lost faith and for that I am truly in their debt.

A final word of acknowledgement is extended to all that enter into this profession to positively enrich the lives of students. It is what you do and how you do it that makes a difference.

--Every student--Every day--100%

Table of Contents

List of Tables

List of Figures

CHAPTER ONE: INTRODUCTION

Background of the Study

The current educational climate demands that students demonstrate academic achievement, and this is exclusively measured through standardized assessments. Medina (2010) reported that many educators do not believe that what is assessed is actually aligned with what makes their schools successful. The principals in his study felt that doing the 'right thing' for students by providing a safe environment, and for teachers by allowing them to work on their classroom instruction. He also found that many of the principal respondents had no concrete ideas on how to use assessment data to improve student learning.

According to Popham (2004), "In education, the linchpin of the accountability system is our students' test performance" (p. 82). He addressed the concept of assessment literacy and its potential impacts on student achievement. He equated teachers who do not possess the necessary assessment literacy skills as committing the ultimate sin of "professional suicide" (p. 83). Assessment literacy is the set of skills and knowledge that allow educators to make judgments about validity, reliability, and appropriate application about testing, results, and corresponding instruction.

What seems to be less understood than need for assessment literacy is teacher attitude toward assessment. Very little empirical research exists at this time on the subject. Brown (2004) completed a study in New Zealand and Queensland, Australia. He studied teacher attitude about assessments and defined attitude as the way a teacher interacts with, responds to, and understands assessment. According to this researcher, the beliefs and behaviors associated with assessment are essential and extremely important. He concluded that teachers characterized assessment in the following four distinct ways: improvement of learning, student accountability,

school accountability, and irrelevancy. Brown concluded that teacher attitude toward assessments were positive when the results were used for school accountability and when improvement of learning and student accountability were associated with formative, not summative, assessment.

In a similar study Brown, Kennedy, Fok, Chan, and Yu (2009) considered Hong Kong teacher attitudes toward assessment. Results of this study differed slightly from Brown's (2004) earlier study. The major difference was Hong Kong teachers reported having a greater belief that assessment was positive for student accountability. The results for learning improvement and school accountability matched the earlier study. Recently, Brown, Lake, and Matters (2011) conducted a study of teacher attitude toward assessment, both summative and formative, and student achievement. This study concluded that the relationship between teacher attitude toward assessment and increased student learning could be the missing link in school improvement policies.

Statement of the Problem

Extensive research focused on assessment has been completed over the last thirty years (Barksdale-Ladd & Thomas, 2000; Black & Wiliam, 1998a, 1998b; Bracey, 2009; Haycock, 2001; Harlan, 2009; Lacina-Gifford & Kher-Durlabhji, 1992; Nicholas & Berliner, 2008a, 2008b; Noddings, 2005; Popham, 2003, 2004, 2006, 2008; Stiggins, 1999, 2002); however, very little research has been conducted with relationship to teacher attitude about assessment. The process of assessment and application of formal measures of student achievement have become a significant part of educational accountability; in spite of that fact, the attitudes of teachers have not been clearly measured.

Purpose of the Study

The purpose of this study was to explore the possible relationships between teacher attitudes and self-reported behaviors about an assessment tool and student academic growth in mathematics and reading as measured by the same assessment tool. The secondary purpose was to assess teachers' beliefs and self-reported behaviors when teachers are grouped by selected demographics. For example, the study considered the following types of questions. Did attitudes and self-reported behaviors of teachers vary across grade levels and/or schools? Did attitudes and self-reported behaviors of teachers vary when groups are established based on the status of the school (Title I versus Non-Title I)?

Significance of the Study

The significance of this study is multidimensional and could be linked to effective accountability efforts. A key to school reform and improvement is increased student achievement. Exploring teacher attitude toward assessment provided valuable information for professional development. Targeted professional development for teachers based on teacher and school demographic information could promote more effective use of assessment results.

Research Questions

Five research questions guided this study were:

1. Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for mathematics as measured by the Measures of Academic Progress assessment be predicted from three independent variables (general assessment belief and MAP belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*])?

2. Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for reading as measured by the Measures of Academic Progress assessment be predicted from three independent variables (general assessment belief and MAP/MPG belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*])?
3. Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established for certain demographics?
4. Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established in Title One and non-Title One designations?
5. Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established by school?

Definition of Terms

- Adaptive Testing: "In this paradigm, the test adapts to match the difficulty of the questions administered to the performance of each student as the student takes the test. The advantages of this **adaptive-testing** paradigm include increased testing efficiency, and tests that are challenging but not frustrating for each student" (Weiss, 1982, p. 67).
- Assessing: "to determine the importance, size, or value of" (Merriam-Webster Online Dictionary, 2010).

- Assessment: “the action or instance of assessing” (Merriam-Webster Online Dictionary, 2010).
- Attitude: “A relatively enduring organization of beliefs around an object or situation predisposing one to respond in some preferential manner” (Rokeach, 1968, p. 112).
- Belief: “Any simple proposition, conscious or unconscious, inferred from what a person says or does” (Rokeach, 1968, p. 113).
- Formative Assessment: “Formative evaluation intended to foster development and improvement within an ongoing activity” (Pearson, 2008): also referred to as Assessment *for* Learning (Stiggins, 1999).
- Intermediate grades: Third, fourth and fifth grade
- Measures of Academic Progress (MAP): Assessment created and distributed by the Northwest Evaluation Association. This assessment is administered in two parts for this study: mathematics and reading. These two assessments are given to second through fifth grade students.
- Measures of Academic Progress for Primary Grades (MPG): Assessment created and distributed by the Northwest Evaluation Association. This assessment is administered in two parts for this study: mathematics and reading. These two assessments are given to kindergarten and first grade students.
- Primary grades: Kindergarten, first grade, and second grade
- Raush Unit (RIT): A type of Item Response Theory (Lord, 1980) scale that allows a measurement scale that relays the relationship between a student’s achievement and test item characteristics. It is based on a one-parameter logistic IRT model.

- Summative Assessment: Evaluation "used to assess whether the results of the object being evaluated (program, intervention, person, etc.) met stated goals" (Pearson, 2008, p.1): also referred to as Assessment *of* Learning (Stiggins, 1999).
- Value: "relative worth, utility, or importance" (Merriam-Webster Online Dictionary, 2011).

Limitations and Delimitations

The study has the following limitations:

1. Findings cannot be generalized beyond the teaching population of this school district. This information could vary in other districts due to size or professional development philosophy.
2. The sample of schools was drawn from a single school district; therefore, results may not be generalizable to all school districts.
3. The study was restricted to elementary school teachers teaching kindergarten through fifth grade in this school district despite the fact that the assessment is utilized through ninth grade in this district. Results may differ if teachers who work at the secondary level were considered.
4. The utilization of a pre-existing survey which was not designed with this study in mind presented some difficulties.
5. Respondents to surveys in a school work environment may fear lack of complete anonymity; therefore, results may have responses that are not accurate of the respondent's true beliefs. This may produce skewed results toward perceived favorable responses.
6. Many variables outside the control of the researcher could have impacted the results of the teacher survey and/or student growth data. These variables may have

included: time given for the survey, the environment of the survey, the principal's influence or attendance during the survey, the proctoring of the survey by the MAP site coordinator, the qualification and experience of the assessment proctor in both testing sessions, and the testing environment for the students at each elementary school.

The delimitations utilized within this study were determined by the desire to understand the relationship of the teacher attitude toward assessment and the academic growth in the elementary schools located in a single school district. Prior to this study, an attitude survey was administered to the elementary school teachers by the district curriculum and assessment office. To reduce the time burden of administering two instruments, it was decided to use existing data collected by school district personnel. This did not allow for addition of questions that may have provided important insights related to this study.

A second delimitation was that the study used only one school district that has less than four thousand elementary students.

Assumptions

This study included the following assumptions: (a) all the teacher participants in the survey answered the questions honestly and accurately; (b) all the students participating in the Fall 2010 season were encouraged by the assessment proctor to do their best; (c) all the students who participated in the Spring 2011 season were encouraged by the assessment proctor to do their best; (d) all testing sessions were completed under the guidelines of the school district's test security and protocols; (e) all obtained data from the school district is accurate.

Organization of the Study

This research study on teacher perception of assessment and the student growth on that assessment are presented in five chapters. Chapter one includes background of the study, the statement of the problem, the purpose of the study, the significance of the study, definition of key terms, research questions and hypothesis, limitations, delimitations, assumptions, and the organization of the study.

The second chapter presents the theoretical framework and the literature review of the known impact of teacher attitude toward different topics, the different types and purposes of assessment, and the influences on student achievement. Chapter three outlines the methodology utilized for this research project. This chapter includes the instrumentation, data acquisition, and data analysis.

Chapter four presents the study's findings including teacher demographics, the testing of the research questions, multiple regression analysis, and the results of the data analysis as it relates to the research questions. Chapter five provides the discussion of the findings, implications for future practice, recommendations for further research, and conclusions.

Conclusion

This chapter provided the background of this study. The importance of increasing student achievement has consequently raised the level of need for assessment literate teachers. Extensive research on assessment has been completed. General research exists on teacher attitudes toward assessment (Jones & Egley, 2004, 2006) but little research exists about the teacher attitude toward assessments and how that may impact student achievement. This study brings to light a possible relationship between teacher attitude and student achievement.

CHAPTER TWO: LITERATURE REVIEW

Introduction

The purpose of this study was to explore the possible relationships between teacher attitude and self-reported behaviors about an assessment tool and student academic growth in mathematics and reading as measured by the same assessment tool. The secondary purpose was to assess teacher attitude toward assessment and self-reported behaviors of that assessment when teachers are grouped by selected demographics. In this chapter, research will be presented about teacher attitudes and the impact those attitudes have on different aspects in education. Also included in this chapter will be theories and research on assessment, the role of assessments within accountability measures, and use of assessments to promote improved academic achievement. Finally, research on student and school demographics will be reviewed and related to teacher characteristics.

Attitude

According to Rokeach (1968) beliefs are any "simple proposition, conscious or unconscious, inferred from what a person says or does" (p. 113). Therefore a belief is said to "have observable behavioral consequences" (p. 1). People have tens of thousands of beliefs that are organized in some fashion. One way of organization is an attitude. Rokeach (1968) contended that "an attitude is a relatively enduring organization of beliefs around an object or situation predisposing one to respond in some preferential manner" (p. 112). Little distinction is made in the literature between beliefs and attitude except that every attitude is made up of a set of beliefs but every belief does not necessarily belong to an attitude. Each belief within an attitude is set up with three components: affective, cognitive, and behavior. These components dictate a person's response to certain stimuli. Rokeach (1968) analyzed the relationship

between attitude and behavior. He found that although "an attitude leads to preferential (or discriminatory) response, the basis for the preferential response is not clear" (p. 121). The behavioral response may be in reaction to a strongly held belief that something is right or wrong. A teacher's attitude toward assessment then, by Rokeach's definition, would cause certain reactions or behaviors to occur. These behaviors may directly or indirectly affect student academic outcomes.

The relationship between teacher attitude toward assessment and student academic achievement is an important research topic. Researchers and theorists have considered general relationships from a variety of approaches, including but not limited to teacher qualifications, instructional practices, and both student and school accountability. Theoretical discussions have included teacher attitudes related to mandatory high-stakes assessments, pay for performance bonuses, student abilities, as well as a wide variety of other topics (Alvidrez & Weinstein, 1999; Brown, 2004; Diamond, 2007; Heneman & Milanowski, 1998) . One research topic that has not drawn much attention is the possible relationship between teacher attitude toward assessment and student achievement. The following section provides a summary of research related to teacher attitudes related to assessment.

For years, salary increases for teachers were based on two things: length of service and advanced coursework. In general, studies have attempted to link this type of salary schedule to student academic achievement. Some researchers have argued for a performance based salary increases based on the students' academic achievement. Heneman and Milanowski (1998) conducted a study involving K-12 teachers. They found that teacher attitude toward school-based performance awards varied depending on the size of the bonus, the satisfaction with the current level of salary, and the perception of fairness of distribution of the funds. Most teachers

felt that although the bonuses were desirable, they did not provide the incentive to increase student achievement. Heneman and Milanowski (1998) found limitations to their study. Some respondents may have felt inclined to choose socially desirable answer choices based on political correctness and not true personal feeling. The researchers concluded that the extrinsic motivator was secondary to the teachers' desire to improve student learning. The motivator for teachers was intrinsic; these motivators included the ability to raise student achievement and to meet specified goals.

Alvidrez and Weinstein (1999) conducted a fourteen year longitudinal study to examine the relationships between teacher perception of abilities of students and the impact of these perceptions on the later success of the same students. The study involved students from age four to eighteen. Teachers were asked to rate the intelligence for four-year-old children based on the behaviors they demonstrated in pre-school. Researchers found significant results based on socioeconomic status. The teachers tended to correlate higher SES level with higher estimates of intelligence. However, gender, ethnicity, and school location were not significant predictors. The judgments that teachers made were strong predictors of future achievement for students. "Teachers make judgments on a regular basis about the ability of students, and their appraisals can have critical implication for curricular and instructional opportunities and for the messages about ability conveyed to children" (p. 731).

Teacher expectations of ethnic minority children or children from lower socioeconomic groups are generally lower than those for other children (Alvidrez & Weinstein, 1999; McCall, Hauser, Cronin, Kingsbury, & Houser, 2006). Likewise, Boaler (2003) found teachers' expectations and students' attitudes toward learning and success impacted student achievement. She explained that teachers' stereotypes threatened student achievement in the at-risk school

studied. If students of a particular social status or ethnic group were told they are expected to do poorly on a mathematics assessment then they performed below a comparison group of peers who were not told that predestination belief. "Confidence in one's ability to succeed in mathematics is an intrinsic part of success and motivation" (p. 505). This research demonstrated the power the teacher attitudes had in relationship to student achievement.

With the renewed education legislation of the No Child Left Behind Act, a focus on teacher background and education gained new importance. The law considers a teacher's educational endeavors in order to establish whether or not a teacher is certified as highly qualified. However "the stipulations established by NCLB may not be the best indicators of teacher effectiveness; rather, aspects such as teacher attitudes and practices may be superior" (Palarady & Rumberger, 2008, p. 111). Wayne and Youngs (2003) found similar findings in their research review. Teacher characteristics, such as the teacher's certification, degree(s) held, test scores, and even college alma mater, were all examined in a causal relationship with student achievement. In this review, teacher characteristics were found to have influence on student academic achievement especially in high school curricular areas with emphasis on mathematics instruction. However, there was little empirical evidence in the twenty-one reviewed studies that supported that the defined teachers' characteristics as indicated above had any significant impact any curricular areas other than mathematics.

Palarady and Rumberger (2008) found very little research evidence that teacher backgrounds have a direct impact on student achievement with the exception of mathematics. There are other aspects that have greater impact, including teacher attitude and instructional practices (Palarady & Rumberger, 2008). "Compared to teacher background characteristics, teacher attitudes and practices have received less attention in the research literature, in part

because they tend to be more difficult to measure or quantify" (p. 112). The researchers found that teachers' instructional practices were closely related to student learning; whereas, teachers' backgrounds, experiences, and attitudes indirectly related to student learning. Attitudes regarding teachers' expectations and teacher efficacy were a central theme. The researchers found that reading achievement was significantly impacted by teacher expectations. The lower the teacher expected the student to perform in reading, the lower the student's achievement. In mathematics teacher efficacy or feelings of satisfaction with their career made a significant difference in students' achievement. This study also demonstrated the power the teacher attitudes had in relationship to student achievement.

Brown's (2004) study of teacher attitude of assessments defined the term of attitude as the way a teacher interacts with, responds to, and understands assessment. The study considered four generalized purposes of assessment. According to Brown:

> teachers' conceptions of assessment can be understood in terms of their agreement or disagreement with four purposes to which assessment may be put, specifically, (a) improvement of teaching and learning, (b) school accountability, (c) student accountability, or (d) treating assessment as irrelevant. (p. 301)

The first purpose of assessment is that it improves teaching and learning. Any assessment that is perceived to have the goal to improve teaching and learning is believed to diagnose areas of deficiencies. These types of assessments produce valid and reliable data that teachers would consider worthwhile. This is the only instance in Brown's study that teachers felt that the assessments had validity and relevancy. The second and third purposes specifically related to accountability, one for school and the other for students. Teachers reported that assessments that only targeted accountability were equally likely to be relevant as they were irrelevant. This

attitude often led to a disregard of the data obtained from assessments geared toward accountability.

The final purpose of assessment was for the summarization of student learning. This purpose of assessment was significantly interpreted as irrelevant by teachers because of the perception that the data was unreliable, invalid, and useless. The teachers in the study felt that this type of assessment had no legitimate place in school. Teachers also felt that if the assessments, or the results of the assessments, were viewed as encroaching on teachers' professionalism or instructional practices, then it would likely be viewed as an irrelevant activity. Brown concluded:

> The introduction of any assessment policy intended not only to diagnosis and monitor student learning but also to improve the quality and quantity of learning should be done in such a way as to minimize association with student accountability and instead maximize association with teachers' commitment to improving their own instruction and the learning of their own students, while taking advantage of teachers' agreement that assessment can identify quality schooling. (p. 315)

In summary, the study warned policy makers to consider teacher attitude of assessment prior to implementing new testing procedures. The author recommended that it was better to associate the assessment with the improvement of teaching and learning. The assessment is more likely to be viewed as relevant and important when given under the guise of improvement rather than accountability. Many educational policies are revolving around assessments and the results from those assessments. "Success or failure of such policies may hang on the conceptions and meanings that teachers give policies" (p. 301).

Summary

Rokeach (1968) argued that beliefs become attitudes and attitudes become behavior. Although it is possible to behave in direct contradiction to one's beliefs, it is not easy to do without conscious effort (Rokeach, 1968). Teachers have a tremendous influence on student academic achievement (Hanushek & Welch, 2006; Marzano, 2007). The beliefs teachers hold can relate to academic success of students. It is vital to understand the relationships between teacher attitudes toward an assessment and possible effects on the academic outcomes of a student.

Assessments to Measure Academic Achievement

History.

Assessment has become both a popular and highly debated topic in education. For some it is the means to an end (Barksdale-Ladd & Thomas, 2000; Haycock, 2001; Harlan, 2009). For others it is the potential downfall of the American education system (Black & Wiliam, 1998a; Bracey, 2009; Lacina-Gifford & Kher-Durlabhji, 1992; Nicholas & Berliner, 2008a, 2008b; Noddings, 2005; Stiggins, 2002). For many educators the intent of assessment is raising accountability for teaching and learning thereby increasing academic achievement. This concept is not new in the realm of education, but it has gained recent legislative attention. Raising standards of learning in our schools is a national priority (Black & Wiliam, 1998a). For more than two and half decades, the federal government has attempted to raise student academic achievement through legislatively mandated standardized testing. The assessment movement has been hurling forward at break neck speeds; yet, there is no greater assurance that students are doing any better (Barksdale-Ladd & Thomas, 2000).

Many educators refer to the 1983 publication *A Nation at Risk* (National Commission of Excellence in Education) as the launch pad for the accountability movement in United States. Twenty years later, Guthrie and Springer (2004) found this report was likely "to render the nation's education system more effective for students and more useful for the larger society" (p. 35). The Nation At Risk (1983) "contributed forcefully to a vastly enhanced federal government presence in American education. A centuries-long American tradition of state plenary authority and local operating discretion is now giving way to a pressing national uniformity of federally imposed accountability requirements" (Guthrie & Springer, 2004, p. 36).

In 1988 the reauthorization of the Elementary and Secondary Education Act (ESEA) added requirements for school districts to use standardized test scores to assess the academic achievement of individual schools. In 1997 an additional component was added to the assessment movement with the passage of Public Law 105-78 commonly referred to as Obey-Porter. This law authorized reform programs to aid failing schools by providing monies to adopt better curricular programs (McDougall, Saunders, & Goldenburg, 2007). In 2001, the ESEA was again amended and became known as the No Child Left Behind (NCLB). The goal of this reauthorization was one hundred percent of the students in the United States will score at or above proficient by the year 2014.

In the past there was a belief that assessments served two purposes: to inform educators about instructional decisions and to motivate learners by setting parameters for academic success (Stiggins, 2002). There are currently two main forms of assessment; however, the intention of each differs greatly. The first is summative assessment, the assessment *of* learning. This assessment type is typically attributed to account for the academic achievement for a child, school, district, or educational system; it provides a summary of past performance. These tests

are the basis of the NCLB legislation. Testing sessions occur infrequently and are attached to rewards or punishments. These kinds of assessments are used to make large scale educational decisions at all levels. These decisions could lead to a change in performance standards, curriculum programming, or teacher preparation and professional development.

The second category of assessment is formative assessment, the assessment *for* learning. This assessment type is used by teachers to make changes in instructional practices within their classroom. Theses assessments are on-going, attached to classroom instruction, and thought to be the crux of educational decision making by most classroom teachers.

Hargreaves (2005) suggested that summative and formative assessments are not polar opposites. She proposed to consider knowledge and results in this fashion:

> A conception of knowledge as external to the learner and fixed- at one extreme; and the conception of knowledge as constructed or co-constructed by the learner/s and as fluid – at the other extreme. It is the view of knowledge held by those who have power in education that is likely to determine what sort of assessment and learning happens, to determine the purpose of assessments, what actually comes to be assessed, and therefore the quality of the student's learning processes and products. (p. 224)

However, there is a gap between the two types of assessment in design and purpose creating a rift between political and practical sides of education.

Summative Assessment.

The intention of summative assessment is to determine the current state of educational achievement and to provide a basis to make appropriate adjustments to promote improvements (Stiggins, 1999; Barksdale-Ladd & Thomas, 2000). This is the underlying theory of the 2001 reauthorization of the ESEA known as NCLB.

The impetus for the use of summative assessments is that Americans believe that students need to be able to compete with international students in the global marketplace; the stigma of failure is a significant political issue (Popham, 2009a). For more than thirty years educators have worked to raise student achievement, particularly students of minority backgrounds and students who live in economically disadvantaged communities. Yet student achievement has not increased. The politicians argue that teachers and administrators have failed to raise achievement; therefore, educators must be held to a greater degree of accountability. Educators are being forced to improve through the legal fiat or receive sanctions for failure. The procedures within the act are increasing pressures intended to promote success, i.e. academic achievement in public schools. However, the United States is not the only nation engaging the use of assessments to increase performance. These pressures are international and are as often misunderstood. Black and Wiliam (1998a) of England, found that "in the U.S. similar pressures have been characterized by a distrust of teachers and a belief that external testing will, on its own, improve learning" (p.142). In 1995, six years prior the enactment of No Child Left Behind push, Rick Stiggins warned:

> It is naïve at best and dangerous to student well-being at worst to believe that we can maximize school effectiveness by inducing fear of public censure for low standardized test scores. While this form of public accountability can motivate action for school improvement and so should not be eliminated, we must understand that merely raising anxiety levels as a means of raising test scores will be nonproductive. (p. 197)

High-stakes testing is typically aligned with summative assessments. "High-stakes testing is the practice of attaching important consequences to standardized test scores, and it is the engine that drive the No Child Left Behind (NCLB) Act" (Nichols & Berliner, 2008b, p.

672). Since the NCLB inception in 2001, this type of testing is widespread and has caused a certain level of panic throughout the educational field. Although assessments are not a new concept, the role they play in schools has changed.

Many additional and substantial issues have been raised with politicians in the hope of creating an appropriate environment of high-stakes testing. One of these is the concept of assessment promoting achievement, which has been left unsupported (Harlan, 2009). Despite the common sense of taking into account mitigating circumstances such as learning disabilities, English language acquisition, or socio-economic status, the progress toward a new system of accountability has not happened. Other ramifications of high-stakes testing continue to fester in the education realm. For example, the harsh consequences attached to test scores are leading to unethical behaviors among educators. The problems of NCLB, according to Guilfoyle (2006), include providing incentives for dishonest professional practices such as encouraging low-performing students to be absent for the assessments. The policies give schools opportunities to demonstrate improvement in test scores, but have failed to raise actual student achievement.

Harlan (2009) contended that in order for sound decision making to occur in relationship to test scores, the assessments must be valid and reliable. It is only when validity and reliability are established that the inferences derived from assessment results are justified and usable. Many teachers and administrators believe that these high-stakes tests have missed the mark on both requirements. On the issue of validity, Popham (2006) discussed a problem common to school improvement efforts. He stated,

> Although these key state standardized tests vary substantially, the vast majority of them are instructionally insensitive- that is, they're unable to detect even striking instructional improvements when such improvements occur. This significant short-coming arises

> because these tests are so strongly linked to students' socioeconomic status that they tend to measure what students bring to school rather than what they are taught there. (p. 2)

Nichols and Berliner (2008b) suggested that tests such as the ones included in the accountability measures of NCLB tend to help the elite or upper class. The reinforcement of lower class status is now reinforced by adequate yearly progress status as well.

Besides student assessment validity, teacher accountability issues are also problematic. For example, measurement of teacher's productivity by using isolated data points does not accurately assess the teaching and learning. Assessment results from tests given over the course of a few days cannot accurately explain the impact a teacher has on a day-to-day basis. The snapshot of student assessment may not take into account the entire educational experience of a school year. The inclusion of threats or incentives to improve success is not aligned to the way schools should operate (Nichols & Berliner, 2008b). "Ethical questions are asked when educational decision making is made based on test results" (Lacina-Gifford & Kher-Durlabhji, 1992, p. 565). Noddings (2005) stated:

> Public schools in the United States today are under enormous pressure to confirm – through improved test scores – that they are providing every student with a thorough and efficient education. The stated intention of No Child Left Behind (NCLB) is to accomplish this goal and reverse years of failure to educate many of our inner-city and minority children. But even if we accept that the motives behind NCLB are benign, the law seems fatally flawed. (p. 8)

Stiggins (2002) indicated that politicians saw themselves in a role of intensifying the level of intimidation that is associated with high-stakes assessment as a means of increasing student achievement. In Stiggins' opinion, a "flawed" view of the assessment role is that it is a

great motivator; students will strive to reach higher if the challenge is continuously raised. This view suggested students are motivated to achieve high test scores. But Nichols and Berliner (2008a) suggested that all levels of performers are experiencing increased levels of apathy due to the overemphasis of test taking. Popham (2006) concluded:

> Given insensitive standardized tests and unrealistic AYP targets, more and more teachers are apt to abandon sound instructional strategies like classroom assessment for learning. You see, even this powerful classroom assessment strategy won't be able to increase students' scores enough so that most schools can avoid AYP assassination. Pressured teachers, then will most likely succumb in desperation to any sort of quick-fix score-raising techniques that offer the promise of AYP success-even though some those techniques, such as relentless test-prep drilling using practice items practically cloned from the state's standardized exam, are educationally unsound. (p.2)

Stiggins (2004) identified several critical mistaken beliefs regarding assessment. One of those mistaken beliefs is that high-stakes tests motivate students to learn; for many students, just the opposite occurs. "High-stakes tests without supportive classroom assessment environments harm struggling students" (p. 24). Another identified false belief is that educators do not need to understand the principles of sound assessment practices and that the testing professionals are the only ones who need to understand assessment practices. "In fact, our collective assessment actions over the past 60 years reveal a fundamental lack of trust in teachers and school leaders to accurately assess the achievement of their students" (p. 26).

In 2006, Popham stated:

> Given insensitive standardized tests and unrealistic AYP targets, more and more teachers are apt to abandon sound instructional strategies like classroom assessment for learning.

> You see, even this powerful classroom assessment strategy won't be able to increase students' scores enough so that most schools can avoid AYP assassination. Pressured teachers then will most likely succumb in desperation to any sort of quick-fix score-raising techniques that offer the promise of AYP success-even though some of those techniques, such as relentless test-prep drilling using practice items practically cloned from the state's standardized exam, are educationally unsound. (p.2)

Wilson (2009) found that "despite substantial progress in test design, standard administration of norm- or criterion-references tests given at the end of the year or unit provides little guidance for classroom instruction" (p. 68).

Another problem related to the high-stakes testing phenomenon is resource depletion. Standardized assessments have drained the resources and time allotments that would be better spent on gathering more dependable information about student achievement on a day to day basis (Stiggins, 2002). Summative assessments serve the purpose of accountability for the federal government. Millions of educational dollars are absorbed into mandated standardized testing and the results have not been conclusive. Strong opinions for both sides exist. Nichols and Berliner (2008a) stated, "In the current high-stakes environment teachers, students, parents, and American education are being hurt by required high-stakes testing. This policy is corrupting our education system and needs to be stopped" (2008a, p. 676). However, an acceptable alternative to high-stakes testing has not been developed.

Formative Assessment

"Assessment for learning is any assessment for which the first priority in its design and practice is to serve the purpose of promoting students' learning" (Black, Harrison, Lee, Marshall, & Wiliam, 2004, p.10). This type of assessment is formative assessment. Many researchers

believe this could lead to significant increases in student achievement (Barksdale-Ladd & Thomas, 2000; Black & Wiliam, 1998a; Black et al., 2004; Black, McCormick, James, & Pedder, 2006; Cimbricz, 2002; Diamond, 2007; Faulkner & Cook, 2006, Guilfoyle, 2006; Hargreaves, 2005, Harlan, 2009; Stiggins, 2002, 2004, 2009). Formative assessment is a process which involves gathering evidence of learning on a consistent basis (Heritage, 2007). It is not the number of times a student is assessed but the nature and purpose of the assessment. The way results are used makes the assessments worthwhile (Black & Wiliam, 1998a; Stiggins, 1995,1999).

The intention of formative assessment is to inform the teacher and the student about the current state of learning as well as future educational needs. Specific and timely feedback on a student's progress is essential. Such feedback is the key to making modifications to teaching practices (Black & Wiliam, 1998a; Guilfoyle, 2006). Assessment results need to be descriptive and provide information to the learner about weaknesses, strengths and future needs. (Stiggins & DuFour, 2009). "In this era of accountability, classroom assessments are the foundation of a truly effective system" (p. 286). Stiggins (2002) summarized the role assessments should promote if educational excellence is achieved within an environment of high-stakes assessments by stating:

> Standards frame accepted or valued definitions of academic success. Accountability compels attention to these standards as educators plan and deliver instruction in the classrooms. Assessment provides evidence of success on the part of the students, teachers, and the system. (p.759)

However, formative assessment is not without problems. First, classroom assessments can lack reliability and validity. Many teachers have not received the proper education on test

design or statistical analysis. According to Guilfoyle (2006), the leading professional development need of both pre-service and practicing educators is sound assessment practice and utilization. Many teachers do not possess the necessary skill set to design, implement, and analyze classroom assessments. If the assessments themselves are not sound in design, the results may be challenged. If the assessment results are not analyzed and understood, the results have limited value (Popham, 2009b). Harlan (2009) felt that not only do teachers need assistance with assessment, but there is also a need for professional development in the creation and use of data sensitive instructional practices. Teachers need to increase analytical and critical thinking activities within their classrooms. The instruction should not only address the state specified curriculum but should also be responsive to students' learning needs.

A second issue is accountability. There must be a system of accountability related to student learning outcomes. Research has supported the use of formative assessment. However, federal mandates do not recognize the validity and reliability of this type of assessment. Formative assessment is considered a best instructional practice; yet, present accountability measures do not include the data from formative assessments.

Research on assessment and corresponding impacts have been significant during the last decade. Assessment has come to play a vital role in the United States educational system. Assessments are intended to improve instructional practices and student achievement (Faulkner & Cook, 2006). The prevalence of high-stakes testing throughout the United States in compliance with federal legislation has sparked many research studies. The focus of much research has been to determine if high-stakes testing has resulted in increased student academic achievement. Researchers seem to disagree not only on the intention but the results as well.

Diamond (2007) found that teachers changed the content upon which they focused in their classrooms because of the subject matter included within the high-stakes assessment. Content adaptations were associated with policy changes concerning standards of instruction or benchmark skills. The teachers used high-stakes test results to inform the content that needed to be emphasized, but teachers did not use these results to gauge the effectiveness of their lesson delivery. Changes in classroom practices included teaching multiple choice test preparation, drilling vocabulary tests, repeatedly reviewing key content, and timed tests practice. Results indicated that although the teachers felt compelled to teach the strategies and participate in test preparation activities, these practices did not dominate classroom instructional time. In conclusion, Diamond (2007) found less than twenty percent of the teachers viewed that their pedagogical decisions changed based on testing, and this suggested that policies in high-stakes testing may influence content more so than instructional practices.

Vogler (2002) indicated that proponents of a measurement-driven reform system contend that by applying increased pressure, student academic achievement will increase; inclusion of mandated tests will ensure that particular subject matter is taught. Testing has appeared to be the best way to account for success or failure of students and teachers. If the standards are clear, then assessing them does not seem harmful. Holding students and staff accountable is logical (Vogler, 2002). On the other hand, policy makers and researchers often hold different views on the issue of assessment and accountability. Policy makers have written policies. Some researchers have used these policies in their studies and have determined that the current testing system is inherently bad for education (Barksdale-Ladd & Thomas, 2000).

Despite years of academic endeavors in eliminating the achievement gap through forced high-stakes testing, there is only a marginal narrowing of the gap between African-American

students, Hispanic students, and students from high poverty with those students of European decent and lower poverty (McCall, Hauser, Cronin, Kingsbury, & Houser, 2006). The achievement growth over time for individuals or groups of students must be measured in order to assess the achievement gap. This is where NCLB has failed to provide for growth data. By taking a moment in time snapshot of student academic achievement and applying benchmark scores, the apparent gap perception is skewed and misrepresented (McCall, Hauser, Cronin, Kingsbury, & Houser, 2006).

Cimbricz (2002) found a disconnection between theory and research when considering assessments, teachers, and instructional practices. The dominating idea is that assessments, if focused on instruction and improved teachers' practices and attitudes, would increase the students' achievement. However, for what little research was available at the time, the relationship between assessment and teacher attitude was not positive. Cimbricz's work in the area of assessment illustrated that research dating back two decades has proved that high-stakes testing has significantly and negatively influenced the actions and beliefs of teachers. She identified five areas that were drawbacks to high-stakes testing at the state level as:

> 1)Narrowing of the curriculum and instruction, 2) fostering anxiety, confusion, fear, shame, anger, and/or mistrust, 3) deskilling of teachers and/or a perception of powerlessness, 4) the invalidity and inadequacies of these tests as accurate measurements of what is taught and learned; and 5) loss of instructional time due to test preparation and testing. (p. 6)

This does not necessarily mean that there is no room for accountability through assessment in today's educational system, but it does indicate that significant changes must take place in order for accountability and improvement to happen. She later stated:

> For if state-mandated testing continues to be viewed as a viable mechanism of educational reform, it is necessary to understand the ways in which this mechanism is mediated through the local contexts and the minds, motives, and actions of teachers. (2002, p.16)

In another study conducted on high-stakes assessment and the system of accountability that is tied to the assessment results, Barksdale-Ladd and Thomas (2000) found the following:

> No teacher argued for a lack of standards or for dismissing assessment. In fact, teachers did not want to lose what they considered best practices in assessment: assessment that (a) provides feedback to help students improve their learning; (b) is a part of a student's work, which can go into a working portfolio; (c) provides flexibility and does not dominate the curriculum; (d) informs instruction to help teachers improve their teaching; thereby ensuring student learning; (e) over time is classroom based; and (f) uses more than one measuring stick for assessing students' learning. (Barksdale-Ladd & Thomas, 2000, p. 395)

They concluded and proposed a sensible balance may lie in the point of view put forth by Christensen (1999) who asserted:

> The question for anyone who cares about kids is how do we retain our critical stance on assessments while preparing students for them? Can we "teach the tests" without compromising what we know to be true about teaching and learning? (p. 396)

The impacts of a whole school reform of student achievement and test scores were examined by McDougall (2007). Strong leadership and a properly chosen reform model were associated with the highest gains. Although assessments were not the focus of this research project the results indicated that teachers felt that assessments were helpful in determining the

learning accomplishments of the students, but were not informative about school accountability. Assessment was useful as a tool to help the teachers' instruction. "No single assessment can answer all these questions. A productive, multi-level assessment system is needed to ensure that all users are served so all instructional decisions can be made well" (Stiggins & DuFour, 2009, p. 642).

Stiggins (2001) argued that for many educators, assessment fits into the educational system because of a desire to rank students based on academic ability. Highly ranked students continued on to college whereas lower ranked students entered the work force immediately after high school. However, this ranking system is no longer acceptable within the contemporary world of work. The business community needs a work force that is successful as critical readers, mathematic problem solvers, and competent writers and has held the educational system to a higher standard in order to receive a prepared work force.

Nichols and Berliner (2008b) stated tests are the current method of measuring productivity in the nation's schools. It is difficult to compare the manufacture of goods with the development of children. In business, if tests are positive then production improves, but when applied to schools the model begins to disintegrate. The productivity in schools must take into account the existence of outside influences. Economic status, cultural background, and language differences all impact the output of a school. The inputs in a school cannot be controlled in the same manner as the inputs in a factory. The outcomes are therefore less predictable for a school than a manufacturing company. The competitive nature of many people in the United States could possibly explain how high-stakes testing has easily been slipped into contemporary education system (Nichols & Berliner, 2008b). In a nation of sport enthusiasts, the competitive nature likens the assessment results into the winner/loser mentality. Using this

same analogy, very little time is taken to consider the contributions of individuals in a team sport just as we take very little time is taken to consider other contributing factors of students' academic achievement beyond their test scores (Nichols & Berliner, 2008b). Brown (2004) defined assessment as "any act of interpreting information about student performance, collected through any of a multitude of means or practices" (p. 304). This type of formative assessment is where research and theory seem to meet and agree.

Summary.

When reviewing the last thirty years of educational reform law one thing seems certain: assessment in American schools is not going away. Summative and formative assessments have very important roles in today's schools. The use of a "single shot" summative assessment to judge a school's academic progress is highly contested by many in the educational community. Also contested is the use of assessment data to prescribe sanctions which include eliminating administrators and teachers as a way to improve students' academic achievement. Formative assessment has been hailed by a few in the educational community as the essential element in real school improvement. When quality teachers have access to valid and reliable formative assessments and the skills to relate them to better instructional practices, student achievement is a likely by-product of these instructional practices.

Much of the literature gathered focused primarily on the theoretical nature of assessment and accountability and less on the empirical nature. High-stake assessments have changed instructional practices when it came to test preparation (Popham, 2009a). A change in content was more evident (Vogler, 2002). Teachers believed high-stakes testing has had a negative impact on instructional practices. Many felt that there had been an abandonment of the best teaching practices for those which are quick and more efficient in getting the tested material

covered (Popham, 2009b). Teachers most often held a negative attitude toward high-stakes assessments (Palarady & Rumberger, 2008).

Student Academic Achievement

Student academic achievement is often defined by test scores. The No Child Left Behind (NCLB) Act of 2002 requires states to report achievement scores for mathematics and reading from grades three through eight. Science achievement is encouraged but the results are not included in adequate yearly progress status reports. Each state is required to set a standard against which all students in the state are measured. The states then are allowed to set benchmark percentages that stair-step to the hundred percent proficiency mark at the year 2014. School districts and schools not meeting designated benchmarks receive federal sanctions and state punitive actions as outlined in the NCLB legislation. Student academic achievement is at the heart of all these legislative efforts.

The increase of student academic achievement is the reason for school improvement initiatives, school reform acts, and countless hours of debate on school accountability. These efforts and others have become operationally defined in terms of improved student test scores (Barksdale-Ladd & Thomas, 2000). Because improved student achievement is the desired target, key factors that drive achievement have been studied extensively.

The factors associated with student achievement are complex, often interact, and vary. Researchers have studied many of these factors, including parental involvement (Casanova, Garcia-Linares, de la Torre, & de la Villa Carpio, 2005; Jacobs & Harvey, 2005), the role of specialty classes such as physical education (Taras, 2005; Podulka Coe, Pivarnik, Womack, & Reeves, 2006), the role of school culture and violence (Chen & Welkart, 2008), and the student's

assessment preference and style (van de Watering, Gijbels, Dochy, van der Rijt, 2008). In the following section, student academic achievement is explored.

Impact of Family Demographics.

Children typically enter public education (kindergarten through 12th grade (K-12)) at the age of five and continue through their eighteenth year. The influence of the first five years is greatly impacted by a child's family demographics. The ethnicity of a child, income level of the parents, educational attainment of the parents, location of the family residence, and many other factors are all included in the demographics of a child. Some researchers have tried to use these factors to predict achievement in primary and secondary schools (Casanova et al., 2005; Jacobs & Harvey, 2005).

Casanova et al. (2005) found that socio-demographics, or income level of the parents was a predicting variable of student achievement for teenagers experiencing normal academic achievement. However for low academic performing students many factors including family discipline style, family socio-demographics, and family involvement could predict low performance. "Our data show lower levels of acceptance, control, involvement, and parental expectations for students with low achievement" (p. 433).

In a study of middle school students who drop out of school, Rumberger (1995) concluded that for adolescents the background of the family was the most significant predictor of students' success in school. He found that the majority of the middle school students who dropped out were from minority, low income families where one or both parents had left the family unit. In a study fashioned from Rumberger's 1995 study, Jacobs and Harvey (2005) examined student academic achievement and family involvement at a school level; they collected data from high, middle and low achieving schools and from parents regarding their beliefs about

the role of schools. They found that parents from the high achieving schools had the highest desired and expected levels of educational attainment for the children. These parents expected their children to not only have high academic achievement but also strive to complete more advanced levels of education. According to the parents from high achieving schools, it is the responsibility of the schools' staff to provide the means for children to reach their fullest potential and continue well prepared on to higher education work. In contrast to parents whose children attend high achieving schools, parents whose children attend low and middle achieving schools reported that they did not believe that success was solely based on educational levels of attainment or academic success. These parents were interested in the schools' role of assisting in personal growth and character building of the children.

A remarkable finding was that parents of all three school groupings were satisfied with the education their children received. Personal expectations were met by each group of parents but for drastically different reasons. The authors concluded that academic success of the students could be predicted based on variables such as family background, educational attainment by the parents, and degree of expectation for school success (Jacobs & Harvey, 2005). This confirms results of the previously reported research from Rumberger (1995) that family demographics and parental influences are vastly important in academic achievement of students.

Impact of School Related Influences.

When students enter school, everything and everyone influence their achievement (Casanova, Garcia-Linares, de la Torre, & de la Villa Carpio, 2005; Chen & Welkart, 2008; Jacobs & Harvey, 2005; Podulka Coe, Pivarnik, Womack, & Reeves, 2006; Taras, 2005). For example, minutes of mathematics instruction, the disposition of the lunchroom staff, the level of school safety, the environment of the playground, and teacher characteristics are factors that

impact education. In addition, the school environment interacts with individual student characteristics.

Researchers have attempted to isolate the factors which cause academic achievement to increase. Research studies that take into account the number of minutes of play, reading, writing, mathematics and other classes such as art, music, and computer skills have been all disaggregated to study the causes of the perceived failure of American schools.

For example, studies have considered the relationship between physical education and academic achievement. In a study of scheduled physical activity, Taras (2005) found that although physical education classes may have increased a student's enjoyment of the school, it did not translate into higher academic test scores. In a similar study, Podulka Coe, Pivarnik, Womack, & Reeves (2006) found that rigorous physical activity in physical education classes affected grades in a positive manner but such activities did not have an effect on standardized test scores.

The topic of safety is an important issue at every school. According to Chen and Weikart (2008), school safety is a part of school order or disorder and that "school disorder and student academic achievements are correlated " (p. 4). They also found that school attendance can be a predictor of student academic success. They concluded that students from lower socio-economic backgrounds attended school less than higher socio-economic peers. They also found that "school culture, as measured by student attendance, has a direct positive effect on student achievement" (p. 7). Lower attendance rates also were correlated with greater amounts of reported school violence and decreased student academic achievement. Chen and Weikart (2008) concluded that school violence and low attendance rates were a byproduct of school culture in

schools with high urban poverty. They argued that the only way to close the achievement gap between high and low achieving students is to reduce urban poverty (Chen & Weikart, 2008).

Reduction of urban poverty to close the achievement gap was supported in the meta-analysis of research performed by Sirin (2005). He wanted to replicate a previous study completed in 1982 with research from 1990 to 2000. The studies concerned the impacts of poverty and low socioeconomic status on student achievement. The achievement gap between students of high socioeconomic status and low socioeconomic status increased as students grew older. Sirin (2005) concluded that the key contributing characteristic with the strongest correlation to student academic achievement is the family socioeconomic status.

> Beyond the main findings, the results from this review also show that the magnitude of the relationship between SES and academic achievement is contingent upon several factors. More specifically, methodological characteristics, such as the type of SES measure, and student characteristics, such as student's grade, minority status, and school location, moderated the magnitude of the relationship between SES and academic achievement. (p. 438)

Caldas and Bankston (1997) considered the effects of a students' school socioeconomic status on individual academic achievement. They analyzed potential impacts that peers played in school success. The findings indicated that not only did student's own SES affect his/her achievement but the socioeconomic status of peers may also influence academic achievement. The investigators concluded that the impact of the school and peer socioeconomics may be due to teachers in lower socioeconomic status schools. These teachers may have lower academic standards for an entire class based on personal bias that certain students from particular student groups are academically weak. This attitude then impacts the entire class of students on an

individual level. In the study, when the income level of the student was controlled, the socioeconomic status of the peers negatively impacted the academic achievement of the individual. However, "although the family background of schoolmates has a demonstrably important influence on individual student achievement, it may not be as important as the influence of the student's own social status" (p. 274). The researchers concluded with the recommendation that the student population should be distributed so that less than thirty percent of the school population is from low socioeconomic backgrounds.

Studies such as those cited above suggest that for students living in economically disadvantaged communities are at risk for achieving academically; yet there are studies that have shown the powerful nature of an effective teacher. There are many attributes of successful teachers that seem to influence student achievement. When research is controlled for student demographics teacher attributes leading to increased student academic achievement can be identified. The manner in which a teacher structures the classroom and the activities for the students can determine how students perceive themselves as successful. Meece, Anderman, and Anderman (2006) found that when a teacher emphasized mastery, understanding, and improvement of skills and knowledge, students were more likely to be internally motivated and develop positive learning patterns. The authors indicated that provisions in the No Child Left Behind Act of 2001 (NCLB) proposed the opposite type of environment in schools. The competitive nature of NCLB and the public scrutiny of failure are contradictory to their findings. The performance goals or benchmarks set by the law are "negatively related to intrinsic motivation, to adaptive form of coping in the presence of challenge and failure … (they) are also correlated with greater self-reported cheating and disruptive behaviors in the classroom, which can reduce learning opportunities for all students" (p. 498).

Finally, student behaviors associated with assessment has been a subject of research. In a study of ninth graders, Alkharusi (2008) found that assessment practices of teachers had a great impact on personal achievement goals of the students. Through his study he concluded:

> …student characteristics such as self-efficacy and a perceived harsh assessment environment as well as class characteristics such as class gender, class average self-efficacy, and class average perceived learning assessment environment provided the best model for explaining differences in mastery goals. (pp. 259-260)

In a similar study by van de Watering et al. (2008), students in Belgium were surveyed about their test preferences and attitude toward the assessment itself. This study found a relationship between preferred assessment methodology and assessment results. A student who was given an assessment in a method preferred format and context had greater achievement scores than when given a non-preferred assessment format on the same academic material. They also found that students' perception of the cognitive demand of the assessment was not typically correct. For example, students perceived multiple choice questions as easier to answer, but they also thought of the multiple choice questions testing at a higher level of cognition. However, multiple choice questions are more likely to be at lower levels of cognition. There was a disconnection in the students' study habits resulting in lower results on multiple choice questions. The investigators also concluded that test anxiety factored into the test results. Those with high test anxiety often favored multiple choice questions which have a level of certainty. In conclusion, they found that "a mismatch in the perception of assessment leads to poorer assessment results" (p. 656).

Summary.

Research suggests that if students from a middle or high income family go to a school where attendance is high, , the environment is safe, where mastery and improvement are equally important, and the right assessment is given at the right time in the correct format then student academic achievement will increase. Many factors affect student achievement but understanding these factors could add vital information for school accountability measures and improved student academic outcomes. If student demographic variables can be controlled then a greater understanding of teacher attitude toward assessment and the possible relationship to student academic achievement might be a focus for future research.

Conclusion

According to legislative mandates like No Child Left Behind Act of 2001, student achievement must increase. NCLB requires that 100% of the student population must reach proficiency on academic standards by 2014. Student achievement, for a typical child, can possibly be predicted almost at birth. If these factors are not enough to explain a student's academic success or failure then the prescribed assessments themselves must be understood to a greater degree. The number of opinions on validity and reliability of the high-stakes tests are plentiful (Popham, 2003, 2004, & 2006).

Theories and research on teachers' influence in the classroom are increasing. From pay for performance research to theories concerning a teacher's perception of students can influence their academic achievement immediately and in the future are topics all taken under consideration. A merging theme is the substantial influence a teacher has on their students. Researchers have concluded that there is no substitute for an effective teacher when considering

school improvement (Chappuis, S., Chappuis, J., & Stiggins, R., 2009; Diamond, 2007). A good teacher is one of the essential keys to unlocking student potential.

The belief system of a teacher can directly impact students. However the question remains whether teacher attitude toward a particular assessment itself can indirectly influence the achievement of the students? This question remains unanswered in the available research. Perhaps it is a socially offensive topic to discuss in light of the assessment environment at the federal, state, and district level. It would be socially unacceptable to highlight how a teacher's poor attitude toward assessment could negatively impact student performance. It might become an ethical issue as discrimination based on ethnicity or social class has over the years. It may be that there are very few assessments that could be directly correlated to teacher attitude available for research. Awareness of the impact an attitude toward assessment can possibly have on student performance may be something that could help turn around some schools from failure. Future research on this topic, the relationship between teacher attitude toward assessment and student academic achievement, could be the missing link in school improvement initiatives.

CHAPTER THREE: RESEARCH METHODOLOGY

Introduction

The purpose of this study was to explore possible relationships between teacher attitude toward assessment and student academic achievement. The study focused on data collected from teachers who work in six elementary schools in a single school district. This study used two sets of existing data. This chapter is divided into four sections: (a) a description of the survey instrument and its administration, (b) a description of Measures of Academic Progress (MAP) and Measures of Academic Progress for Primary Grades (MPG) assessments and their administration, (c) the process of acquisition from the school district and the subsequent filtering and coding for the purpose of this study, and (d) data analysis.

Professional Development Survey

Data obtained from the District Staff Survey for Professional Development (SSPD), a forty-four question computerized survey, was used in this study. The existing data was collected by the district statistician and contained no identifiable information. The authors of this survey were Dr. Ricky Medina, the school district statistician, and Dr. Steven Pradere, the director of grants and special projects for the school district. The original survey was designed to collect information related to the Measures of Academic Progress (MAP) and Measures in Academic Progress for Primary Grades (MPG) on the following topics:

- The extend of usage of the for
    - Academic planning
    - Academic progress reporting
    - Evaluation of individual student academic growth
- The professional development adequacy as it pertained to MAP/MPG

- The adequacy of time given within the school day to utilize the different components of MAP/MPG
- General demographic of participants

For the purpose of this study, the survey was combined into three subscales, which are related to general belief about assessment, perceived value of MAP/MPG, and teacher self-reported behaviors associated with MAP/MPG. This instrument had two Likert-like forced choice formats. The first format consisted of the following six choices: very strongly disagree, strongly disagree, disagree, agree, strongly agree, and very strongly agree. This format related to the subscales of general belief of assessment and perceived value of assessment. The second format consisted of the following four choices: to no extent at all, to a slight extent, to some extent, to great extent. This format related to the subscale of self-reported behavior of use of assessments.

According to Sprinthall (2007) content and face validity can be established without the use of statistical measures. Face validity can be determined "by simply looking the test over and judging whether it has the look and feel of what it intends to measure" (p. 515). Content validity is based on whether or not the items are representative of the general domain of questions. This type of validity is important to ensure that the instrument has a range of questions for each intended target and the balance of subjects within is evident. In order to establish content and face validity of this instrument the authors developed a committee of experts. The committee consisted of fifteen elementary teachers, an elementary vice principal, and the Associate Superintendent of Curriculum. This committee screened each item in terms of understandability, wording, content overlap, compliance within the teachers' union negotiated language, and time necessary for item completion. Two items were removed to avoid contracted

language conflicts. The remaining items were included in the survey and were considered valid measures of attitude and self-reported behavior. The survey had not been used by the school district before and was designed in summer of 2010 for implementation in fall of 2010. A copy of the Professional Development survey is presented in Appendix A.

Reliability of the instrument was established using the approaches suggested by Sprinthall (2007). "Cronbach's alpha is a measure of internal consistency, that is, how closely related a set of items are as a group" (Introduction to SPSS, 2010). Due to available data, test/retest and alternate reliability were not possible. Cronbach's coefficient alpha can be used with Likert-type scales. An advantage of using this measure of reliability was the ability to identify which items do not contribute to the overall reliability of the instrument because "each and every item has to be individually assessed for variability" (Sprinthall, 2007, p. 512). Alpha is a correlation coefficient; therefore, the range is -1 to +1. As the value of alpha approaches +1, the strength of the internal consistency of the instrument increases. Cronbach's reliability coefficient for the entire instrument was 0.952 as reported by the district statistician.

Data Collection.

The participants who completed the survey were teachers in six elementary schools in a small city school district. The participants consisted of182 classroom teachers from kindergarten through fifth grade. The study used a population of convenience. Because the population size was relatively small, there was no attempt to sample. All teachers were encouraged to participate in the survey during a fall 2010 faculty meeting. Teachers who were absent the day of the data collection had the survey emailed to them and were encouraged by the site MAP coordinator to complete the survey in a timely fashion. Although special education teachers and English as a second language teachers (ESL) participated in the district survey,

these two categories of teachers do not have classroom MAP data. The classroom MAP data was the dependent variable for the research question one and two. Special education and ESL teachers do not have test data for MAP so these two groups of teachers were excluded from the regression portion of this study.

Northwest Evaluation Association Assessments

The Northwest Evaluation Association's (NWEA) Measures of Academic Progress (MAP) was utilized in the areas of mathematics and reading. This instrument has two different formats that are grade level dependent. The form of the MAP given to kindergarten and first grade students was referenced throughout this study as the Measures of Academic Progress for Primary Grades (MPG) and the assessment given to second through fifth grade students is the Measures of Academic Progress Survey 2-5 (MAP). Each of these assessments are divided into two academic areas, mathematics and reading.

Measure of Academic Progress Assessment for Primary Grades.

The Northwest Evaluation Association's (NWEA) Measures of Academic Progress for Primary Grades (MPG) in mathematics and reading was used to assess student academic growth in mathematics and reading for students in kindergarten and first grade. The assessment was created by NWEA and has been aligned with content standards for mathematics and reading in the state where the study was conducted. A key reference for this assessment can be found in the *Technical Manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades* (2008).

The mathematics assessment consisted of two sets of thirty-five items. The items were presented in a forced choice computerized format utilizing several different answer options. The questions were audio enhanced to increase reliability due to reading ability limitations. There

were six subscales for the MPG for mathematics assessments. Those subscales were number sense and computation, patterns and algebra, measurement, geometry, data analysis and probability, and problem solving. The skills in the subscale of number sense included counting, identifying and representing whole numbers, relative position and magnitude, place value with the base-ten system. The skills represented in the computation subscale included addition, subtraction, and the readiness for multiplication and division. For the subscale of measurement and geometry, the skills sets included attributes, measuring and estimating, identifying 2-D and 3-D attributes, and spatial transformations and symmetry. The skills for the subscale of statistics and probability included data collection, organization, display, and analysis, as well as probability and predictions. For the subscale of algebra, the skill set included attributes, patterns, and functions as well as understanding and application of algebraic concepts. In the subscale of problem solving, the skills included understanding and representing word problems, solution strategies, and logic and reasoning. These are universal subscales and are not state specific.

The reading assessment consisted of two sets of thirty-five items. The items were presented in a forced choice computerized format utilizing several different answer options. The questions were audio enhanced to increase reliability due to reading ability limitations. There were six subscales for the MPG for reading assessments. Those subscales were phonological awareness, phonics, concepts of print, vocabulary and word structure, writing, and reading comprehension. The skills covered in the subscale of phonological awareness included phoneme identification, blending, rhyming, and phonemic manipulation of sounds and syllabication. The skills in the subscale of phonics included consonants, vowel patterns, spelling patterns and rhyming, and sound manipulation and syllabication. The skills in the subscale of concepts of

print included developmental reading and writing skills as well as environmental print. The skills in subscale of vocabulary and word structure included sight words, content vocabulary and context clues, synonyms, antonyms, homonyms, homographs, and homophones, base words with prefixes and suffixes, compound words, and contractions. The skills in the subscale of comprehension included literal, interpretive, and evaluative comprehension. The skills in the final subscale of writing included writing process and conventions of language, language structure, phrase, sentence, and paragraph, and grammatical patterns. These are universal subscales and are not state specific.

The computerized adaptive tests were administered in the fall 2010 testing window and again in spring 2011 testing window. The Achievement and Growth (ASG) report was generated at the close of the spring 2011 window. A portion of the information contained in this report was used as a part of this study.

The state in which this study was conducted did not participate in the reliability testing of this particular assessment. However, the MAP Mathematics and Reading for Primary Grades assessments have established reliability and validity. NWEA uses a standard error of measurement (SEM) as a basis for reliability. "The standard error of measurement is an estimate of variability…. The standard error of measurement combines information about the test's reliability as well as the variability occurring" (Sprinthall, 2007, p. 526). The standard error of measurement provides an estimate of variance between obtained scores and the true scores (Best & Kahn, 2006).

As summarized in Figures 1 and 2 the Standard Error of Measurement (SEM) study was conducted using data from other states. The optimal fixed form assessments for kindergarten, first, and second grade are represented in the lighter dashed lines and the MAP for Primary

grades are represented in with the darker dashed lines for each grade. The SEM means maintain between a 3.0 and 4.0 for the MPG in mathematics and reading. These standard error measurement means indicated that at any point a students' true score is within 3 to 4 points of the theoretical true score. The SEM for MAP for Primary Grades is significantly different than the fixed item tests for those testing at both ends of the achievement spectrum. Students achieving in the higher Raush Unit (RIT) levels or the RIT levels maintained a SEM of 3 to 4 for MAP but for a fixed item assessment the SEM varied depending on achievement level.

Figure 1.

Mean standard errors of measurement for MAP for Primary Grades Mathematics tests in grades kindergarten through 2 and for optimal fixed form tests. (p. 175).

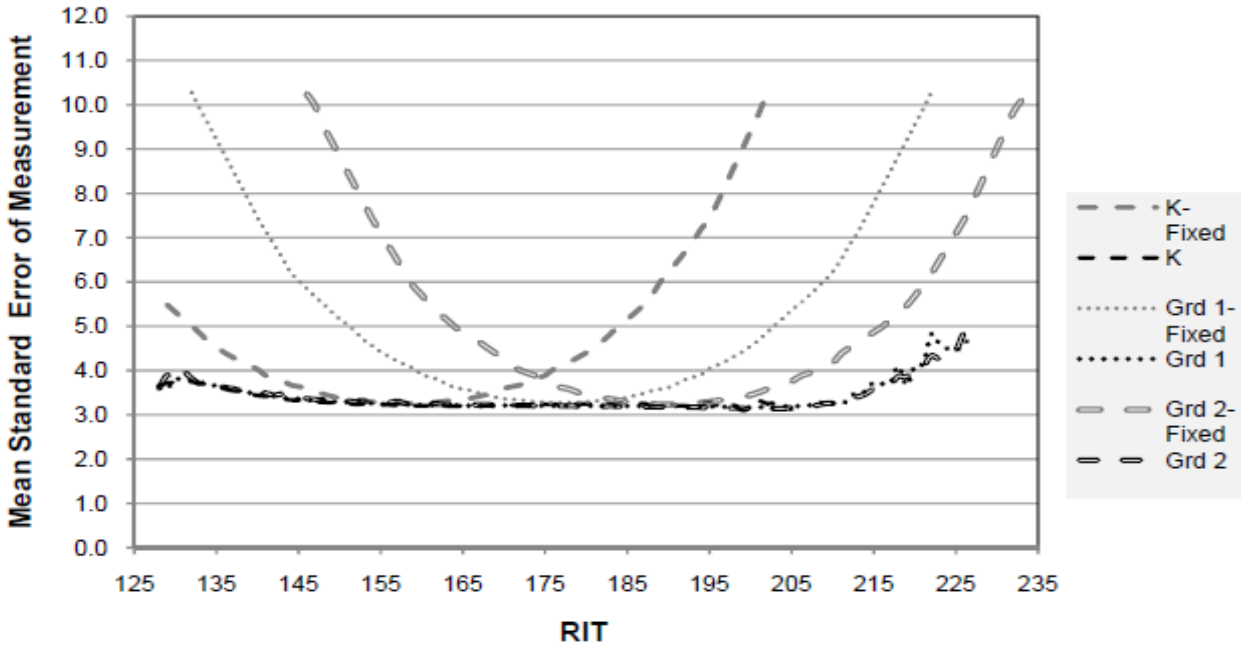
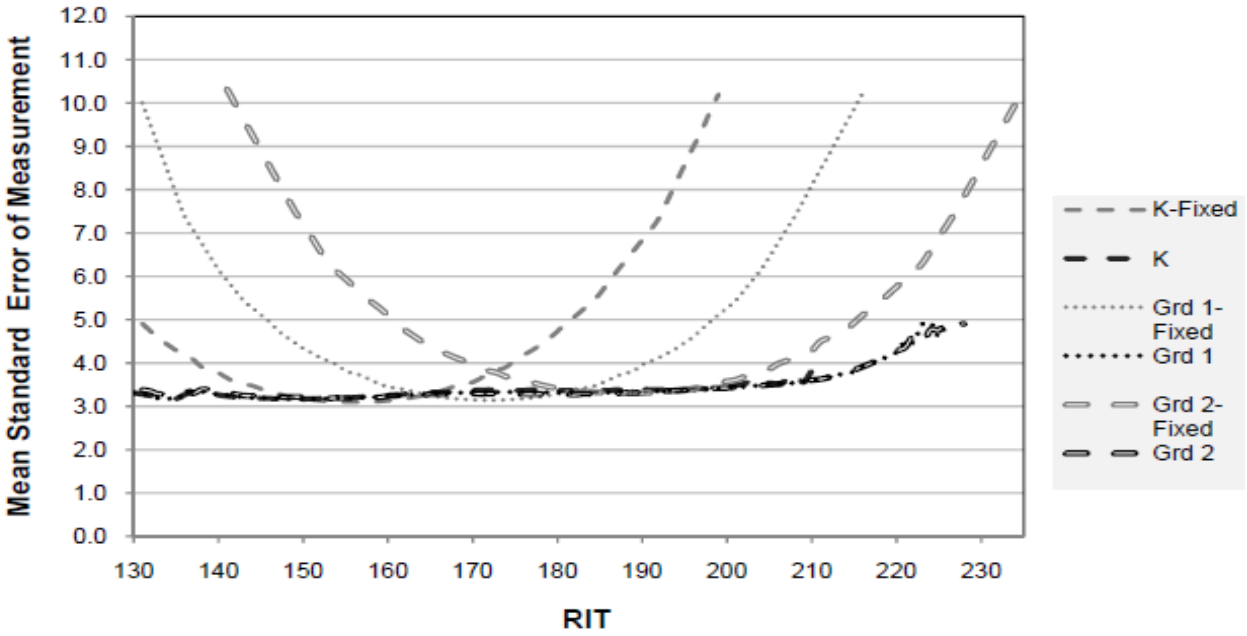


Figure 2.

Mean standard errors of measurement for MAP for Primary Grades Reading tests in grades kindergarten through 2 and for optimal fixed form tests. (p. 170).



reading for students in second through fifth grade. The assessment was created by NWEA and has been aligned with the state content standard for mathematics and reading. The purpose of this assessment was to measure the academic achievement and growth in mathematics and reading. A key reference that provides a discussion of reliability and validity for this instrument can be found in the *Technical Manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades* (2008).

The MAP assessment for mathematics consisted of fifty-two items. The items were presented in a forced four multiple choice computerized format. There were five subscales for the MAP for mathematics assessments. These subscales were number sense and computation, algebra, measurement, geometry, and data analysis. The subscale of number sense and computation assessed the student's abilities in place value including decimals, comparing and ordering of numbers, fractions, order of operations, number theory, and the basic four operations of addition, subtraction, multiplication, and division. The subscale of algebra included algebraic

expressions and equations as well as patterning. The subscale of measurement included questions relating to time, temperature, money, linear measurements, perimeter, area and volume. The geometry subscale incorporated coordinate planes, two- and three-dimensional shapes, symmetry, transformations, similarity, and congruency. The last subscale of data analysis accounted for questions concerning central tendency, probability, and data collection.

The MAP assessment for reading consisted of forty-two items that assess four subscales that were aligned with state content standards. The four subscales were word analysis, reading strategies, literary text, and expository text. The subscale of word analysis included questions concerning vocabulary development and comprehension, phonics and structural analysis. In the subscale for reading strategies included before, during, and after reading strategies. For the subscale literary text, language, mood, tone, irony, setting, plot, characterization and point of view were included. This subscale also included the ability to predict and react to literary pieces. The expository subscale included text features, prediction, inferring, and drawing conclusions.

Based on state data, reliability of the NWEA MAP mathematics and reading assessments have been established using a test/retest methodology. For math the reliability coefficients are presented in Table 1. The coefficients of reliability are all above .811. The largest N or number of pairs has a significance above .08, therefore this correlation coefficient is significant at $\alpha = .01$ level.

Table 1

Spring 2007 - Fall 2007 Test-Retest Correlations for State Content-Aligned MAP MATHEMATICS Tests With Common Item Pool Structures Administered to a Minimum of 300 Students. (p. 89)

Grade	Test	Correlation Coefficient	N
2	Math Goals Survey 2-5 NV V3	0.811	593
3	Math Goals Survey 2-5 NV V3	0.847	1016
4	Math Goals Survey 2-5 NV V3	0.868	988
5	Math Goals Survey 2-5 NV V3	0.868	1046

For reading the reliability coefficients are presented in Table 2. The coefficients of reliability were calculated and were all above 0.79 and were significant at α level = .01. The data for these tables is available in the *Technical Manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades* (2008).

Table 2

Spring 2007 - Fall 2007 Test-Retest Correlations for State Content-Aligned MAP Reading Tests With Common Item Pool Structures Administered to a Minimum of 300 Students. (p. 68)

Grade	Test	Correlation Coefficient	N
2	Reading Goals Survey 2-5 NV V3	0.791	590
3	Reading Goals Survey 2-5 NV V3	0.843	1016
4	Reading Goals Survey 2-5 NV V3	0.842	983
5	Reading Goals Survey 2-5 NV V3	0.793	451

Validity is also evident in the NWEA MAP mathematics and reading assessment. "MAP test scores can be used to classify students into performance levels for accountability purposes. For this reason, educational decisions that rely on classification accuracy and consistency are important" (Technical Manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades, 2009, p.64). The summary of results from the 2008 NWEA study for mathematics is presented in Table 3. The validity coefficients were calculated and were all at or above 0.854.

Table 3
Validity data for third through fifth grade students in MAP Mathematics test

Grade	Student RITs	Standard Deviation	Estimated State Proficiency Cut Score	N	Classification Accuracy	Decision Consistency
3	201.9	11.81	202	1859	0.900	0.854
4	210.9	12.44	209	1905	0.902	0.858
5	218.3	13.92	217	1808	0.928	0.892

The validity data is presented for reading in Table 4. The validity coefficients were calculated and all were at or above 0.854. Again these correlations show a decision consistency significant at the $\alpha = .01$ level. Due to the nature of the state's high-stakes proficiency examinations, second grade was not included in the validity measures.

Table 4

Validity data for third through fifth grade students in MAP Reading test

Grade	Student RITs	Standard Deviation	Estimated State Proficiency Cut Score	N	Classification Accuracy	Decision Consistency
3	196.7	14.90	198	1934	0.910	0.870
4	205.1	13.78	203	1947	0.906	0.864
5	210.3	13.89	213	1757	0.898	0.855

Concurrent validity for the state was established in 2003. "In concurrent validity, the test scores are correlated with an already established and accepted measure of the construct under study" (Sprinthall, 2007, p. 516). This allows test creators the ability to evaluate the relationship between the test and other criterion. The results for math and reading by grade level are presented in Table 5. Pearson *r* was calculated and the values were between 0.76 and .86 and were significant for both reading and mathematics in both third and fifth grades. For the number of participants in the study, the results indicated significance at the $\alpha = .01$. The correlations are all relatively strong. Fourth grade was not included in this study because the state of Nevada did not report fourth grade scores until 2005.

Table 5

Concurrent Validity Evidence for State Criterion Referenced Assessment scale scores and MAP scores (NWEA reliability and validity estimates, 2005).

Grade	Term	Subject	N	r
Third Grade	Spring	Mathematics	1,087	0.76
		Reading	1,084	0.82
Fifth Grade	Spring	Mathematics	1,155	.086
		Reading	1,184	0.83

Data Acquisition

The district statistician provided the data in electronic format. The data from the Staff Survey for Professional Development and the MAP/MPG was paired for analysis by the district statistician. This data was exported to an Excel spreadsheet. This spreadsheet contained the original code of the participants. Upon receipt of the data, the teacher identification codes were recoded using a random number generator. Thus any ability to connect data to individual teachers was lost. No identifiable teacher and student information was exchanged as part of the data collection.

Data Analysis

This study employed a quantitative methodology for data collection and data analysis. Descriptive statistics were computed for all the variables. These statistics included mean, standard deviations, and correlations.

The data analysis addressed five research questions. These questions are as follows:

1. Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for mathematics as measured by the Measures of Academic Progress assessment be predicted from the three independent variables (general assessment belief and MAP/MPG belief [*belief*], value of [*value*], and self-reported behaviors [*behavior*])?
2. Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for reading as measured by the Measures of Academic Progress assessment be predicted from the three independent variables (general assessment belief and MAP/MPG belief [*belief*], value of [*value*], and self-reported behaviors [*behavior*])?
3. Are there significant mean differences in attitudes and beliefs regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-

reported behaviors regarding MAP/MPG) when groups of teachers are established for certain demographics?

4. Are significant mean differences in attitudes and beliefs regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established in Title One and non-Title One designations?
5. Are there significant mean differences in attitudes and beliefs regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established by school?

For the data analysis for Question One and Two, three independent variables from the staff survey were used to compute a step-wise regression with the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for math and reading as the respective dependent variable. The independent variables were general assessment belief [*belief*], value of [*value*], and teacher self-reported behavior associated with MAP/MPG [*behavior*]. Step-wise regression was selected because the order of importance of the independent variables was considered when the model was developed (Draper & Smith, 1998). The regression analysis was conducted using appropriated programs from the Statistical Package for Social Science (SPSS).

A multiple stepwise regression is the appropriate method of analysis when the research problem involves a single dependent variable presumed to be related to multiple independent variables. In a systematic process the computer program will select the best independent variable from the set of independent variables as the first predictor variable. Then, the computer program will select the next best variable to add to the regression equation. The process is

repeated until all variables of significance are considered. As each variable is added, its contribution of new variance is considered. At each step the computer program tests for significance.

At each step in the process values for R, R^2, and adjusted R^2 are calculated. The R-value of the model indicates the multiple correlation coefficients (Draper & Smith, 1998; Moore & McCabe, 2006). The multiple correlation coefficients represent the association between the independent variables in the model with the dependent variable in the model. In multiple regression analysis, the R-value can take on a range of values from zero to one. The R^2-value then represents the amount of variation in the dependent variable that is explained by the independent variables included in the model. The R^2-value can take on a range of values from 0 to 1 with a value of one indicating that the independent variables explain all of the variation in the dependent variable.

For the mathematics related information Table 6 provides a summary of the various subscales and the corresponding items within each subscale. For the regression related to mathematics achievement the independent variables were general assessment belief [*belief*], value of [*value*], and teacher self-reported behavior associated with MAP/MPG [*behavior*]. The dependent variable was the percentage of students within a given classroom who met or exceeded their growth targets. The subscale of general assessment belief explored how the participant felt about the assessment concept in general terms and MAP/MPG. For the subscale of teacher perceived value questions focused on adequacy and accuracy of the MAP/MPG assessment as well as the professional development. For the subscale of teacher self-reported behavior associated with MAP/MPG questions focused on how the teacher utilizes the different

components of MAP/MPG and the information contained in the MAP/MPG data. Some questions were excluded from this research study and reserved for further studies.

Table 6

Mathematics Survey Items and Variables Used for Multiple Regression Analysis

Variable type	Variable Name	SSPD Survey Item Numbers	Scale with more positive scoring higher
Independent	General assessment and MAP/MPG belief	1, 2, 3, 4, 5, 6, 8, and 16 33, 34, 35, and 36	1 to 6 1 to 4
Independent	Value of assessment	10, 12, and 14	1 to 6
Independent	Teacher self-reported behavior	21, 23, 25, 27, 29, and 31	1 to 4
Dependent	% of students meeting or exceeding growth target in mathematics as based on MAP/MPG	Percentage of students in a teacher's classroom meeting or exceeding their growth targets	0% to 100%

For the reading related information, Table 7 provides a summary of the various subscales and the corresponding items within each subscale. For the regression related to reading achievement, the independent variables were general assessment belief [*belief*], value of [*value*], and teacher self-reported behavior associated with MAP/MPG [*behavior*]. The dependent variable was the percentage of students within a given classroom who met or exceeded their growth targets for reading. Some survey questions were excluded from this research study and reserved for further studies.

Table 7

Reading Survey Items and Variables Used for Multiple Regression Analysis

Variable type	Variable Name	SSPD Survey Item Numbers	Scale with more positive scoring higher
Independent	General assessment and MAP/MPG belief	1, 2, 3, 4, 5, 7, 9, and 17 33, 34, 35, and 36	1 to 6 1 to 4
Independent	Value of assessment	11, 13, and 15	1 to 6
Independent	Teacher self-reported behavior	22, 24, 26, 28, 30, and 32	1 to 4
Dependent	% of students meeting or exceeding growth target in mathematics as based on MAP/MPG	Percentage of students in a teacher's classroom meeting or exceeding their growth targets	0% to 100%

For research questions three, four, and five a multivariate analysis of variance (MANOVA) was used to test the significant difference between the designated groups. The MANOVA is a multivariate version of the univariate analysis of variance (ANOVA). The MANOVA tests differences between and among groups when the independent variables are considered. The purpose of the MANOVA is to determine if the vectors of means of the various groups are equal or unequal. For this study one-way MANOVAs were conducted. "Wilks Lamda is the most commonly reported MANOVA statistic" (Mertler & Vannatta, 2010, p.123). Homogeneity of variance-covariance using Box's test was also included. To address questions three, four, and five, groups were established using teaching experience, Title One status of teachers' schools, and teachers' schools. When significant differences were found, then appropriate post hoc analysis was conducted.

Summary

This chapter restated the purpose of this study and the research questions. The study utilized two existing data sets provided by the school district statistician. The three instruments included in this study were also included in this chapter. Reliability and validity for these three instruments was presented and discussed. Data acquisition procedure was detailed as well as the filtering and coding of the data. Finally, data analysis methodology for each research question was outlined. The intention to use a multiple regression analysis for the first two research questions was presented. The intention of MANOVA was presented to address three of the proposed research questions. The following chapter will include the results and data analysis.

CHAPTER FOUR: RESULTS

Introduction

The purpose of this study was to explore relationships between teacher attitude toward assessment and student academic achievement. The study focused on the data collected from teachers who work in six elementary schools in a school district. This study used two sets of existing data collected by the district statistician. The chapter is divided into four sections: (a) descriptive statistics, (b) data for each of the five research questions, (c) additional analysis, and (d) summary of the results.

Descriptive Statistics

The original data set for the survey results had a total of 182 participating teachers. Through examination of the data, four cases were eliminated due to univariate outliers. Six additional cases were identified as multivariate outliers and eliminated. Participants in the survey who identified themselves as special education teachers or teachers who teach English as a second language (ESL) were excluded. After cleaning the data, 148 survey results were included to study research Questions Three, Four, and Five. For research Questions One and Two, 143 survey results were utilized; for these questions it was required to match teacher survey results with corresponding student academic data.

Professional Development Survey Results.

For this study, six subscales of the District Staff Survey for Professional Development (SSPD) were analyzed: a) average math self-reported behavior, b) average math belief, c) average math value, d) average reading self-reported behavior, e) average reading belief, and f) average reading value. These subscales were discussed in detail in Chapter 3. For each response, the averages of the items for each subscale were computed and used for analysis. As a

result, the means of the subscales were approximately the same. For each subscale the mean, standard error of mean, standard deviation, minimum average score, and maximum average score are presented in Table 8.

Table 8

Descriptive Statistics for the Professional Development Survey Subscales.

	# of items	Mean	SD	Min.	Max.
Average math behavior scores	6	3.99	1.12	1.50	6.00
Average math belief scores	12	3.83	0.77	1.50	5.13
Average math value scores	3	3.30	0.87	1.00	5.33
Average read behavior scores	6	3.90	1.12	1.50	6.00
Average read belief scores	12	3.83	0.77	1.50	5.13
Average read value scores	3	3.29	0.87	1.00	5.33

The statistics for the subscales of belief, self-reported behavior, and value for the mathematics and reading were similar. The average math self-reported behavior mean was 3.99 and the average reading self-reported behavior was 3.90. The average math value mean was 3.30 and the average reading belief was 3.29. For the belief subscale math and reading both had a mean of 3.83.

Measures of Academic Progress.

Student academic growth was measured by the percentage of a teacher's class that met or exceeded the growth target as set from the fall 2010 mathematics and reading MAP/MPG scores. The descriptive statistics for the mathematics and reading MAP/MPG tests are presented in Table 9. The mean of the percentage of classes that met or exceeded reading growth targets

was 10.7% higher than the percentage of classes that met or exceeded growth targets in mathematics. The range of scores for the percentage of a class that met or exceeded growth targets varied from 8.33% to 92%, whereas the percentage of a class having met or exceeded growth targets varied from 25% to 100%. The mean percentage of classes that met or exceeded math growth targets was 48.36% which was almost 11% less than the mean of classes that met or exceeded reading targets.

Table 9

Descriptive Statistics for Reading and Mathematics Measures of Academic Progress Percentages of Students per Class that Met or Exceeded the Norm Rate of Growth.

	N	Mean	SD	Min.	Max.
Percentage of class that met or exceeded math growth targets	143	48.36	20.23	8.33	92.00
Percentage of class that met or exceeded reading growth targets	143	59.06	15.24	25.00	100.00

Data Analysis for Question One

Research Question One was: Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for mathematics as measured by the Measures of Academic Progress be predicted from three independent variables (general assessment belief and MAP/MPG belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*])? The descriptive statistics for the variables related to this question are displayed in Table 10.

Table 10

Descriptive Statistics for the Mathematics Measures of Academic Progress and the Professional Development Survey Mathematics Subscales.

	N	Mean	Std. Deviation
Percentage of class that met or exceeded math growth targets	143	48.36	20.23
Average math belief [*belief]*	143	3.82	0.78
Average math value [*value*]	143	3.28	0.88
Average math behavior [*behavior*]	143	3.98	1.13

The mean of the class percentages that met or exceeded growth targets in mathematics was 48.36%. The variables of average math belief [*belief*] (M = 3. 82) and average math self-reported behavior [*behavior*] (M = 3.98) varied 0.16, whereas average math value [*value*] (M = 3.28) varied from 0.54 and 0.7 respectively. Teachers rated themselves higher on survey questions related to their self-reported behaviors of the math assessment than on questions related to the value of the math assessment.

The Pearson correlation (r) was used to measure the degree of association between each of these variables and the percentage of class that met or exceeded growth targets. The computed correlations are all significant and are presented in Table 11. Thus, the data suggested that linear relationships exist among all variables: average math belief, average math value, average math self-reported behavior, and percentages of a teacher's class meeting or exceeding mathematic growth targets. The association of percentage of class that met or exceeded math growth targets and average math belief scores of teachers was strong and positive. The association of percentage of class that met or exceeded math growth targets and average math value scores of teachers was strong and positive. The association of percentage of class that met or exceeded

math growth targets and average math self-reported behavior scores of teachers was strong and positive. In other words, teachers' positive responses on the survey were related to higher percentages of students having met or exceeded math growth targets.

Table 11

Correlation Coefficients for Math Subscale for Survey Variables

	% of class that met or exceeded math growth targets	Ave. math belief	Ave. math value	Ave. math behavior
% of class that met or exceeded math growth targets	1.000			
Ave. math belief	0.24**	1.00		
Ave. math value	0.16*	0.52***	1.000	
Ave. math behavior	0.24**	0.72***	0.24***	1.000

Note: No. of participants= 143

*** Correlation is significant at the 0.001 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

For math, a stepwise multiple regression was conducted to determine if a predictive model could be developed when the independent variables (general assessment belief and MAP/MPG belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*]) were used to predict the percentage of students in a class meeting or exceeding the norm rate of growth on the mathematics Measures of Academic Progress. Regression results indicated that the overall model significantly predicted percentage of a class meeting mathematics norm growth, R^2 = 0.058, R^2_{adj} = 0.052, $F(1,141) = 8.736$, p=0.004. However, this model accounted for only 5.8% of variance in the percentages of teachers' classes meeting or exceeding norm rate of growth in mathematics. The only variable that entered into the step-wise regression model was the average math self-reported behavior [*behavior*], p = 0.004. When the results from Table 11 and Table 12 are considered jointly, the results are clearer. The correlation between percentages of

teachers' classes meeting or exceeding norm rate of growth in mathematics and math self-reported behaviors of the teachers was 0.24 which when squared provides the values of 0.058 (5.8% variance accounted for). In addition, given the high correlations among the three independent variables, the results suggest that the shared variance with the dependent variable was accounted for by one independent variable. Thus, the other independent variables failed to add a new variance accounted for in the model.

Table 12

Summary Statistics for Step-wise Regression Model for Coefficient to Predict Math Achievement.

	B	*β*	*t*	*p*
Average math behavior	4.34	0.24	2.96	0.004

Data Analysis for Question Two

Research Question Two was: Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for reading as measured by the Measures of Academic Progress be predicted from three independent variables (general assessment belief and MAP/MPG belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*])? The descriptive statistics for this question are displayed in Table 13.

Table 13

Correlation Coefficients for Reading Subscale for Survey Variables.

	N	Mean	Std. Deviation
Percentage of class that met or exceeded reading growth targets	143	59.06	15.24
Average read belief scores	143	3.82	0.78
Average read value scores	143	3.28	0.87
Average read behavior scores	143	3.89	1.13

For the dependent variable of the percentage of class that met or exceeded reading growth targets, the mean was 59%. The teachers rated survey items related to their self-reported behaviors associated with the reading MAP/MPG assessment higher than the rest of the reading assessment related items on the survey. The mean for the reading self-reported behavior subscale was 3.89. The teachers tended to score survey items related to the beliefs toward the reading MAP/MPG assessment higher than the survey items related to the value of the reading MAP/MPG assessment. The mean for the subscale of belief was 3.82 whereas the mean for the value subscale was 3.28.

The Pearson correlation coefficient (r) was used to measure the degree of association between each of the variables and the percentage of class that met or exceeded growth targets. The correlation matrix is presented in Table 14. The reading subscale and the math subscale correlations presented in Table 11 are very similar. The difference is between the independent variable and dependent variable. The correlations between the dependent variable and the independent variables were not significant. That is, the respective correlations between reading belief, reading value, and reading self-reported behavior, the percentage of a class percentage that met or exceeded growth targets were not significant.

Table 14

Correlation Coefficients for Reading Subscale for Survey Variables.

	% of class that met or exceeded reading growth targets	Ave. math belief	Ave. math value	Ave. math behavior
% of class that met or exceeded reading growth targets	1.00			
Ave. math belief	0.03	1.00		
Ave. math value	0.11	0.53***	1.00	
Ave. math behavior	-0.08	0.73***	0.39***	1.00

Note: No. of participants= 143 *** Correlation is significant at the 0.001 level (2-tailed).

A stepwise multiple regression was conducted to determine if a predictive model could be developed when the independent variables (general assessment belief and MAP/MPG belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*]) were used to predict the percentage of students in a class meeting or exceeding the norm rate of growth on the reading Measures of Academic Progress. Regression results indicated that the overall model was not significant and did not predict the percentage of a class meeting reading norm growth. No variables entered into the regression model.

Data Analysis for Question Three

Question Three of this research study was: Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP /MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established based on selected demographics? The demographic variables used for this analysis were the education level of the teachers (Bachelor's, Bachelor's plus additional credits, Masters, Masters plus additional credits, and Doctorate), years of experience (Before Vested, Raise Based on Service, and Longevity), and grade level taught by the participants (kindergarten, first grade, second grade, third grade, fourth grade, and fifth grade). Each analysis was conducted utilizing a multivariate analysis of variance (MANOVA). Each individual analysis is discussed in separate sections.

As discussed above, data screening was performed on the original 182 district survey participants. Ten responses were identified as univariate or multivariate outliers; therefore, they were eliminated. Twenty four responses were excluded because they did not meet the requirement of teaching in a regular education classroom. As a result, the number of responses used for these analyses was 148.

Education Level Analysis.

To determine the educational level each respondent was asked to indicate the highest level of education completed from five categories: Bachelor's, Bachelor's Plus, Masters, Masters plus, Doctorate. The results are summarized in Table 15. The majority of the teachers have a Bachelor's degree with additional college credit (76 of 148). In the district where the study was conducted, a teacher can advance all the way across the pay scale using additional educational college credits without completing a Masters. None of the teachers indicated that they had completed a doctorate at the time of this study.

Table 15

Number of Teachers by Educational Level Category.

	Bachelor's	Bachelor's +	Masters	Masters +	Total
N	26	76	19	27	148

The survey subscales in this analysis are the same as the variables utilized in the two regression analyses. The six variables used in this analysis were the average scores in six categories: math belief, math self-reported behavior, math value, reading belief, reading self-reported behavior, and reading value. Descriptive statistics for the education level of the teachers corresponding to each of the subcategories of the Professional Development Survey are presented in Table 16 and Table 17.

Table 16

Group Means and Standard Deviations for Education Level for Math Subscales from the Professional Development Survey.

		Ave. math belief		Ave. math behavior		Ave. math value	
Grade	N	M	SD	M	SD	M	SD
Bachelor's	26	3.93	0.76	3.83	1.33	3.13	0.98
Bachelor's +	76	3.73	0.79	3.85	1.10	3.32	0.91
Masters	19	3.84	0.71	4.52	0.79	3.35	0.84
Masters +	27	4.00	0.74	4.25	1.06	3.38	0.69
Total	148	3.83	0.77	3.99	1.12	3.30	0.87

Table 17

Group Means and Standard Deviations for Education Level for Reading Subscales from the Professional Development Survey.

		Ave. read belief		Ave. read behavior		Ave. read value	
Grade	N	M	SD	M	SD	M	SD
Bachelor's	26	3.91	0.78	3.68	1.29	3.19	0.99
Bachelor's +	76	3.73	0.80	3.78	1.13	3.29	0.91
Masters	19	3.85	0.71	4.42	0.86	3.35	0.83
Masters +	27	3.98	0.73	4.06	1.03	3.35	0.64
Total	148	3.82	0.77	3.90	1.12	3.29	0.87

For educational level, a one-way multivariate analysis of variance (MANOVA) was conducted to test for group differences on the survey results. The summary statistics are Wilks' $\Lambda = 0.818$, $F(18, 394) = 1.607$, $p = 0.055$, multivariate $\eta^2 = 0.065$. The MANOVA results revealed no significant differences among the groups established by the education levels when the subscales were considered. As a result, additional post hoc analyses were not performed.

Length of Teaching Experience.

An additional analysis was conducted for the demographic of teaching experience. To determine experience level, each respondent was asked to indicate the number of years that he/she had been in the teaching profession. For the purpose of this analysis, the data was aggregated into three groups: Before Vested, Raise Based on Service, and Longevity. "Before Vested" is the category that included all teachers with four or less years of teaching. This category is for teachers who are not vested in the state retirement system. The second category was named "Raise Based on Service". This category included all teachers who have from five years through fifteen years of experience. In the district where this study was conducted, salary increases are based on the year of service up until the end of the fifteenth year. The final group of teachers was "Longevity". These teachers do not receive yearly salary increases based on service but rather they received stipends based on longevity. Teachers who begin their sixteenth year are given one lump sum at the end of each year. The amount increases at five year increments. The number of teachers in each category is presented in Table 18. There were very few new teachers at the time the survey was administered. The other two categories of experience were fairly equal. This third category distribution was the most equitable way to split the years of service due to a wide range of responses. The range in teaching experience in this district was 41 years with brand new teachers and one who has taught 42 years.

Table 18

Number of Teachers Group Related to the Years of Experience Category.

	Before Vested	Raise Based on Service	Longevity	Total
N	13	65	70	148

The survey subscales used in this analysis were the same variables utilized in the two regression analyses. The six variables used in this analysis are the average scores in the following six categories: math belief, math self-reported behavior, math value, reading belief, reading self-reported behavior, and reading value. Descriptive statistics for the length of experience of the teachers corresponding to each of the subcategories of the Professional Development Survey are presented in Table 19 and Table 20. It was noted that mean responses decreased as the experience level in years increased.

Table 19

Group Means and Standard Deviations for Experience Level for Math Subscales from the Professional Development Survey.

		Ave. math belief		Ave. math behavior		Ave. math value	
Grade	N	M	SD	M	SD	M	SD
Before Vested	13	4.15	0.56	4.12	0.57	3.67	0.75
Raise based on Service	65	3.91	0.75	3.90	0.75	3.27	0.83
Longevity	70	3.70	0.81	3.70	0.93	3.26	0.92
Total	148	3.83	0.77	3.82	0.77	3.30	0.87

Table 20

Group Means and Standard Deviations for Experience Level for Reading Subscales from the Professional Development Survey.

		Ave. read belief		Ave. read behavior		Ave. read value	
Grade	N	M	SD	M	SD	M	SD
Before Vested	13	3.72	0.68	4.37	0.62	4.31	0.74
Raise based on Service	65	3.27	0.83	4.00	1.12	3.92	1.14
Longevity	70	3.23	0.92	3.91	1.19	3.80	1.16
Total	148	3.97	0.87	3.99	1.12	3.89	1.12

For years of teaching experience, a one-way multivariate analysis of variance (MANOVA) was conducted to test for group differences on the survey results. The summary statistics for the MANOVA are Wilks' $\Lambda = 0.907$, $F(12, 280) = 1.169$, $p = 0.305$, multivariate $\eta^2 = 0.048$. The MANOVA results revealed no significant differences among the groups established by the education levels when the subscales were considered. As a result, additional post hoc analyses were not performed.

Grade Level Taught by the Teacher.

An additional analysis was conducted for the demographic of grade level taught by each teacher. For the grade level taught to be determined each survey respondent was asked to indicate the grade level that best described the grade level taught for the 2010-2011 school year. Each grade level from kindergarten through fifth grade was used for this analysis. The number of teachers in each category is presented in Table 21. The majority of the teachers in this school district reported teaching in one of the primary grades of kindergarten, first grade, or second grade (81 of 148).

Table 21

Number of Teachers Grouped by Grade Level Taught.

	Kinder	First Grade	Second Grade	Third Grade	Fourth Grade	Fifth Grade
N	14	35	32	25	23	19

The survey subscales used in this analysis match the variables utilized in the regression analysis. The six variables used in this analysis are the average scores in the following six categories: math belief, math self-reported behavior, math value, reading belief, reading self-reported behavior, and reading value. Table 22 and Table 23 present the means and standard deviations for each of the survey variables by grade level taught. It is noted that for the majority of variables, the scores increased as the grade level taught increased.

Table 22

Group Means and Standard Deviations for Grade Level Taught for Math Subscales from the Professional Development Survey.

		Ave. math belief		Ave. math behavior		Ave. math value	
Grade	N	M	SD	M	SD	M	SD
Kinder	14	3.18	0.97	3.27	1.23	2.69	1.14
First Grade	35	3.47	0.73	3.24	1.14	3.14	0.76
Second Grade	32	3.81	0.75	3.99	0.97	3.29	0.87
Third Grade	25	4.12	0.52	4.41	0.79	3.51	0.80
Fourth Grade	23	4.18	0.62	4.37	0.74	3.42	0.76
Fifth Grade	19	4.20	0.6	4.91	0.85	3.63	0.89
Total	148	3.83	0.77	3.99	1.12	3.30	0.87

Table 23

Group Means and Standard Deviations for Grade Level Taught for Reading Subscales from the Professional Development Survey.

		Ave. read belief		Ave. read behavior		Ave. read value	
Grade	N	M	SD	M	SD	M	SD
Kinder	14	3.19	0.96	3.25	1.22	2.72	1.15
First Grade	35	3.46	0.73	3.09	1.16	3.17	0.74
Second Grade	32	3.81	0.77	3.87	0.99	3.28	0.83
Third Grade	25	4.12	0.53	4.34	0.72	3.43	0.83
Fourth Grade	23	4.16	0.62	4.30	0.76	3.42	0.82
Fifth Grade	19	4.19	0.59	4.83	0.81	3.61	0.88
Total	148	3.82	0.77	3.90	1.12	3.29	0.87

For grade level taught by the teacher, a one-way multivariate analysis of variance (MANOVA) was conducted to test for group differences on the survey results. The summary statistics for the MANOVA are Wilks' $\Lambda = 0.609$, $F(30, 550) = 2.421$, $p < 0.001$, multivariate $\eta^2 = 0.094$. The MANOVA results revealed significant differences among the groups established by the grade levels taught by the teachers. As a result, post hoc analyses were conducted.

An analysis of variance (ANOVA) was conducted on each dependent variable as a follow up test to MANOVA. Grade levels taught by the teacher were significant for the variable of average math belief, $F(5,142) = 7.421$, $p < 0.001$, partial $\eta^2 = 0.207$. The mean for the average math belief subscale was 3.83. Differences for average reading belief were also significant, $F(5,142) = 7.121$, $p < 0.001$, partial $\eta^2 = 0.200$. The mean for the average reading belief subscale was 3.82. Another significant difference was for average math self-reported behavior, $F(5,142) = 10.189$, $p < 0.001$, partial $\eta^2 = 0.276$. The mean for the average math self-reported behavior

subscale was 3.99. A significant difference for average reading self-reported behavior was also indicated, $F(5,142) = 10.749$, $p < 0.001$, partial $\eta^2 = 0.289$. The mean for the average reading self-reported behavior subscale was 3.90. Differences for the variable average math value were also significant, $F(5,142) = 2.641$, $p = 0.026$, partial $\eta^2 = 0.085$. The mean for the average math value subscale was 3.30. However, differences in average reading value were not significant, $F(5,142) = 2.208$, $p = 0.057$, partial $\eta^2 = 0.072$.

Additionally, multiple t-tests were conducted and the Bonferroni post hoc correction was used. The Bonferroni correction is a procedure that limits the probability of a Type I error when multiple comparisons are used. "Pairwise comparisons control the familywise error by correcting the level of significance for each test such that the overall Type I error rate (α) across all comparisons remains at .05" (Field, 2005, p. 339). This correction ensures that when multiple tests are conducted the overall level of significance is maintained.

The pairwise comparisons revealed many significant differences. For the variable of average math belief, kindergarten (M = 3.18) was significantly different from third (M = 4.12, p = 0.001), fourth (M = 4.18, p = 0.001), and fifth (M = 4.20, p = 0.001) grades. A similar finding occurred for first grade (M = 3.47). The third grade mean for the math belief subscale was 4.12 and p = 0.007. The fourth grade mean for the math belief subscale was 4.18 and p = 0.003. The fifth grade mean for the math belief subscale was 4.20 and p = 0.005.

For the variable average reading belief, kindergarten (M = 3.19) was significantly different from third (M = 4.12, p = 0.002), fourth (M = 4.16, p = 0.001), and fifth (M = 4.19, p = 0.001) grades. A similar finding occurred for first grade (M = 3.46) for the reading belief variables. The third grade mean for the reading belief subscale mean was 4.12 and p = 0.007.

The fourth grade mean for the reading belief subscale mean was 4.16 and $p = 0.004$. The fifth grade mean for the reading belief subscale mean was 4.19 and $p = 0.006$.

For the variable of average reading self-reported behavior, kindergarten ($M = 3.25$) differed significantly from third grade ($M = 4.34$, $p = 0.014$), fourth grade ($M = 4.30$, $p = 0.23$), and fifth grade ($M = 4.83$, $p < 0.001$); just as first grade ($M = 3.09$) differed with third ($p < 0.001$), fourth ($p < 0.001$), and fifth ($p < 0.001$) grades on the reading self-reported behavior variable.

For the variable of average math self-reported behavior, the kindergarten ($M = 3.27$) significantly differed from the third ($M = 4.41$, $p = 0.009$) and fifth ($M = 4.91$, $p = 0.016$) grade; whereas the first grade ($M = 3.24$) differed significantly from third ($M = 4.41$, $p < 0.000$), fourth ($M = 4.37$, $p < 0.001$), and fifth ($M = 4.91$, $p < 0.001$) grades.

In summary, kindergarten and first grade means on the subscales tended to be significantly different from the corresponding means of the upper grade means. Second grade did not factor into any of significant differences of means for any of the subscales.

Data Analysis for Question Four

Question Four of this research study was: Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/ MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established in Title One and non-Title One designations? Title One status is determined by the free and reduced lunch status of the students in each school. The school district where this study was conducted has four schools that are designated as Title One schools and two that are not designated as Title One schools (Non-Title One).

To determine Title One status each respondent was asked to indicate the school where he/she taught; this information was coded by the district statistician prior to transmission of data. The schools were coded as a 1 for "Title" and a 2 for "Non-title". The frequencies of responses by Title Status are presented in Table 24. The majority of the teachers in this school district teach at a Title One designated school; that is, 68.24% of the teachers work in Title One schools. This statistic is parallel to the percentage of the schools that are Title One schools (66.6%).

Table 24

Number of Teachers Grouped by Title One Status.

	Title 1	Non-Title 1	Total
N	101	47	148

The survey subscales in this analysis are the same as the variables utilized in the two regression analyses. The six variables used in this analysis were the average scores in six categories: math belief, math self-reported behavior, math value, reading belief, reading self-reported behavior, and reading value. Descriptive statistics for the Title One status of teachers' schools corresponding to each of the subcategories of the Professional Development Survey are presented in Table 25 and Table 26. The mean for the subscale average math belief was 3.83 for both Title One and Non-Title One teachers. For the subscale of average reading belief the means were 3.83 and 3.82 for Title One and Non-Title One teachers respectively. For the subscales of average math and reading self-reported behavior, Title One teachers scored the questions higher with means of 4.04 and 3.85 respectively than the Non-Title One teachers with means of 3.89 and 3.85 respectively. For the subscale of average math value, the mean score was 3.26 for Title One and 3.38 for Non-Title One teachers. For the subscale of average reading value, the mean scores were 3.17 for Title One teachers and 2.72 for Non-Title One teachers.

Table 25

Group Means and Standard Deviations for Title One Status for Math Subscales from the Professional Development Survey.

	Ave. math belief		Ave. math behavior		Ave. math value	
Status	M	SD	M	SD	M	SD
Non-Title	3.83	0.77	3.89	1.15	3.38	0.78
Title One	3.83	0.78	4.04	1.11	3.26	0.91
Total	3.83	0.77	3.99	1.12	3.30	0.87

Table 26

Group Means and Standard Deviations for Title One Status for Reading Subscales from the Professional Development Survey.

	Ave. read belief		Ave. read behavior		Ave. read value	
Status	M	SD	M	SD	M	SD
Non-Title	3.82	0.77	3.85	1.14	2.72	1.15
Title One	3.83	0.78	3.92	1.12	3.17	0.74
Total	3.82	0.77	3.90	1.12	3.29	0.87

For Title One status a one-way multivariate analysis of variance (MANOVA) was conducted to test for group differences on the survey results. The summary statistics are Pillai's Trace V = 0.042, $F(6, 141) = 1.026$, $p = 0.411$, multivariate $\eta^2 = 0.042$. Pillai's Trace was utilized because the Box's Test was significant at $p < 0.001$. "Typically, if Box's Test was significant at $p < 0.001$ and the group populations were extremely unequal then robustness cannot be assumed due to unequal variances among groups. In such a situation, a more robust MANOVA test statistic, Pillai's Trace, is utilized when interpreting the MANOVA results" (Mertler & Vannatta, 2010, p. 124). The MANOVA results revealed no significant differences

among the groups established by the Title One status of a school when the subscales were considered. As a result, additional post hoc analyses were not performed.

Data Analysis for Question Five

Question Five of this research study was as follows: Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established by school?

To determine the school, each respondent was asked to indicate the school in which he/she was employed for the 2010-2011 school year. These results were then coded A-F by the district statistician prior to data transmission. The frequencies by school are presented in Table 27.

Table 27
Number of Teachers Grouped by School.

	School A	School B	School C	School D	School E	School F	Total
N	26	21	26	30	21	24	148

The survey subscales in this analysis were the same as the variables utilized in the two regression analyses. The six variables used in this analysis were the average scores in six categories: math belief, math self-reported behavior, math value, reading belief, reading self-reported behavior, and reading value.

It was noted that School D had higher means than the other schools for the variables of math belief and math value. School D had higher means for the all schools with the exception of School F for the variable of math value. For reading, School D had higher means on all reading variables when compared to all the other schools. School E had the lowest means for all schools

for all three of the mathematics related variables as well as all the reading variables. In general, the mathematics variables of belief, self-reported behavior, and value had higher means than the corresponding subscales for reading. The summary statistics for the school corresponding to each of the subcategories of the Professional Development Survey are presented in Table 28 and Table 29.

Table 28

Group Means and Standard Deviations for School for Math Subscales from the Professional Development Survey.

		Ave. math belief		Ave. math behavior		Ave. math value	
Grade	N	M	SD	M	SD	M	SD
School A	26	3.93	0.66	3.92	1.05	3.50	0.66
School B	21	3.69	0.88	3.84	0.89	3.22	0.90
School C	26	4.01	0.58	4.00	0.74	3.32	0.91
School D	30	4.10	0.74	4.69	0.83	3.37	0.79
School E	21	3.40	0.81	3.39	1.46	2.90	1.03
School F	24	3.83	0.80	3.83	1.03	3.39	0.94
Total	148	3.83	0.77	3.99	1.12	3.30	0.87

Table 29

Group Means and Standard Deviations for School for Reading Subscales from the Professional Development Survey.

		Ave. read belief		Ave. read behavior		Ave. read value	
Grade	N	M	SD	M	SD	M	SD
School A	26	3.95	0.66	3.91	1.05	3.47	0.73
School B	21	3.69	0.90	3.76	1.28	3.85	0.89
School C	26	3.99	0.91	3.90	0.74	4.00	0.87
School D	30	4.12	0.79	4.63	0.89	4.69	0.77
School E	21	3.39	1.03	3.19	1.39	3.39	1.00
School F	24	3.66	0.80	3.70	1.09	3.83	0.92
Total	148	3.82	0.77	3.90	1.12	3.29	0.87

For school groups, a one-way multivariate analysis of variance (MANOVA) was conducted to test for group differences on the survey results. The summary statistics are Wilks' $\Lambda = 0.69$, $F(30, 550) = 1.784$, $p = 0.007$, multivariate $\eta^2 = 0.072$. The MANOVA results revealed significant differences among the groups established by school building and the six professional development survey categories.

Analysis of variance (ANOVA) was conducted on each dependent variable as a follow up test to MANOVA. The school groups were significant for the variable of average reading belief, $F(5,142) = 3.181$, $p = 0.009$, partial $\eta^2 = 0.101$. Differences for average math self-reported behavior were also significant, $F(5,142) = 4.122$, $p = 0.002$, partial $\eta^2 = 0.127$. Another difference that was significant was for average math belief, $F(5,142) = 3.181$, $p = 0.013$, partial $\eta^2 = 0.096$. A final significant difference for average reading self-reported behavior was also indicated, $F(5,142) = 5.545$, $p < 0.001$, partial $\eta^2 = 0.149$. Differences in average math value

were not significant, $F(5,142) = 1.265$, $p = 0.282$, partial $\eta^2 = 0.043$. Likewise, differences in average reading value were not significant, $F(5,142) = 1.402$, $p = 0.227$, partial $\eta^2 = 0.047$.

Additionally, multiple t-tests were conducted and the Bonferroni post hoc correction was applied. Several pairwise differences were revealed. The variable of average math belief School D (M=4.10) differed significantly from School E, (M = 3.40, p = 0.017). The variable of average reading belief School D (M = 4.12) differed significantly from School E, (M = 3.39, p = 0.012). The variable of average math self-reported behavior School D (M = 4.69) was again significantly different from School E (M = 3.39, p = 0.001). The variable of average reading self-reported behavior School D (M = 4.63) was significantly different from School E, (M = 3.19, $p < 0.001$), and School F, (M = 3.70, p = 0.025)

Additional Analyses

Additional regression analyses were conducted to explore possible relationships between the dependent variables and the independent demographic variables. The first additional analysis was a standard multiple regression. For this analysis the dependent variable was the percentage of a teacher's class who met or exceeded norm rate of growth for mathematics. The independent variables included: the sum of the averages for the Professional Development Survey, school, Title One designation, grouped grade level, experience level, and educational level of the teacher. For this regression analysis, the following variables were dummy coded: school, experience level, educational level, and grade level taught. "Dummy coding is a way of representing groups of people using only zeros and ones" (Fields, 2005, p. 208). Dummy coding allows categorical data to be utilized in statistical analysis such as regression (Fields, 2005; Draper & Smith, 1998).

For the variable of school, responses were classified using dummy coding. For experience level of the teacher, responses were classified using dummy coding as either "1" for teachers with four years or less of teaching experience, "2" for teachers who have five through fifteen years of teaching experience, and "3" for those who have taught sixteen or more years. For educational level, responses were either dummy coded as a "1" for teachers who had completed a Bachelor's degree only, a "2" for teachers who had completed a Bachelor's plus additional course work, a "3" for teachers who had completed a Masters only, and a "4" for teachers who had complete a Masters plus additional course work. For grade level taught, responses were dummy coded as either "1" for primary or "2" for intermediate. Primary grades included kindergarten, first grade, and second grade. Intermediate grades included third grade, fourth grade, and fifth grade.

A standard multiple regression was conducted using independent variables (total sum of the average professional development survey [*survey*], school [*school A, school B, school C, school D, school E, school F*], experience level [*before vested, raise based on service, longevity*], education level [*bachelor's, bachelor's plus, masters, masters plus*], grade grouping [*before vested, raise based on service, longevity*], and Title One status [*title, non-title*]) for predicting the percentage of students in a class meeting or exceeding the norm rate of growth on the mathematics Measures of Academic Progress. Regression results indicated that the overall model significantly predicts percentage of a class meeting mathematics norm growth, $R^2 = 0.232$, $R^2_{adj} = 0.167$, $F(12,141) = 3.548$, $p < 0.001$. This model accounted for 23.2% of variance in the percentages of teachers' classes meeting or exceeding norm rate of growth in mathematics. A summary of the regression coefficient is presented in Table 30 and indicated that no variables significantly contributed to the model.

Table 30

Summary Statistics for Step-wise Regression model for Coefficient for the Standard Regression Model to Predict the Percentage of Class Meeting or Exceeding Mathematics Growth Targets.

Variable	*B*	*B*	*t*	*P*
School B	-10.26	-0.18	-1.95	0.05*
School C	7.53	0.14	1.43	0.16
School D	1.25	0.02	0.25	0.80
School F	-4.47	-0.09	-0.89	0.37
zero to five years	-1.67	-0.02	-0.30	0.76
Fifteen years plus	0.34	0.01	0.11	0.92
Masters plus	-1.18	-0.02	-0.29	0.78
Masters	-0.73	-0.01	-0.15	0.88
Bachelor's	1.05	0.02	0.25	0.80
Primary grades	-12.82	-0.32	-0.95	0.34
Intermediate Grades	5.13	0.13	0.38	0.70
Title status	8.06	0.19	1.59	0.12

* Significant at the 0.05 level (2-tailed).

A similar analysis of the above regression model using the dependent variable of the percentages of a class meeting or exceeding reading norm rate of growth was also calculated. This analysis was not significant.

Additional analyses were calculated utilizing the MANOVA statistic. These analyses replaced the survey variables as the dependent variables. These analyses were calculated using the percentage of a class meeting or exceeding in both reading and mathematics as the dependent variables and keeping all other independent variables the same as presented in Questions Three, Four, and Five. Only one analysis proved significant: grade level taught.

This analysis utilized percentage of classes that met or exceeded reading growth targets and percentage of classes that met or exceeded mathematics growth targets as dependent variables. The fixed variables were the grade level taught. Table 31 presents means and standard deviations for each of the academic variables by grade level taught. It was noted that first grade had the highest percentage of students meeting or exceeding reading growth targets of all the grades; however, first grade had the second lowest percentage of classrooms meeting or exceeding mathematics growth targets. Second grade classrooms had the lowest percentage of classes meeting or exceeding both reading and mathematics. Fourth grade had the highest percentage of classes meeting or exceeding reading growth targets.

With the exception of fifth grade, the percentage of classes meeting or exceeding growth targets in reading was higher than in mathematics. The mean kindergarten class percentage of those students who met or exceeded the math growth target was 45.95%; however, the mean kindergarten class percentage of those students who met or exceeded for reading growth targets was 59.06%. The mean first grade class was 44.44%, but the mean first grade class percentage of those students who met or exceeded for reading growth targets was 65.69%. The mean second grade class percentage of those students who met or exceeded the math growth target was 32.66%, whereas the mean second grade class percentage for those students who met or exceeded for reading growth targets was 52.50%. The mean third grade class percentage for those students who met or exceeded the math growth target was 54.29%, but the mean third grade class percentage for those students who met or exceeded for reading growth targets was 57.45%. For the intermediate grades of fourth and fifth, the opposite was true. The mean fourth grade class percentage for those students who met or exceeded the reading growth target was 53.76%, however the mean fourth grade class percentage for those students who met or exceeded

for reading growth targets was 58.95%. The opposite is true for fifth grade. The mean fifth grade class percentage for those students who met or exceeded the math growth target was 67.14%; whereas the mean fifth grade class percentage for those students who met or exceeded for reading growth targets was 56.98%.

Table 31

Group Means and Standard Deviations for Grade Level Taught for the Measures of Academic Progress for Mathematics and Reading Assessments.

		% meeting reading		% meeting mathematics	
Grade	N	M	SD	M	SD
Kinder	13	59.06	20.28	45.95	17.24
First Grade	35	65.69	19.12	44.44	19.42
Second Grade	29	52.50	11.28	32.66	17.13
Third Grade	25	57.45	14.79	54.29	18.10
Fourth Grade	22	58.95	8.44	53.76	14.19
Fifth Grade	19	56.98	11.18	67.14	17.34
Total	148	59.06	15.23	48.36	20.22

A two-way multivariate analysis of variance (MANOVA) was conducted to determine group differences between academic performance variables and the grade level taught by the teacher. The summary statistics are Wilks' $\Lambda = 0.611$, $F(10, 272) = 7.601$, $p < 0.001$, multivariate $\eta^2 = 0.218$. MANOVA results revealed significant differences among the grade levels taught and the two areas of academic growth. As a post hoc procedure one way analysis of variance (ANOVA) was conducted on each dependent variable as a follow up test to MANOVA. Grade level taught by the teacher was significant for the variable of percentage of class that met or exceeded math growth targets, $F(5,137) = 10.399$, $p < 0.001$, partial $\eta^2 = 0.275$.

Differences in the variable of percentage of class that met or exceeded reading growth targets were significant, $F(5,137) = 2.782$, $p = 0.020$, partial $\eta^2 = 0.092$.

Additionally, multiple t-tests were conducted and the Bonferroni post hoc correction was used. For the variable of percentage of class that met or exceeded reading growth target first grade (M = 65.09) differed significantly from second grade, M = 52.50 with p = 0.009. For the variable of percentage of class that met or exceeded mathematics growth target, kindergarten (M = 45.95%) differed significantly from fifth grade, M = 67.14 with p = 0.15 as did first grade (M = 44.44) from fifth grade, p < 0.001, second grade (M = 32.66) differed significantly from third grade, M = 54.29 with p < 0.001, fourth grade, M= 53.76 with p = 0.001, and fifth grade, p < 0.001.

Summary

In this chapter, the analysis related to each of the five research questions was presented as well as additional data analyses. Data screening procedures were outlined for each question. Descriptive statistics were given for the overall sets of data as well as for individual research questions. The analyses associated with significant results are summarized below.

For research Question One results revealed a significant regression model that included only one variable. This regression model only accounted for 5.8% of variance in the percentage of a class that met or exceeded growth targets. The third research question was addressed in three different analyses. The third demographic analyzed in Question Three did produce significant differences among teachers at different grade levels. In general, means for primary grades teachers' students differed from the means for intermediate grades teachers' students. Question Five utilized a MANOVA statistic and revealed significant differences in survey results

when disaggregated by school site. Several pairwise comparisons were significant between the different schools.

Additional analyses were performed on all of the data. For the multiple regression questions, dummy variables were used as independent variables to determine if these variables could significantly predict the academic outcomes. These dummy variables for school and teacher demographics replaced the Professional Development Survey variables used in research Questions One and Two. The results were significant for math but not for reading.

Additional MANOVAs were conducted using various demographic variables to establish groups, and percentages of a class that met or exceeded in mathematics and reading were utilized as dependent variables. Only one test, grade level taught, revealed significant differences.

The next chapter will discuss a summary of the study. Also included in the next chapter will be a discussion of the findings as well as implications for current practice. Limitations recognized in this study will be outline and recommendations for further research will be presented.

CHAPTER FIVE: DISCUSSION

Introduction

The results of analysis of the data for this study were reported in the previous chapter. Chapter five is divided into six sections: a) summary of the study, b) discussion of the findings, c) limitations, d) implications for practice, e) recommendations for further research, and f) summary. These sections will summarize concepts and clarify results obtained associated with relationships between teacher attitude toward assessment and student academic achievement. Suggestions for further research on the relationship of attitudes and assessment, as well as the interaction of teacher and school demographic with attitude and student achievement, are presented.

Summary of the Study

The problem identified in this study was the lack of empirical evidence of the possible relationships between teacher attitude toward assessment and student academic achievement as measured by assessments. Although extensive research focused on assessment has been completed over the last thirty years (Barksdale-Ladd & Thomas, 2000; Black & Wiliam, 1998a, 1998b; Bracey, 2009; Haycock, 2001; Harlan, 2009; Lacina-Gifford & Kher-Durlabhji, 1992; Nicholas & Berliner, 2008a, 2008b; Noddings, 2005; Popham, 2003, 2004, 2006, 2008; Stiggins, 1999, 2002), very little research has been conducted concerning relationships among teacher attitudes and assessment. The process of assessment and application of formal measures of student achievement has become a significant part of educational accountability. In spite of that fact, the attitudes of teachers toward assessments have not been clearly explored.

The purpose of this study was to explore the possible relationships between teacher attitude toward assessment, self-reported behaviors about assessment, and student academic

growth. The secondary purpose was to study teacher beliefs toward assessment and self-reported behaviors about assessment when teachers were grouped by selected demographics.

The study used existing data from the Professional Development Survey. Survey results were analyzed from 148 teachers whose main teaching assignments were kindergarten through fifth grade from the six elementary schools in a school district. In addition to the survey data, existing student academic achievement data was obtained. The data from the Measures of Academic Progress (MAP/MPG) mathematics and reading assessments was utilized to assess student achievement. Percentages of teachers' classes were calculated based on the number of students in the class that met or exceeded norm rate of growth targets between the fall and spring testing sessions. For this portion of the study 143 matched survey results and class percentages were utilized. This study considered five research questions:

1. Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for mathematics as measured by the Measures of Academic Progress assessment be predicted from three independent variables (general assessment belief and MAP/MPG belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*])?
2. Can the percentage of a teacher's class meeting or exceeding the norm rate of growth targets for reading as measured by the Measures of Academic Progress assessment be predicted from three independent variables (general assessment belief and MAP/MPG belief of [*belief*], value of [*value*], and self-reported behaviors [*behavior*])?
3. Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established for certain demographics?

4. Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established in Title One and non-Title One designations?
5. Are there significant mean differences in attitudes regarding assessment (as measured by the combination of general assessment belief, value of MAP/MPG, and self-reported behaviors regarding MAP/MPG) when groups of teachers are established by school?

A step-wise regression model was used to address Questions One and Two utilizing the matched data from the Professional Development Survey and the class percentage data obtained through the Measures of Academic Progress. For Question One mathematics MAP/MPG percentages were used, whereas Question Two incorporated reading data from the MAP/MPG. The Staff Survey for Professional Development was utilized for research Questions Three, Four and Five. Teacher demographic data including teacher education level, teaching experience, and grade level taught were obtained from the survey and analyzed in conjunction with the assessment related portions of the survey. The data in Questions One and Two highlighted possible relationships between attitudes and achievement; however, the data for Questions Three through Five highlighted the interaction between demographics and attitudes.

Discussion of the Findings

Findings for this study will be discussed with reference to the theoretical structure developed by Rokeach (1968). According to Rokeach (1968), beliefs are any "simple proposition, conscious or unconscious, inferred from what a person says or does" (p. 113). Therefore a belief is said to "have observable behavioral consequences" (p. 1). People have tens of thousands of beliefs that are organized in some fashion. One approach to organizations of

beliefs has been described in terms of attitudes. Rokeach (1968) contended that "an attitude is a relatively enduring organization of beliefs around an object or situation predisposing one to respond in some preferential manner" (p. 112). Attitudes are complex and difficult to define; however, they can be characterized by affective, cognitive, and behavioral components. In essence these components dictate a person's response to stimuli. For example, a behavioral response may be in reaction to a strongly held belief that something is right or wrong. Rokeach (1968) analyzed relationships between attitudes and behaviors. He found that although "an attitude leads to preferential (or discriminatory) response, the basis for the preferential response is not clear" (p. 121). For this study, it was posited that a teacher attitude toward assessment then would lead to preferential or discriminatory behaviors. These behaviors may directly or indirectly affect student academic outcomes. In the case under consideration, teachers' response may be associated with strongly held beliefs about instructional behaviors related to assessments. Brown (2004), when studying teacher attitude toward assessment, defined the term of attitude as the way a teacher interacts with, responds to, and understands assessment.

There were two salient findings of this study. First, the math self-reported behavior subscale contributed to the prediction of math achievement. That is, if the teachers marked higher scores concerning behaviors in conjunction with mathematics assessment, then those teachers were more likely to have a higher percentage of students who met or exceeded math growth targets. The relationship was significant; however, it was not strong. The data suggested that teachers' self-reported behaviors related to math were related to math achievement. This finding is consistent to Rokeach's (1968) theory of attitude, that the behaviors of the teachers were associated with student outcomes.

The second major finding was that attitudes of teachers were different when groups were established based on the grade level taught. Post hoc analysis revealed significant differences for all the subscales with the exception of the average reading value. In summary, kindergarten and first grade teachers differed significantly from third, fourth, and fifth grade teachers on the subscales of math belief, math value, math self-reported behavior, reading belief, and reading self-reported behavior. This finding of a difference among the grade levels in teacher attitude was also supported through the additional analysis that revealed academic achievement differences among the grade levels. These results could be associated with several factors.

First, the skill, knowledge, and abilities of kindergarten and first grade students have been found to be quite different from those characteristics of upper grade level students (Kurdek & Monitz-Rodgon, 1975; Guarino, Santibanez, & Daley, 2006). For example, the expected levels of reading and math for kindergarten students are quite different than the expected levels for fourth and fifth grade students. Educators have a belief that assessments of primary students lack reliability and validity (Popham, 2009a). Van Scoy (1994) found that primary teachers were as different from intermediate teachers as primary students were different from intermediate students. Teaching techniques, as well as opportunities for teacher to student conversation, were significantly different between these two groups in Van Scoy's study.

In addition, some believe that early assessments are not predictive of later achievement (Alvidrez & Weinstein, 1999; Boaler, 2003). As a result primary level teacher attitude may be significantly influenced by their beliefs about the abilities of standardized assessment to accurately demonstrate proficiency for kindergarten, first grade, and second grade students. Such attitudes may change if the same teachers were assigned to teach in an upper grade. For example, there may be a noticeable difference for intermediate grade teachers coming from primary grades

in regard to test preparation because these grades include high-stake assessments (Lai & Waltman, 2008).

Historically, high-stakes assessments in the state where the study was conducted have been administered in elementary school in grades 3, 4, and 5, which would suggest that upper grade level teachers have had more exposure to assessment, effective use of results, and a culture of data based decision making. In addition, teachers tend to teach at an assigned grade level for many years; that is, most primary level teachers tend to teach at the primary level year after year. In effect, high-stakes testing is newer to kindergarten, first grade, and second grade teachers than it is to upper grade level teachers. In this state's public schools only these three grade levels do not have state mandated testing. Connected to this historical lack of assessment experience are factors that could influence primary teachers' attitude toward assessments. For example, appropriate professional development, continuous use of data to improve instruction, available data from previous years, and a lack of a district emphasis on data for primary students may predispose the primary teachers to have less assessment literacy, and therefore, question the validity and reliability of assessing young students (Popham, 2004).

This limited involvement with measured accountability may have contributed to more negative attitudes toward assessments from primary teachers. These teachers are not as familiar with testing statistics and protocol as intermediate grade teachers (Popham, 2006). In the particular school district under study, MAP/MPG for reading and mathematics were included for kindergarten, first grade, and second grade students in 2008 whereas third grade, fourth grade, and fifth grade students have participated in MAP or the Achievement Level Test (ALT), the MAP predecessor, since 2002.

It is also possible that teacher negative attitudes toward assessment, particularly in the mathematics assessment, led to decreased effort from students (Brady & Woolfson, 2008). There is a chain of emphasis suggested by Rokeach (1968) that teacher attitude impacts student attitude. If the assessment was not given value by the teacher or the student, then the likelihood that success would be demonstrated on the assessment would not be high (Stiggins, 2002). In 2004 Stiggins warned that if an assessment was used to determine success or failure then the effectiveness of the assessment could possibly decrease. If the assessment results have been utilized to make judgments of teachers' effectiveness, there is a higher likelihood that the assessment itself would be viewed negatively (Brown, 2004). This may have occurred for teachers in the lower primary grades.

Beyond the two findings discussed above a secondary finding of this study was that school differences were associated with differences in teacher attitudes. As stated in Chapter Three, the identity of the schools was removed before the data was provided to the researcher. Therefore, additional analysis into the characteristics of the schools was not possible. As a result the following speculative discussion will be based on possible events and public school characteristics.

School D had higher mean scores on all the survey subscales. Many factors could have contributed to these results. How an assessment is used and defined for educational purposes can influence teacher attitude toward that particular assessment (Brown, Lake, & Matters, 2011; Stiggins, 2004). It is possible that stakeholders from School D gave the assessment high instructional value as well as consistent application of assessment data for student accountability. As such, assessment and related activities could have become part of the culture of School D; therefore, these activities were given greater favor by the teachers at this particular school. Such

a relationship is consistent with recent research (Brown, Kennedy, Fok, Chan, & Yo, 2009; Faulkner & Cook, 2006). Additionally, in School D data based decision making could be a part of the plan for school improvement. The two MAP/MPG assessments could be extensively utilized within School D to support instruction, which would account for the more positive scores on the math and reading self-reported behavior subscales (Stiggins, 1999).

The attitude scores for School E were lower than the corresponding scores from the other schools. Using the above logic, it could be posited that the opposite characteristics could be attributed to School E. These findings are consistent with other studies that consider school differences; quality schools have higher achievement than lower quality schools (McDougall, Saunders, & Goldenburg, 2007). Higher achievement schools also tend to have higher quality and more informed teachers (Hanushuk & Welch, 2006).

Non-significant results were obtained for predictability of reading subscales and reading achievement. Other non-significant findings included identifying differences of teacher attitude toward assessment when groups were established by the teachers' education level, teachers' experience level, and Title One status of the teachers' school.

Limitations

The major limitation of this research study as stated in Chapter 1 was the Professional Development Staff Survey, which was designed by district personnel. In addition to the impacts predicted as limitations several other characteristics of the survey may have impacted the results of this study. First, the correlations among the survey subscales for beliefs and self-reported behaviors were extremely high, indicating that the survey subscales were virtually the same. For example, the value subscale for math was highly correlated with the value subscale for reading.

Another limitation of the survey is related to the reliability and validity of the instrument. The instrument was designed to identify professional development needs of the teachers based on low efficacy scores. Lower mean scores indicated certain professional development areas. The subscales used by the school district were not the subscales used for this study. The validity of the instrument is subject to question: Does the instrument accurately assess the purported subscales? Although face validity was established, construct validity was not established. In addition, the reliability has not been established on populations beyond this district.

In addition to the five limitations as outlined in Chapter 1, other limitations were identified. One additional limitation identified in this study was the limited scope of attitude measurement. This study measured only teacher attitude toward assessment. There are many attitudes that could interact with and have impact on student achievement. The narrow focus of this study did not consider possible interactions between teacher attitudes and student achievement.

A final limitation of this study is the measure of student achievement. The selected measure may have artificially reduced the variance of student achievement. That is, the achievement measures used in this study were based on the percentage of a teacher's class that met or exceeded growth targets. If the measure of academic achievement was more direct, such as fall to spring growth per student, then more student variability would have been included and possibly have produced different results.

Implications for Practice

Because of the limitations of the study, the implications for practice are primarily applicable to the district where the study was conducted. The findings in this study have several district-wide implications. This study identified several key differences between groups of

teachers based on grade level taught. The data indicated that there was a divide between primary and intermediate elementary grades. It could benefit the district to investigate characteristics that distinguish the two groups. The study indicated that primary students did not meet or exceed their growth targets to the extent of their older peers. Primary students were more likely to meet or exceed the academic growth targets in reading than in math. Mathematics instruction could be improved for primary grades. This result suggests that additional support for primary teachers may be beneficial. The district could investigate the different levels of achievement between math and reading. After review, district personnel may acknowledge this difference and make adjustments accordingly.

Popham (2003) found that state-mandated achievement tests have become the key component in state accountability systems. "Data from most states' accountability tests, unfortunately, have almost no value for improving teaching and learning" (p. 48). Such high-stakes summative testing failed to produce the meaningful shifts in instruction that improve student learning. However, Popham (2008) defined a different approach to assessment, formative assessment, as a "planned process in which assessment-elicited evidence of students' status is used by teachers to adjust their ongoing instructional procedures or by students to adjust their current learning tactics" (p. 6). This type of assessment focuses on the information and the instructional shifts that occur to increase student learning; as such, formative assessments can promote improved student achievement. Perhaps the school district personnel may want to investigate other assessment methods that benefit teacher assessment literacy as well as student involvement in assessment results.

The school effect was significant. School level personnel could investigate reasons and possible causes for observed differences across the six schools. These were stark differences in

the survey subscale means between schools as well as the MAP/MPG growth rates. These results could be a component of a needs assessment, which would have implications for professional development and resource allocation.

With respect to professional development, this study highlighted the differences between grade levels as well as between schools. The data indicated the possible need for additional professional development in assessment literacy for primary teachers. Popham (2004) stated that assessment literacy is lacking in teacher development programs and yet is one of the most important aspects of teachers' education in today's school environment. Assessment literacy is the set of skills and knowledge that allow educators to make judgments about validity, reliability, and appropriate application about testing, results, and corresponding instruction.

Relationship to Other Studies

A study completed in the same school district indicated that principals did not value the current accountability measures because principals felt that the assessments included in the accountability system were not measuring the correct things (Medina, 2010). The principals believed that the measurements of student achievement were flawed because they did not accurately reflect what was actually going on inside their particular buildings. According to Medina (2010), principals were forced to reconcile personally held beliefs and those dictated through accountability standards. The failure of principals to be able to do this successfully caused a dissonance within the school district. Rokeach (1968) would have called this collection of beliefs an attitude toward assessment. He posited:

> an attitude predisposes one to respond preferentially not only to the attitude object or situation, but also to other objects – individuals and groups who agree with or oppose us

> with respect to the attitude. (…) Finally, the preferential response may be directed toward the maintenance or preservation of the attitude itself. (Rokeach, 1968, p. 122)

Applying Rokeach's work to that of Medina (2010), principals would have a difficult time maintaining positions that were in conflict to their belief systems.

Logic would then prevail that principals would have preferential responses toward assessments that were not in alignment with district improvement initiatives. These behaviors or attitudes would be on display for teachers to see. Thus, it could be argued that the principal's attitudes are reflected in the teachers' attitudes at a school and key processes within the school. For example, if a principal does not feel that an assessment has value or can accurately assess student achievement, then it would not seem out of place for a teacher to hold that same set of beliefs or attitude. This logic could explain the differences as found across schools in relationship to the fifth research question.

Another connection to previous research is that of teacher experience and educational level. Winkler (2002) found that how an assessment was perceived related almost solely to experience. New teachers were more likely to see assessments as opportunities for professional growth whereas veteran teachers were more likely to view assessment as an infringement on professional freedom and power. Winkler found that veteran teachers conspired to include activities or lessons deemed inappropriate by administration. The lessons were deemed inappropriate if they did not align to the accountability assessments, but veteran teachers were more likely to create a lesson that seemed "legit" regardless of the instructional value of the lesson.

In a review of research, Day & Gu (2009) found similar views of veteran teachers. According to their findings, veteran teachers could alter their beliefs toward educational topics depending on their experience and desire to do so. Day and Gu (2009) stated:

> Throughout their professional lives these teachers will have been confronted by professional, situational and personal pressures and tensions which at times at the very least are likely to have challenged their values, beliefs, practices, for some, their willingness to remain in the job, and for others, their capacity to continue to do their best in the classroom as commitment becomes eroded (p. 442).

In their review, Day and Gu found that veteran teachers were more likely to challenge change because of vast experiences that may contradict the need or the desire to accept to the systemic change. If the required change is in deep conflict with a strongly held belief system or attitude of a teacher, then there is a likelihood that a change in attitude or behavior will not occur regardless of the consequences (Day & Gu, 2009).

On the topic of teacher education addressed in this study, several studies support this study's finding that there was no significant difference in attitude or student academic achievement when teachers were grouped based on educational attainment. The National Council on Teacher Quality (2002) found that degree attainment or advanced certification of teachers did not reflect teaching quality. In research previously discussed in this study, no factor determines student academic success more than the teacher (Hanushuk & Welch, 2006; McDougall, Saunders, & Goldenburg, 2007). It is the quality of the individual teacher that should be examined rather than experience or degree attainment. These findings are parallel to the non-significant differences of attitude or student academic achievement in this study when teachers were grouped based on length of teaching or educational level.

Recommendations for Further Research

The goal of this research study was to add to the limited body of research of how a teacher's attitudes toward assessment could directly or indirectly relate to student achievement. Data was acquired to test five research questions. The information was analyzed, but only a few significant results were obtained. This study did not consider qualitative data; a future study might consider teacher interviews related to attitudes toward assessment. This study used existing data; therefore, follow-up interviews of teachers could not be conducted and related to survey results. Based on the limitations of the survey, the teacher attitude toward the MAP/MPG assessment may not have been accurately portrayed.

Future research into this subject should continue and increase. The impacts of teacher demographics as well as school demographics could not be completely determined in this study. Alvidrez and Weinstein (1999) pointed out the importance of teacher attitude toward student ability and how it impacts the student academic success. Future research should continue with the following suggestions:

- Utilize a proven instrument with established validity and reliability for assessing teacher attitudes
- Widen the scope of investigation to include more than just teacher attitude toward assessment
- Analyze teachers from across different school districts
- Include more precise measurements of student achievement, for example pre and post scores may provide additional student achievement movements

Summary

In this chapter the three findings of this study were discussed. The limitations for this study were outlined and related to previous research. The implications of practice were itemized for stakeholder groups: school district personnel, principals, and teachers. Further research recommendations were also put forth to advance the limited body of empirical research on the topic of the influence of teacher attitude toward assessment.

The empirical evidence provided by this study was limited; however, the theoretical basis for the original study seems valid. As a result, it is suggested that additional studies could consider the same theoretical constructs; however, a more reliable measure of teacher attitude should be used. Teacher assessment literacy remains a dominant concern because data-informed accountability within schools will continue to drive organizational improvement efforts.

References

Alkharusi, H. (2008). Effects of classroom assessment practices on students' achievement goals [Electronic version]. *Educational Assessment, 13,* 243-266.

Alvidrez, J., & Weinstein, R.S. (1999). Early perceptions and later student academic achievement [Electronic version]. *Journal of Educational Psychology, 91*(4), 731-746.

Assessing. (n.d). In *Merriam-Webster online dictionary*. (2010). Retrieved July 4, 2010 from http://www.merriam-webster.com/assessing.

Assessment. (n.d). In *Merriam-Webster online dictionary*. (2010). Retrieved July 4, 2010 from http://www.merriam-webster.com/assessment.

Barksdale-Ladd, M.A., & Thomas, K.F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out [Electronic version]. *Journal of Teacher Education, 51,* 384-397.

Best, J.W. & Kahn, J.V. (2006). *Research in Education, Tenth Edition.* Boston. Pearson Education, Inc.

Black, P., Harrison, C., Lee, C., Marshall, D., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom [Electronic version]. *Phi Delta Kappan, 86*(1). 9-21.

Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry [Electronic version]. *Research Papers in Education, 21*(2), 119-132.

Black, P., & Wiliam, D. (1998a). Inside the black box: Raising standards through classroom assessments [Electronic version]. *Phi Delta Kappan, 80*(2), 139-148.

Black, P., & Wiliam, D. (1998b). Assessment and classroom learning [Electronic version]. *Assessment in Education: Principles, Policy, & Practice, 5*(1), 7-74.

Boaler, J. (2003). When learning no longer matters: Standardized testing and the creation of inequality [Electronic version]. *Phi Delta Kappan, 84*(7), 502-506.

Bracey, G.W. (2009). Identify and observe effective teacher behaviors [Electronic version]. *Phi Delta Kappan, 90*(10), 772-773.

Brady, K. & Woolfson, L. (2008). What teacher factors influence their attributions for children's difficulties in learning? [Electronic version]. *British Journal of Educational Psychology, 78*, 527-544.

Brown, G.T.L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development [Electronic version]. *Assessment in Education, 11*(3), 301-318.

Brown, G.T.L., Kennedy, K.J., Fok, P.K., Chan, J.K.S., & Yu, W.M. (2009). Assessment for improvement: Understanding Hong Kong teachers' conceptions and practices of assessment [Electronic version]. *Assessment in Education: Principles, Policy, & Practice, 16*(3), 347-363.

Brown, G.T.L., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes [Electronic version]. *Teaching and Teacher Education, 27,* 210-220.

Caldas, S. J. & Bankston, C. (1997). Effect of school population socioeconomic status on individual student achievement [Electronic Version]. *The Journal of Educational Research 90(5),* 269-277.

Casanova, P. F., Garcia-Linares, M.C., de la Torre, M. J., & de la Villa Carpio, M. (2005). Influence of family and socio-demographic variables on students with low academic achievement [Electronic version]. *Educational Psychology, 25*(4), 423-435.

Chen, G. & Weikart, L. A. (2008). Student background, school climate, school disorder, and student achievement: An empirical study of New York City's middle schools [Electronic version]. *Journal of School Violence 7*(4), 3-20.

Cimbricz, S. (2002). State mandated testing and teacher beliefs and practice. *Educational Policy Archives, 10*(2), 1-21. Retrieved on November 9, 2009, from http://epaa.asu.edu.epaa/ v10n2.html

Chappuis, S., Chappuis, J., & Stiggins, R. (2009). The quest for quality [Electronic version]. *Educational Leadership, 67*(3), 14-19.

Day, C. & Gu, Q. (2009). Veteran teachers: Commitment, resilience, and quality retention [Electronic version]. *Teachers and Teaching: theory and practice, 15*(4). p. 441-457.

Diamond, J.B. (2007). Where the rubber meets the road: Rethinking the connection between high stakes testing policy and classroom instruction [Electronic version]. *Sociology of Education, 80,* 285-313.

Draper, N. & Smith, H. (1998). *Applied regression analysis (3rd).* New York, NY. John Wiley & Sons.

Faulkner, S.A., & Cook, C.M. (2006). Testing v. teaching: The perceived impact of assessment demands on middle grade instructional practices. *Research in Middle Level Education Online, 29*(7), 1-13. Retrieved August 23, 2009, from http://www.nmsa.org/Publications/RMLEOnline/Articles/Vol29No7/tabid/731/Default.aspx

Field, A. (2005) *Discovering statistics using SPSS.* London: Sage Publication Ltd.

Guarino, C.M., Santibanez, L, & Daley, G.A. (2006) Teacher recruitment and retention: A review of recent empirical literature [Electronic version]. *American Educational Research Association, 76*(2), 173-208.

Guilfoyle, C. (2006). NCLB: Is there life beyond testing [Electronic version]. *Educational Leadership, 64*(3), 8-13.

Guthrie, J.W. & Springer, M.G. (2004) A nation at risk revisited: Did "wrong" reasoning result in "right" results? At what cost?. *Peabody Journal of Education, 79*(1), 7-35.

Hanushuk, E., & Welch, F. (Ed.). (2006). Teacher quality. *Handbook of the Economics of Education: Vol. 2.* (p. 2-24). Doi : 10.1016/S1574-0692(06)02018-6.

Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box [Electronic version]. *Cambridge Journal of Education, 35*(2), 213-224.

Harlan, W. (2009). Improving assessment of learning and for learning. *Education3-13, 37*(3), 247-257. Retrieved August 23, 2009, from http://dx.doi.org/10.1080/03004270802442334.

Haycock, K. (2001). Closing the achievement gap [Electronic version]. *Educational Leadership, 58*(6), 6-12.

Heneman, H.G. & Milanowski, A.T. (1998). Teacher attitudes about teacher bonuses under school-based performance award programs [Electronic version]. *Journal of Personnel Evaluation, 12(4), 327-341*.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? [Electronic version]. *Phi Delta Kappan, 89*(2), 140-145.

Introduction to SAS. UCLA: Academic Technology Services, Statistical Consulting Group. Retrieved on July 29, 2010 from http://www.ats.ucla.edu/stat/sas/notes2.

Jacobs, N. & Harvey, D. (2005). Do parents make a difference to children's academic achievement? Differences between parents of higher and lower achieving students [Electronic version]. *Educational Studies, 31*(4), 431-448.

Jones, B.D., & Egley, R.J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing [Electronic version]. *Educational Policy Analysis Archives, 12*(39), 1-29.

Jones, B.D., & Egley, R.J. (2006). Looking through different lenses: Teachers' and administrators' views of accountability [Electronic version]. *Phi Delta Kappan, 87*(10), 767-771.

Kurdek, L. A. & Monitz-Rodgon, M. (1975). Perceptual, cognitive, and affective perspective taking in kindergarten through sixth grade students [Electronic version]. *Developmental Psychology, 11*(5), 643-650.

Lacina-Gifford, L.J., & Kher-Durlabhji, N. (1992). The price of high-stakes testing on children [Electronic version]. *Education, 112*(4), 565-566.

Lai, E.R. & Waltman, K. (2008). Test preparation: Examining teacher perceptions and practices [Electronic version]. *Educational Measurement: Issues and Practice. 27*(2), 28-45.

Lord, F.M. (1980). *Applications of item response theory to practical testing problems.* Mawhaw, New Jersey: Lawrence Erlbaum Associates.

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction.* Alexandria, VA: Association for Supervision and Curriculum Development.

McCall, M.S., Hauser, C., Cronin, J., Kingsbury, G.G., & Houser, R. (2006). Achievement gaps: An examination of differences in student achievement and growth. *Northwest Evaluation Association.* 1-4, Retrieved on October 23, 2009, from http://www.nwea.org/research /gap/asp

McDougall, D., Saunders, W.M., & Goldenburg, C. (2007). Inside the black box of school reform: Explaining the how and why of change at *Getting Results* schools [Electronic version]. *International Journal of Disability, Development, and Education, 54*(1), 51-89.

Medina, R. L. (2010) *The beliefs that school administrators hold about the role of data in driving school improvement* (Unpublished doctoral dissertation). University of Nevada Reno, Reno, Nevada.

Meece, J. L., Anderman, E. M., & Anderman, L.H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57,* 487-503.

Mertler, C.A. & Vannatta, R.A. (2010). *Advanced and multivariate statistical methods: Practical application and interpretation (4th ed.).* Glendale, CA; Pyrczak Publishing.

Merriam-Webster Online Dictionary. (2010). Retrieved July 4, 2010 from http://www.merriam-webster.com/

Moore, D., & McCabe, G. (2006). *Introduction to the practices of statistics (5th ed.).* New York, NY; Freeman.

National Council on Teacher Quality. (2002). *Expanding the teacher quality discussion.* (Report No. ED477730). Washington, D.C.: ERIC.

Nichols, S.L., & Berliner, D.C., (2008a). Testing the joy out of learning [Electronic version]. *Educational Leadership, 65*(6), 14-17.

Nichols, S.L., & Berliner, D.C. (2008b). Why has high-stakes testing so easily slipped into contemporary America? [Electronic version]. *Phi Delta Kappan, 89*(9), 672-676.

Noddings, N. (2005). What does it mean to educate the whole child? [Electronic version]. *Educational Leadership, 63*(1), 8-13.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 145 (2002).

NWEA. (2008) Technical manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades. Retrieved on May 24, 2010, from http://www.nwea.org/technicalmanual.

Palarady, G.F. & Rumberger, R.W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30*, 111-140. Retrieved on August 23, 2009, from http://epa.sagepub.com/cgi/content/abstract/30/2/11

Pearson Inc. (2008). What is formative assessment? [Electronic version]. *Research Bulletin No.5.* 1-2.

Podulka Coe, D., Pivarnik, J.M., Womack, C.J., & Reeves, M.J. (2006). Effects of physical education and activity levels on academic achievement in children. *American College of Sports Medicine,* 1515-1519. Retrieved July 1, 2010, from http://www.acsm-msse.org.

Popham, W.J. (2003) The seductive allure of data [Electronic version]. *Educational Leadership 60*(5), 48-51.

Popham, W.J. (2004) Why assessment illiteracy is professional suicide [Electronic version]. *Educational Leadership 62*(1), 82-83.

Popham, W.J. (2006). Needed: A dose of assessment literacy [Electronic version]. *Educational Leadership 63*(6), 84-85.

Popham, W.J. (2006, May). Assessment for learning: An endangered species. *Mathematics Teacher, 2*(7), 1-2. Retrieved November 9, 2009, from http://www.aps.edu/aps/tis/index.htm

Popham, W. J. (2008). Transformative Assessment. Alexandria, VA: Association for Supervision and Curriculum Development.

Popham, W.J. (2009a). A process: Not a test [Electronic version]. *Educational Leadership, 66*(7), 85-86.

Popham, W.J. (2009b). Assessment literacy for teachers: Faddish or fundamental? [Electronic version]. *Theory into practice,48*, 4-11.

Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change.* San Francisco: Jossey-Bass Inc.

Rumberger, R.W. (1995). Dropping out of middle school: A multilevel analysis of students and schools [Electronic Version]. *American Educational Research Journal, 32*(3), 583-625.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A Meta-analytic review of research [Electronic version]. *Review of Educational Research,75(3),* 417-453.

Sprinthall, R.C. (2007). *Basic statistical analysis, eighth edition.* Boston: Allyn & Bacon.

Stiggins, R.J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238-246. Retrieved November 9, 2009, from http://0-search.ebscohost.com.innopac.library.unr.edu/login.aspx?direct=true&db=aph&AN=9512053723&site=ehost

Stiggins, R. J. (1999). Assessment, student confidence, and school success [Electronic version]. *Phi Delta Kappan, 81*(3), 191-198.

Stiggins, R.J. (2002). Assessment crisis: The absence of assessment for learning [Electronic version]. *Phi Delta Kappan, 83*(10), 758-765.

Stiggins, R.J. (2004). New assessment beliefs for a new school mission [Electronic version]. *Phi Delta Kappan, 86*(1), 22-27.

Stiggins, R. J. (2009). Assessment for learning in upper elementary grades [Electronic version]. *Phi Delta Kappan, 90*(6), 419-421.

Stiggins, R.J., & DuFour, R. (2009). Maximizing the power of formative assessments [Electronic version]. *Phi Delta Kappan, 90*(9), 640-644.

Taras, H., (2005). Physical activity and student performance at school [Electronic version]. *Journal of School Health, 75*(6), 214-218.

Van de Watering, G., Gijibels, D., Dochy, F., & van der Rijt, J. (2008). Students' assessment preferences, perceptions of assessment and their relationship to study results [Electronic version]. *Higher Education, 56,* 645-658.

Value. (n.d). In *Merriam-Webster online dictionary*. (2010). Retrieved July 4, 2010 from\ http://www.merriam-webster.com/value.

Van Scoy, I. J. (1994). Differences in teaching between six primary and five intermediate teachers in on school [Electronic version]. *The Elementary School Journal 94* (3), 347-356.

Vogler, K. (2002). The impact of high-stakes, state mandated student performance assessment on teachers' instructional practices. *Education, 123*(1), 39-55. Retrieved on November 23, 2009 from http://teachers.dadeschools.net

Wayne, A. J. & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73(1).* 89-122.

Weiss, D. J. (1992). Improving measurement quality and efficiency with adaptive testing \ [Electronic version]. *Applied Psychological Measurement, 6*(4). 473-492.

Wilson, M. (2009). Assessment from the ground up. *Phi Delta Kappan, 91*(1), 68-71. Retrieved November 9, 2009, from http://pdkintl.org

Winkler, A. (2002). Division in the ranks: Standardized testing draws lines between new and veteran teachers [Electronic version]. *Phi Delta Kappan, 84*(3). p. 219-226.

Appendix A

Questions 1 through 19 will be answered with the following scale
Answers in Likert-type forced response

Very Strongly Disagree
Strongly Disagree
Disagree
Agree
Strongly Agree
Very Strongly Agree

1. Assessment is the best way to evaluate a student's achievement.
2. The current system of assessment at my grade level is sufficient to accurately evaluate my students' academic achievement.
3. The assessments I use in my classroom are more informative than the assessments my students take from outside resources.
4. The combination of assessments, including my classroom, district, and state assessments used in my classroom are the best way to gauge academic achievement.
5. MAP/MPG assessments for reading and math are generally useful.
6. My students' academic **achievement** in math is accurately reported by the MAP/MPG assessment.
7. My students' academic **achievement** in reading is accurately reported by the MAP/MPG assessment.
8. My students' academic **growth** in math is accurately reported by the MAP/MPG assessment.
9. My students' academic **growth** in reading is accurately reported by the MAP/MPG assessment.
10. The time students spend **testing** on the math MAP/MPG assessment is worthwhile.
11. The time students spend **testing** on the reading MAP/MPG assessment is worthwhile.
12. The time spent during the school day to **discuss** math MAP/MPG data is adequate.
13. The time spent during the school day to **discuss** reading MAP/MPG data is adequate.
14. The time spent during the school day to **plan** math instruction based on MAP/MPG data is adequate.
15. The time spent during the school day to **plan** reading instruction based on MAP/MPG data is adequate.
16. The MAP/MPG data for math is **more accurate** than my classroom assessment data.
17. The MAP/MPG data for reading is **more accurate** than my classroom assessment data.
18. Professional development for **accessing** MAP/MPG reports on the NWEA website at my school site is adequate.
19. Professional development for **interpreting** MAP/MPG reports on the NWEA website at my school site is adequate.
20. Professional development for **planning instruction** using MAP/MPG data reports at my school site is adequate.

Questions 21 through 36 will be answered with the following scale

Answers in Likert-type forced response of:
To a great extent
To some extent
To a slight extent
To no extent at all

Please answer the following questions by completing the statement: In general, I believe…..

21. I use MAP/MPG data to **assess** my students' math achievement.
22. I use MAP/MPG data to **assess** my students' reading achievement.
23. I use MAP/MPG data to **assess** my students' **growth** in math achievement.
24. I use MAP/MPG data to **assess** my students' **growth** in reading achievement.
25. I use MAP/MPG data to report my students' math **achievement** to **parents.**
26. I use MAP/MPG data to report my students' reading **achievement** to **parents.**
27. I use MAP/MPG data to report my students' **growth** in math achievement to **parents.**
28. I use MAP/MPG data to report my students' **growth** in reading achievement to **parents.**
29. I use MAP/MPG data to **plan instruction** in math.
30. I use MAP/MPG data to **plan instruction** in reading.
31. I use MAP/MPG data to **evaluate** my own instruction effectiveness in math.
32. I use MAP/MPG data to **evaluate** my own instruction effectiveness in reading.
33. The Class by RIT report is useful.
34. The Class Report is useful.
35. The Growth Summary Report is useful.
36. The instructional tools including the NWEA DesCartes and the Staircases are useful.

Answer choice for question 37 is forced response and uses the following choices:

Fall
Spring
Both are equally important
Neither

37. In general, I believe the most useful MAP assessment is given in:

For the following demographic questions, please base your answers on the current school year.

38. I have taught for _____ years. (Fill in number response)
39. I have taught in the [redacted] School District for _____ years. (Fill in number response)
40. I am a: (Forced response)
    a. Male
    b. Female

41. I currently teach: Choose one that most describes your current position (Forced response)
    a. Kindergarten
    b. First Grade
    c. Second Grade
    d. Third Grade
    e. Fourth Grade
    f. Fifth Grade
    g. Special Education
    h. English as a Second Language
    i. Other: Please specify _____
42. How many school **districts,** including the [redacted] school district, have you worked in? _____ (Fill in number response)

43. Please indicate your highest level of education: (Forced response)
    a. A Bachelors degree
    b. A Bachelors degree and some graduate work
    c. A Masters degree
    d. A Masters degree and some additional higher level course work
    e. A Doctoral degree

44. Where do you currently work: (Forced response)
    a. [redacted] Elementary School
    b. [redacted] Elementary School
    c. [redacted] Elementary School
    d. [redacted] Elementary School
    e. [redacted] Elementary School
    f. [redacted] Elementary School