1-1-2012

Using Mathematics-Curriculum Based Measures to Predict Outcomes on the Mathematics Portion of the Mississippi Curriculum Test, Second Edition.

Eutrophia Lenora Hoganamuel

USING MATHEMATICS-CURRICULUM BASED MEASURES TO PREDICT OUTCOMES ON THE MATHEMATICS PORTION OF THE MISSISSIPPI CURRICULUM TEST, SECOND EDITION

By

Eutrophia Lenora Hogan-Samuel

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Elementary, Middle, and Secondary Education Administration
in the Department of Leadership and Foundations

Mississippi State, Mississippi

May 2012

USING MATHEMATICS-CURRICULUM BASED MEASURES TO PREDICT OUTCOMES ON THE MATHEMATICS PORTION OF THE MISSISSIPPI CURRICULUM TEST, SECOND EDITION

By

Eutrophia Lenora Hogan-Samuel

Approved:

Debra L. Prince
Associate Professor of Leadership and Foundations
(Director or Dissertations)

James E. Davis
Associate Professor of Leadership and Foundations
(Committee Member)

Amanda Taggart
Assistant Professor of Leadership and Foundations
(Committee Member)

Dana Franz
Associate Professor of Secondary Education - Mathematics
(Committee Member)

Dwight Hare
Professor and Graduate Coordinator of Leadership and Foundations

Richard Blackbourn
Dean of the College of Education

Name: Eutrophia Lenora Hogan-Samuel

Date of Degree: May 11, 2012

Institution: Mississippi State University

Major Field: Elementary, Middle, and Secondary School Administration

Major Professor: Dr. Debra L. Prince

Title of Study: USING MATHEMATICS-CURRICULUM BASED MEASURES TO PREDICTOUTCOMES ON THE MATHEMATICS PORTION OF THE MISSISSIPPI CURRICULUM TEST, SECOND EDITION

Pages in Study: 62

Candidate for Degree of Doctor of Philosophy

The nation is challenged with improving the mathematics achievement of its students. No Child Left Behind holds schools, districts, and states accountable for improving student achievement. Because high stakes tests are given at the end of the school year, schools are presented with the challenging task of developing or purchasing reliable assessments that provide accurate information describing how well students understand the skills that will be measured on the end-of-the-year high stakes tests. Curriculum based measurements are used periodically to measure student progress toward meeting objectives during the school year. The problem exists that schools are utilizing limited resources of time and money on a tool with little evidence of effectiveness in increasing mathematics scores on state assessments.

The purpose of this study was to determine the relationships between the scores of the three assessments of the Mathematics-Curriculum Based Measures (M-CBM) and the scores from the mathematics MCT2 for sixth-grade students. A correlational research design is used to fulfill the purpose and test the three null hypotheses. Hypothesis 1 for this study states that there is no relationship between the August 2009 M-CBM scores

and the May 2010 mathematics MCT2 scores. Hypothesis 2 states that there is no relationship between the December 2009 M-CBM scores and the May 2010 mathematics MCT2 scores. Hypothesis 3 states that there is no relationship between the April 2010 M-CBM scores and the May 2010 mathematics MCT2 scores. To test the three null hypotheses, correlation coefficients were computed using the Pearson *r*.

The results from all three hypotheses indicated that there were moderate positive correlations between scores of the M-CBM and scores of the mathematics portion of the MCT2, with the strongest relationship being between the April M-CBM and the mathematics MCT2 scores. Further analysis was done to determine if the relationship between M-CBM and mathematics MCT2 scores continued to exist when examined by mathematics MCT2 proficiency levels. The results of this analysis indicated that relationships between M-CBM and the mathematics portion of the MCT2 scores only existed for students scoring in the proficient range.

DEDICATION

I would like to dedicate this research to my parents, Emma and Johnnie Hogan, my son, Trace Samuel, my husband, Marvin Samuel, and all 13 of my siblings. I would also like to dedicate this research to my uncles, aunts, nieces, nephews, cousins, friends, and all of the students whom I have come into contact with throughout my educational career.

ACKNOWLEDGEMENTS

I would like to first thank my Lord and Savior Jesus Christ for his many blessings. Secondly, I would like to thank my major professor, Dr. Debra L. Prince, for her time, commitment, encouragement, and straight-forwardness. She allowed me to take up several of her weekend and holiday hours to accomplish my goal. I would also like to acknowledge my committee members, Dr. Dana Franz, Dr. Amanda Taggart, and Dr. James E. Davis for their feedback and support. In addition, I would like to thank my study group, Stephanie, Susan, Andrea, Roderick, and Tim, for going through this process with me and reminding me that we could finish the course. I want to thank all of those people who always reminded me of my mission to complete this process and encouraged me by believing that I could do it when I did not believe it myself. I would also like to thank the New Zion United Methodist Church family and my pastor, Reverend Tyrone Stallings, Sr., for their concern, motivation, and prayers. Last but not least, I would like to thank my parents, Emma and Johnnie, and my twin sister, Ledora, not only for their support, but also for their willingness to take care of my son, Trace, throughout this process. A special thank you goes to my mother for encouraging me to pursue this degree. Without her encouragement, I would not have attempted this process.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The foundation of former President George W. Bush's *No Child Left Behind Act of 2001*(NCLB) was a promise to raise standards for all children and to help all children meet those standards (U. S. Department of Education [US DOE], 2001). Standards are the defining skills that are to be mastered by students at particular grade levels, and standardized tests are used to assess for mastery of these standards (Olson, 2002). According to the 2009 National Assessment of Educational Progress (NAEP), the nation's students are not meeting the standards. The results of the 2009 NAEP indicated that many of the nation's students are not proficient in reading or mathematics (US DOE, 2009). Nevertheless, in the NCLB legislation, the goal of having all students score proficient in reading and mathematics by 2014 was established, with penalties for failure to achieve the goal (Robelen, 2002; Simpson, LaCava, & Graner, 2004). Furthermore, the national education agenda is calling for more rigorous standards and placing more demands on schools to increase student performance as measured by standardized, high-stakes tests. Because of these demands, high-stakes tests are driving our nation's schools (Buchanan, 2007).

Based on standards that are set by each state, state tests drive much of the educational decision-making for schools. In Mississippi, the Mississippi Curriculum Test (MCT) was the measure used to determine if students were meeting the standards

(Mississippi Department of Education [MDE], 2002). However, according to the MDE (October, 2008), revisions to the MCT were made in an effort to meet the demands of NCLB and to increase the rigor of questions on the MCT so that the questions would be more parallel to those on the NAEP. In May of 2007, the first pilot of the revised MCT, the Mississippi Curriculum Test, 2nd Edition (MCT2), was given just days after the MCT was administered for the final time. While the results of the MCT2 pilot were not made public, results from the final MCT showed that many of the state's students were not proficient in reading or math. Of particular interest to this study were the high percentages of sixth-grade students who were not proficient in math. According to the results of the 2007 MCT (MDE, 2011), 30% of Mississippi's sixth graders performed at the minimal and basic levels in mathematics. While NAEP does not assess sixth graders, the results of the 2007 NAEP assessment in mathematics indicated that 79% of fourth graders and 86% of eighth graders in Mississippi were not proficient in math (NAEP, 2008). This discrepancy in rates of proficiency between the MCT scores and NAEP scores justified the need to revise the MCT. Consequently, in May 2008, the MCT2 was given to all students in Grades 3 through 8 in mathematics and language arts, and it became the measure that would be used to determine if students were meeting the standards (MDE, 2009).

A comparison of the 2009 NAEP mathematics scores for the state of Mississippi and the 2009 MCT2 scores continued to reveal a discrepancy between the two measures. The 2009 NAEP results indicated that 78% of Mississippi's fourth graders and 85% of Mississippi's eighth graders scored less than proficient on the NAEP assessment (National Center for Education Statistic [NCES], 2009) while the results of the state-wide 2009 MCT2 mathematics assessment indicated that 43% of fourth graders and 46% of eighth graders scored less than proficient in mathematics (MDE, 2009). Consequently, there is evidence to suggest that the Mississippi standards as measured by the MCT2 are not as rigorous as the national standards assessed by NAEP. Nevertheless, Mississippi's schools are rated contingent on how well students are performing in reference to state standards measured by the state tests.

Schools in Mississippi are rated with the following levels: Star School, High Performing, Successful, Academic Watch, Low Performing, At-Risk of Failing, and Failing (MDE, 2010). Schools that are identified as failing can have federal and state sanctions imposed upon them (Robelen, 2002). Therefore, schools continuously search for strategies and programs that will increase student achievement, thereby avoiding state and federal sanctions. Moreover, if the strategies and programs selected by the schools are funded by federal dollars, then they must be research-based methods that have demonstrated effectiveness in increasing student achievement (Simpson et al., 2004).

One research-based strategy mandated by NCLB and employed by school districts throughout the United States is the use of student progress monitoring systems (Simpson et al., 2004; US DOE, 2001). According to NCLB regulations, schools are required to track the progress of all students and initiate a response to intervention (RTI) for students identified by the assessment as being at-risk for academic failure (Helwig, Anderson, & Tindal, 2002; Katsiyannis, Zhang, Ryan, & Jones, 2007; Simpson et al., 2004). Progress monitoring systems are scientifically-based practices that are designed to track students' academic progress as the first step of RTI. In most cases, private companies design and develop these systems, and school district administrators determine which system their schools will use (Ysseldyke & McLeod, 2007). However, the responsibility for implementing the progress monitoring system is designated to building-level teachers who, more often than not, have had little input on its selection and have limited knowledge regarding the purpose and content of the system (Simpson et al., 2004).

One particular system, Systems to Enhance Educational Performance (STEEP), is a student progress-monitoring system being used by districts across the nation (STEEP, 2007). According to Witt and Chun (2006), STEEP is an inclusive system that is being used to aide teachers, administrators, and other professionals in making provisions for students at the right time. This system provides evidence-based tools that help schools achieve the ambitious goals set forth by NCLB (STEEP, 2007).

STEEP is comprised of five components. These components are a universal screener, Can't Do/Won't Do assessments, intervention/remediation, and progress monitoring with RTI, and a data management system. The primary component and the component that informs all of the other components of STEEP is the universal screener. With STEEP, curriculum based measurements (CBM) serve as the universal screener (STEEP, 2007). The CBM, which measures students' content knowledge of particular curricula, is administered to the entire student population, scored by a team of assessors, and recorded into the STEEP database. The database, which serves as the monitoring system, identifies the class median as it relates to the instructional standard (VanDerHeyden, Witt, & Gilbertson, 2007). If the class median falls below the instructional standard, a class-wide intervention is prescribed; however, if the class median is at or above the instructional standard, then students whose scores represent the bottom 16th percentile are identified as having individual problems (STEEP, 2007). Students scoring in the bottom 16th percentile are then administered the Can't Do/Won't Do assessment. The actual items on the Can't Do/Won't Do assessment are identical to the items on the assessment that was initially given; however, the conditions under which it is given have been altered. This time, the Can't Do/Won't Do assessment is given one-on-one, and the student is rewarded with a prize of his choice if improvement is demonstrated. Students who continue to score in the bottom 16th percentile are provided with interventions and remediation as required by the RTI clause in NCLB (STEEP, 2007). This process is repeated three times a year, in August, December, and April, for the entire student population. Therefore, STEEP is a time-consuming process for schools utilizing the system. In addition to STEEP being a time consuming process, the system requires significant resource allocations from the district.

According to J. Fancher (personal communication, July 17, 2010), the Director of Interventions for a district with a student population of 4,000, the initial implementation of STEEP required over $100,000. Fancher also stated that the initial cost did not include the hiring of additional personnel or the recurring cost of providing the assessments. In total, the first year of operating STEEP in Fancher's district was over $500,000. In spite of the enormous cost associated with using student progress monitoring systems, NCLB requires that districts utilize them (Simpson et al., 2004).

Statement of the Problem

School districts across the United States spend thousands of dollars on professional development, supplemental materials, and programs to improve student achievement. NCLB holds schools, districts, and states accountable for improving student achievement (Simpson et al., 2004). Because of these accountability demands for academic gains, NCLB has increased the demands for early identification and intervention for students who are at risk of academic failure (Batsche et al., 2005). Most states use comprehensive state assessments to evaluate school districts; however, these assessments are given at the end of the school year (McGlinchey & Hixson, 2004). Because these high stakes tests are given at the end of the school year, schools are presented with the challenging task of developing or purchasing reliable assessments that provide accurate information describing how well students understand the skills that will be measured on the end-of-the-year high stakes tests.

CBM are used to measure student progress toward meeting curriculum objectives periodically during the school year (Helwig et al., 2002; Hintze & Silberglitt, 2005; Jiban & Deno, 2007; Jitendra, Sczesniak, & Deatline-Buchman, 2005). Data gathered from CBM are immediately available to teachers and can be used to guide instruction. However, the use of CBM is time consuming and expensive and the efficacy of their use with mathematics is not well supported. While there are studies that suggest reading CBM scores are related to state assessments in reading (McGlinchey & Hixon, 2004; Wiley & Deno, 2005), very little research has been done examining the relationship between mathematics CBM and state assessments in mathematics. Therefore, the problem exists that schools are utilizing limited resources of time and money on a tool with little evidence of effectiveness in increasing mathematics scores on state assessments.

Purpose of the Study

The nation is challenged with improving the mathematics achievement of its students and many school districts have implemented programs such as STEEP to help increase mathematics achievement (Simpson et al., 2004). An integral element of STEEP is the CBM. A CBM is a set of specific measurement methods for assessing student progress over time and for identifying students in need of additional instructional support and/or further diagnostic testing (McGlinchey & Hixson, 2004). However, with the implementation of CBM with mathematics, schools are utilizing valuable time and money on a program that does not have strong empirical support for its success. The purpose of this study is to determine the relationships between students' scores on the three Mathematics-Curriculum Based Measures (M-CBM) and the scores from the mathematics MCT2 for sixth grade students.

Research Hypotheses

The M-CBM is administered three times a year and the MCT2 is administered once a year. To fulfill the purpose of this study, the following three null hypotheses will be tested:

1. There is no relationship between the August 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.
2. There is no relationship between the December 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.
3. There is no relationship between the April 2010 M-CBM scores and the May 2010 mathematics MCT2 scores.

Significance of the Study

Educators are constantly searching for ways to improve student achievement. More specifically, they are trying to increase performance levels on state tests so that all students are proficient in reading, language arts, and mathematics by 2014 (Katsiyannis et al., 2007; Robelen, 2002; Simpson et al., 2004). CBM are being used as one method of monitoring students' progress toward proficiency. The purpose of this study is to determine the relationships between the scores of three M-CBM and the scores from the mathematics MCT2 for sixth grade students. This study is significant for several reasons. First, although the use of progress monitoring systems is required by NCLB, knowing whether or not this particular system is effective would help the decision makers of particular school districts make informed decisions and could save districts time and money (Deno, 2003; Helwig et al., 2002; Jiban & Deno, 2007; Robelen, 2002). Moreover, if the results of this study indicate that the M-CBM is a valid and reliable indicator as to how students will perform on the state assessment, teacher buy-in for the M-CBM could increase. Secondly, the STEEP system utilizes a CBM that is given to all students three times a year. If data gathered from one testing cycle provides as much information as data provided by three testing cycles, then school districts could save a great deal of time and money by conducting the M-CBM once a year. On the other hand, if the results of this study indicate that the M-CBM scores are not related to mathematics MCT2 scores, then district personnel would have support for discontinuing the use of M-CBM and begin searching for other tools that would provide more valid and reliable information.

Delimitations

This study is delimited to one school and one grade level. All data used in this study were collected and archived by one school district in Mississippi during the 2009-2010 school year. Currently, there are only two school districts in the state of Mississippi that use STEEP. The school district that will be used in this study was selected because of convenience and because of the amount of time the system has been used in the district. This district is in the second year of implementing the system whereas the other district began using the system in 2004. This study will also be delimited in that the population for the study will be limited to sixth grade students who took the MCT2 in mathematics in May of 2010. Sixth grade students were chosen because of convenience as well as the fact that at the sixth grade level, students begin middle school, and according to Geary (2004), this is the time when the mathematics performance of American students begins to decrease. Consequently, although generalization of the results of this study beyond this one school district and grade are not warranted, the results of this study could justify replication of the study with a broader population.

Theoretical Framework

The theory that provides the foundation for this study is that of constructivism as it relates to formative assessments. The most significant objective of formative assessments is the notion that information obtained from the results would improve instruction and thereby improve student learning (Black & Wiliam, 1998; 2004; Fuchs & Fuchs, 1986; Ketterlin-Geller, McCoy, Twyman, & Tindal, 2003)**.** Consistent with the constructivist theory, the information gathered from the M-CBM should inform teachers and help them construct plans for further instruction. While several theorists are associated with the theory of constructivism (Brooks & Brooks, 1999; Bruner, 1966; Piaget, 1965; Vygotsky, 1978), the theory of Black and Wiliam (1998) is most directly related to the current study. According to Black and Wiliam (2004), effective formative assessments improve student learning. The authors also contend that effective formative assessments have the greatest impact on low achieving students. Moreover, Wiliam and Thompson (2008) asserted that, "with formative assessments, the 'big idea' is that evidence about student learning is used to adjust instruction to better meet student needs— in other words, that teaching is *adaptive* to the student's learning needs" (p. 64).

Consequently, M-CBM, used as a formative assessment, should monitor student progress over time and provide information to teachers that will help them adapt their teaching to meet the needs of their students.

Definition of Terms

The following is a list of terms that were used throughout this study to describe several concepts:

1. *Computational fluency* refers to the accuracy, efficiency, and flexibility with which students can compute (Russell, 2000).
2. *Curriculum Based Measurements* are a set of specific measurement methods for assessing student progress over time and for identifying students in need of additional instructional support and/or further diagnostic testing (McGlinchey & Hixson, 2004).
3. *High Stakes Testing* is any testing program whose results have important consequences for students, teachers, schools, and/or districts if certain performance levels are not met (Johnson & Johnson, 2002).
4. *Response To Intervention (RTI)* is a method used to identify learning disabilities and providing early intervention to all children at risk for school failure (Webb, 2007).
5. *Universal screening* refers to the systematic assessment of all children within a given class, grade, school building, or school district, on academic and /or social-emotional indicators that the school personnel and community have agreed are important (Ikeda, Neesen, & Witt, 2008).

CHAPTER II

LITERATURE REVIEW

When people struggle with mathematics, they are limited in their career options (Hadley, 2005). However, according to the National Council of Teachers of Mathematics (NCTM, 2000), those who are able to do mathematics and clearly understand its content will have more opportunities and options in deciding their futures. Consequently, being competent in mathematics provides a productive future; whereas the lack of competence in mathematics will delete many job opportunities. But, according to NCTM (2000), schools are not producing the mathematics excellence required for global economic leadership and homeland security in the 21st century. America's schools are being challenged to improve mathematics achievement; therefore, research examining the predictive value of mathematics CBM as it relates to scores on state assessments is warranted.

The purpose of Chapter II is to provide a review of the literature related to the need and efficacy of CBM. Included in this review are the regulations set forth by NCLB regarding adequate yearly progress (AYP) and the need to monitor student progress, research related to formative assessments, and empirical evidence related to the use of CBM. The chapter will conclude with a brief summary of the literature reviewed.

No Child Left Behind

When former President George W. Bush proposed that NCLB be enacted, his main focus was to strengthen elementary and secondary schools all over the United States by ensuring that every child had a fair, equal, and significant opportunity to obtain a high quality education and reach a level of proficiency as measured by state academic assessments (Simpson et al., 2004). Entering the 21st Century, President Bush noticed that many of the nation's students were being left unnoticed in the area of reading. During that time, almost 70% of fourth graders in inner city schools were unable to read on a basic level, according to national reading tests (White House, 2001). On the international level, high school seniors in the United States trailed those in Cyprus and Africa on international mathematics tests. On the college level, one-third of college freshmen had to take remedial courses before they were able to enroll in regular college-level courses (White House, 2001). In recognizing the failings of the education system on all levels, former President Bush realized that the federal government would need to have a more active role in holding states accountable for the educational achievement of its students.

NCLB was built on the theory that setting high academic standards and creating measurable goals would not only increase educational accountability but also improve individual outcomes in education (Guthrie & Springer, 2004). At the heart of the act were measures that would help improve student achievement while holding states and schools more accountable for monitoring the progress of all students (US DOE, 2001). The intention of NCLB was to encourage a national and statewide school reform movement that would challenge the academic standards for all students (Jitendra et al., 2005). The goal of NCLB is for every student to meet state-identified academic standards in mathematics and reading with proficiency by 2013-2014 (Katsiyannis et al., 2007; NCLB, 2001; Shapiro, Keller, Lutz, Santoro, & Hintze, 2006; Shriberg & Kruger, 2007; Silberglitt, Burns, Madyun, & Lail, 2006). In order to determine whether schools are achieving AYP, this law requires that all students are assessed in reading and mathematics in grades three through eight. From these assessments, the school's accountability is decided (Katsiyannis et al., 2007; NCLB, 2001; Shapiro et al., 2006; Shriberg & Kruger, 2007; Silberglitt et al., 2006). Consequently, as a mechanism for accountability, the federal government could withhold funding from states that fail to demonstrate progress toward that goal (Guthrie & Springer, 2004; US DOE, 2001).

According to NCLB regulations, schools are also required to track the progress of all students by monitoring student growth over time and initiate an RTI for students identified by the assessment as being at risk for academic failure (US DOE, 2001). Progress monitoring is used to ensure that students are mastering academic content and that student growth rates are adequate to justify current instructional and placement decisions (Vaughn, Linan-Thompson, & Hickman, 2003). Although NCLB mandates standardized, high-stakes tests from each state, this type of test is not administered frequently enough nor is it effective in providing information that could help identify at-risk students or help guide instruction (US DOE, 2001).

Formative Assessment

With the demand to frequently assess students to monitor progress as opposed to the one administration of the high stakes test, the demand for formative assessments has become very popular (Pinchok & Brandt, 2009). Researchers provide several definitions for formative assessment. Formative assessments are used as an on-going process within the class to adjust teaching to meet student needs and provide feedback (Black & Wiliam, 1998). Heritage, Kim, Vendlinski, and Herman (2009) defined formative assessment as a systematic process to continuously gather evidence about learning. Various tools and strategies may be used to determine what students know, to identify gaps in understanding, and to plan future instruction to improve learning. In this process, students and instructors become better informed as to the degree to which materials have been mastered (Algozzine, Ysseldyke, & Elliott, 1997).

Any type of assessment can be used in formative assessment practices whether it is performance-based or utilizes multiple choice, observation checklists, journals, or portfolios. According to Burch (2006), a growth in commercially-created formative assessments has been seen. Regardless of whether the formative assessment is teacher made, district made, or commercially created, there are key aspects to be considered to effectively use formative assessments: purpose, cycle of use, and planned integration with instruction (Frohbieter, Greenwald, Stecher, & Schwartz, 2011).

When assessing formatively, the most important aspect knows how the information is to be used (Black & Wiliam, 2004; Burns, VanDerHeyden, & Jiban, 2006). Not only must teachers know how the information will be used, but they must also follow through with a plan of action that consists of adapting the results to modify teaching and learning (Black & Wiliam, 1998). When teachers are aware of how students are progressing or where they are having trouble, strategies and activities can be used that will assist in meeting the needs of individual students (Black & Wiliam, 2010). In adapting the results to modify teaching and learning, teachers must involve students in the process of improving their learning.

According to Popham (2008), formative assessment is always a planned process and should never be done on the spur of the moment without thought. Formative assessments should be implemented frequently, and results should be produced and shared rapidly.

Finally, formative assessments should have a planned integration of the assessment with instruction (Frohbieter et al., 2011; Popham, 2008). The content of the formative assessment must be connected to the current curriculum that is being taught. Formative assessments are created in a variety of ways to address content. Assessments may address basics that are needed to function, particular skills (single or multiple), or certain benchmarks for various subjects. Regardless of the content, formative assessments must be purposefully implemented in such a way that evidence of student learning in certain areas is obtained, shared with students, and utilized to improve instruction. When learners receive feedback, they are aware of skills that are needed to improve achievement.

Stiggins (2006) argued that educators must be focused on every child and must use formative assessments to help all students in an effort to meet the standards set by the state. The use of formative assessments throughout the learning process provides students with opportunities to show what they know without the pressures that are associated with high-stakes tests (Clarke, 2005). The frequency of completing the formative assessments also motivates students to improve. However, the lack of formative assessments in the classroom is a detriment to at- risk students (Clarke, 2009).

Curriculum-Based Measurements

According to Deno (1985), CBM are among the best methods of formative assessment. CBM are specific measurement methods for assessing student progress over time and for identifying students in need of additional instructional support and/or further diagnostic testing (McGlinchey & Hixson, 2004; VanDerHeyden & Burns, 2005; VanDerHeyden et al., 2007). According to the literature, CBM have been used to screen basic skills, benchmark student performance, establish local norms, conduct brief experimental analysis, guide classroom instruction, and guide intervention implementation (Crawford, Tindal, & Stieber, 2001; Hintze & Silberglitt, 2005; Hosp & Hosp, 2003; Shapiro et al., 2006; Shinn, 2002; Simmons et al., 2002; Wiliam, 2005). With CBM, teachers are provided with information that can be used to guide instructional decision-making (Allsopp et al., 2008; Helwig et al., 2002; Jiban & Deno, 2007). By directly measuring the desired functional outcome of instruction, data from CBM inform teachers about their students' progress and allow teachers to adjust instruction promptly according to what the data suggests in terms of students' learning (Helwig et al., 2002; Jiban & Deno, 2007). CBM have been studied extensively and have been found to be a very reliable and valid process.

Administering CBM has been found to be easy, brief, and cost effective for schools and districts (Pemberton, 2003). CBM are also more sensitive to student growth and allow for continual progress monitoring (VanDerHeyden, Witt, & Naquin, 2003; Ysseldyke & Algozzine, 2006). With the use of CBM, not only are teachers able to monitor student progress over a period of time, but they can also set goals for students to improve (Stecker, Fuchs, & Fuchs, 2005). The scoring of CBM accounts for very little time in the process and provides rapid feedback. In monitoring student progress, teachers know immediately by graphical representation whether there is a class wide problem or if individual students are at risk (Hosp, Hosp, & Howell, 2007). The results that are provided are both easy for teachers to read and easy to communicate to parents (VanDerHeyden et al., 2003). CBM are useful for screening students who are at risk of failing, and they can also be used to develop interventions for students and to evaluate progress on the interventions. In addition, CBM can measure long-term effects by frequent tests utilizing equal level probes that are administered throughout the year (Fuchs, 2004).

Although CBM have many advantages, there are also weaknesses. CBM procedures have been criticized for their lack of identifying a disability and their inability to make inferences from national norms and situational bias (Stecker et al., 2005). CBM may not carry enough weight in defining a disability in terms of special education. A student with a mathematics deficit in one school may not have a deficit in another school. The disability may only be good in that particular environment or local context.

A CBM worksheet may be only as good as the school district that developed it. The curriculum could be good, poor, or culturally biased. Unless a school's curriculum closely resembles the national curricula, valid inferences cannot be made about students' performance to others across the nation (Stecker et al., 2005).

CBM may be subject to situational biases, especially timing effects. CBM relies on timing to measure fluency rates. During CBM, students' performance is timed. These timing procedures allow measurement of students' rates of academic behavior (e.g., word correct per minute or digits correct per minute). Because CBM procedures are rate- or fluency-based measures, these rate measures have been shown to be extremely sensitive to changes in academic skills (Evans-Hampton, Skinner, Henington, & McDaniel, 2002).

CBM in Reading

Research examining the use of CBM in reading is abundant. In the late 1980s, Marston (1989) examined the literature and summarized the results of several studies that examined the validity of oral reading rates on CBM in predicting reading skills. According to Marston, the correlation coefficients for these studies ranged from .63 to

.90. However, the research reviewed by Marston examined the relationship between more global measures of reading achievement and CBM. Another focus of research that is more related to the current study is the relationship between reading CBM and state standardized tests.

Several studies have examined the relationships between CBM and state tests (Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Shapiro et al., 2006; Silberglitt, 2005; Stage & Jacobsen, 2001). Most curriculum-based measurements of oral reading fluency (CBM-R) measure oral reading rates. According to Deno (2003), CBM-R assess the accuracy and the fluency with connected text. The number of words read correctly by a student is measured and compared to the number of words read correctly by others according to the school or district norm.

Stage and Jacobsen (2001), using a correlational research design, established reading CBM scores that could accurately predict which fourth-grade students would meet the standards of the Washington Assessment of Student Learning (WASL). In their study, students were assessed in oral reading fluency during the months of September, January, and May. The results of their study indicated that of the three reading CBM that were administered, the first administration of the CBM, conducted in September, had scores the most related to the scores on the WASL and could be used to predict whether or not fourth-grade students would pass the state test. Stage and Jacobsen (2001) found that CBM-R estimated and predicted scores on the WASL.

In a replicated study of Stage and Jacobsen (2001), McGlinchey and Hixon (2004) conducted a study on using CBM to predict performance on the Michigan Educational Assessment Program's (MEAP) fourth- grade reading assessment. The study took place at an elementary school in an urban school district for seven years (1994 –

2001). The researchers investigated the correlation and predictive value of CBM procedures for performance on the MEAP. A one-minute oral reading fluency probe was administered two weeks before the MEAP was given. A moderately strong relationship was found between the oral reading rates and performance on the MEAP. The positive predictive power of CBM was .74. Results from this study supported the use of CBM for monitoring reading progress and establishing which students were at risk for low reading skills and failing state tests.

McGlinchey and Hixon (2004) conducted another study in a kindergarten through sixth grade elementary school over an eight-year period to examine whether CBM reading probes could predict student performance on the MEAP. Students were administered reading probes two weeks prior to being giving the MEAP. The study found that 74% of the students who received a passing score on the MEAP were able to read 100 words per minute.

Crawford and colleagues (2001) also analyzed the relationship between oral reading rates on CBM and scores on the Oregon state test in reading and math. A group of blended second- and third-grade students from a rural school district in Oregon were administered three different oral reading probes on one day during the month of January. The average of the three scores was then correlated to the results from the reading and mathematics questions from the state test. A strong relationship between the CBM-R and the performance on both mathematics and reading standardized tests was found (Crawford et al., 2001). The number of words read correctly per minute on the second grade CBM was found to have a .66 correlation with the state reading assessment and a .53 correlation with the statewide mathematics assessment. The number of words read correctly per minute on the third grade probes was found to have a .60 correlation with the statewide reading assessment and a .46 correlation with the statewide mathematics assessment.

In a similar study, Good, Simmons, and Kame'enui (2001) also found reading CBM scores to be a good predictor of performance on the Oregon State Assessment and confirmed the findings of Crawford and colleagues (2001). The study assessed 1,439 students in kindergarten through third grades using CBM-R scores to predict their outcomes on the state test (Good et al., 2001). The results indicated that 96% of students who scored at the third-grade benchmark on the CBM were also able to meet reading expectations on the state test.

The relationship between oral reading fluency and outcomes on state tests was also examined in a study done by Powell-Smith (2004), who reported results from studies that examined the relationship between a CBM testing oral reading fluency and outcomes on statewide assessments in Colorado, Florida, Illinois, Michigan, Minnesota, North Carolina, Oregon, and Washington. These studies found the correlation between these two variables at the end of third and fourth grade ranged between .44 (Washington) and .79 (Illinois). On average, most studies reported correlations in the .60 to .75 range (Powell-Smith, 2004).

Wiley and Deno (2005) examined the relationship of CBM scores and Minnesota's state reading assessment scores and particularly examined whether adding a maze reading measure would help in predicting the performance on the state test. A multiple regression model was employed in an attempt to predict third and fifth grade English language learners' (ELLs) performance on the state's assessment using CBM for both oral reading fluency scores and maze scores. The maze is a multiple choice cloze procedure that is completed by selecting the correct word choice to complete a sentence.

Moderate to moderately strong correlations were found between both the CBM-R scores and maze scores when related to students' performance on the state wide assessment. Pearson correlations of .73 or higher were found between the maze task scores and the state assessment scores among ELLs in third and fifth grade. A correlation of .73 was found between the oral fluency task scores and the state assessment scores for third grade ELLs.

Another study examined the predictive validity of reading CBM scores as related to performance on the Minnesota Comprehensive Assessment (MCA; Silberglitt & Hintze, 2005). In this study, CBM-R probes were administered to students through first, second and third grades during the fall, winter, and spring. These scores were then compared to the students' performance on the MCA at the end of third grade. The results of this study revealed that CBM-R scores were strongly associated with performance on the MCA (Silberglitt & Hintze, 2005). CBM-R scores were found to be accurate and efficient in predicting those students who were likely to pass the reading portion of the MCA. Researchers noted that scores from the spring administration of the CBM-R showed a stronger correlation to scores the MCA.

Moreover, Kung (2007) used culturally and linguistically diverse students to compare the criterion validity of CBM on oral reading rates to predict students' performance on state tests. The study found that CBM scores on oral reading measures are moderately to highly correlated with scores on state tests, thus serving as an early indicator of students' performance on state standards tests.

Cabrera (2008) investigated the accuracy of Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency (ORF) on predicting state scores on the California Standardized Testing and Reporting (STAR) Program English-Language Arts scores. This study examined scores for 212 third- through sixth-grade ELL and English Only (EO) students, and found that DIBELS was a good indicator of how ELL students would score on the STAR. DIBELS ORF scores also predicted English-Language Arts scores for students who were deemed at risk for school failure.

Multiple studies have examined the use of CBM in predicting later measures of reading achievement. For the most part, the CBM that have been used have been measures of oral reading rate or fluency. The evidence from this extensive and increasing number of studies demonstrates that scores on CBM assessing oral reading rates is a significant predictor of scores on high-stakes tests in reading for several states. Consequently, fluency, as measured by students' oral reading rates, has an impact on reading comprehension scores as measured by standardized reading assessments.

CBM in Mathematics

Although research has found CBM to predict performance on state assessments, most of the studies have been centered on reading. There are limited studies that examine the relationship between CBM scores and scores on state tests in mathematics. Mathematics-CBM (M-CBM) may assess computational fluency relating to the grade level curriculum standards. Other forms of M-CBM include estimation tasks, concepts and applications, concept-grounded problems, and problem solving (Foegen, Jiban, & Deno, 2007; Fuchs et al., 1994; Helwig et al., 2002).

Foegen and Deno (2001) compared the results of scores from a CBM using mathematics estimation task with the scores from the California Achievement Test for students in Grades 6 through 8. This particular CBM was made up of 20 computational problems and 20 word problems involving basic operations. The correlation between the

CBM scores and the scores from the concepts subtest of the standardized achievement test was .55.

Helwig and colleagues (2002) examined the effectiveness of CBM mathematics concept tasks at predicting the scores of eighth-grade students on a computerized mock version of the Oregon state test that was provided by the Department of Education. The CBM and the computerized assessment were given within two to three weeks of each other. Results showed that the mathematics CBM scores were effective at predicting scores on the computer adaptive test. Scores from the CBM probe predicted with 87% accuracy the students who would meet the state standards for math. Helwig and colleagues (2002) concluded that the CBM grounded in mathematics concepts predicted scores on the state mathematics test for the entire group with .60 and for individual general education students with .53.

Shapiro and colleagues (2006) conducted a study to examine the relationship between CBM and standardized assessments (state and published norm-referenced standardized achievement tests –SAT-9, MAT-8) in Pennsylvania for students in third, fourth, and fifth grades. The areas researched included reading, mathematics computation, and mathematics concepts/application. Reading passages were from AIMS web, and mathematics computation and mathematics concepts/application were assessed by Monitoring Basic Skills Progress (MBSP). A moderate-to-strong relationship was found between scores on CBM reading measures and scores on state assessment measures with correlations at about .70. This was shown across third and fifth grades in both districts.

Christ and Schanding (2007) examined the influence of task novelty and incentives on group level multiple-skill M-CBM performance on 90 students in Grades 2 through 5. Three assessment conditions (novel, reward, and neutral) were examined over a three-day period. The novel condition involved students who had not been exposed to the assessment. The reward condition gave students the opportunity to receive a reward based on their performance. The neutral condition involved students who had taken the assessment before but did not receive a reward. On the first day, both reward and neutral groups out-performed the novel group (Christ & Schanding, 2007). The difference approximated three digits correct per minute, and there were no significant differences in performance between neutral and reward conditions. Christ and Schandling indicated that the M-CBM is an indicator of optimal performance after the initial day of assessment.

Although the literature for CBM and mathematics is not as extensive as it is in reading, various forms of M-CBM have a positive relationship to outcomes on state tests. Studies have compared scores on M-CBM to standardized tests, teacher ratings, and state tests. A strong relationship between the M-CMB scores and scores from other mathematics measures may provide teachers, schools, and districts with data relative to their overall mathematics proficiency.

Summary

The goal of NCLB is to meet state-standards in mathematics and reading with proficiency by 2013-2014 (Katsiyannis et al., 2007; Shapiro et al., 2006; Shriberg & Kruger, 2007; Silberglitt et al., 2006). According to NCLB, schools are responsible for tracking the progress of students through the use of progress monitors. The use of progress monitors ensures that students are mastering set skills. When implementing progress monitors, frequent assessments are required. To assess students frequently, formative assessments are used. Formative assessments have become very popular and are used as a means of gathering evidence about student learning (Heritage et al., 2009; Pinchok & Brandt, 2009). Regardless of the type or format of the formative assessment, the purpose, the cycle of use, and the plan of integration with instruction are needed for them to be effective (Frohbieter et al., 2011). There must also be follow through with the plan of action that must involve modifications in teaching and learning (Black & Wiliam, 1991).

One of the best methods of formative assessment is CBM. CBM are being used as a way to monitor students' progress toward proficiency (Helwig et al., 2002; Hintze & Silberglitt, 2005; Jiban & Deno, 2007; Jitendra et al., 2005). While studies support the use of CBM scores to predict outcomes on state reading assessments, there is very little research that examines the relationship between M-CBM scores and scores on state assessments in mathematics (Hintze & Silberglitt, 2005; McGlinchey & Hixon, 2004; Shapiro et al., 2006; Stage & Jacobsen, 2001; Wiley & Deno, 2005).

CHAPTER III

METHODOLOGY

The purpose of this study was to determine the relationships between students' scores of the three M-CBM and the scores from the mathematics MCT2 for sixth-grade students. This chapter describes the methods and procedures that were used in this study. Specifically, the following null hypotheses were tested:

1. There is no relationship between the August 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.
2. There is no relationship between the December 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.
3. There is no relationship between the April 2010 M-CBM scores and the May 2010 mathematics MCT2 scores.

This chapter consists of the following sections: research design, participants, instrumentation, procedure, and methods of data analysis.

Research Design

A correlational research design was used in this study. According to Gall, Gall, and Borg (2003), when the purpose of a research study is to predict outcomes, a correlational research design should be used. The authors also stated that correlational statistics are used to determine if a relationship exists between two or more quantifiable variables and if a strong relationship exists, then measures on one variable (the predictor variable) can be used to predict measures on another variable (the criterion variable).

Correlational research provides a correlation coefficient that is a numerical index of the direction and strength of the relationship between two or more quantitative variables. The coefficient ranges from -1 to 1. The closer to the absolute value of 1, the stronger the relationship. The positive and negative values identify the direction, or type of relationship. Positive values indicate that the scores on the two variables vary in the same direction and negative values indicate that the two scores vary in different directions. Specifically, positive coefficients indicate that as one measure increases, so does the other measure. Negative coefficients indicate that as one value increases, the other value decreases (Johnson & Christensen, 2008).

The correlational research design is the most appropriate design to fulfill the purpose of this study. The purpose of this study was to determine the relationships between the student scores on three M-CBM and the scores from the mathematics MCT2 for sixth-grade students. For this study, the M-CBM assessments served as the predictor variable and the MCT2 served as the criterion variable.

Participants

Scores were from students attending an intermediate sixth-grade school located in a school district in rural east central Mississippi. The school was accredited by the Southern Association of Colleges and Schools and was the only school in the district that served sixth graders. The structure of the school was divided into four teams which consisted of three teachers, an inclusion teacher, and between 72 and 77 students per team.

Student scores were from all 310 students enrolled in this school during the 2009-2010 school year. In the 2009-2010 school year, 64% of the students were African-American and 33% of the students were Caucasian. The remaining 3% of the student population were Asian, Native American, or Hispanic. Females accounted for 52% of the population and males accounted for 48% of the population.

Instrumentation

Archived consisting of three sets of student scores M-CBM and the 2009-2010 mathematics portion of the MCT2 were used in this study. The instruments used in this study were the M-CBM provided by STEEP and the mathematics portion of the MCT2.

M-CBM Probe

The mathematics probe that was purchased by the district from STEEP was the Mathematics Fluency-Multiple Skills in Multiplication and Division for grade 6 (STEEP, 2011). Sixth-grade schools were given the option of using either mathematics fluency assessments or a mathematics concepts and application assessment. According to J. Witt (personal communication, December 8, 2011), senior scientist for STEEP, the reliability of the Mathematics Fluency-Multiple Skills probe ranges from .90 to .94. The validity of the probe ranges from .55 to .60 coefficients.

In administering the probe, teachers followed a scripted set of procedures. The M-CBM is a two-minute assessment that was administered to an entire class at the same time. After every class was tested, the same trained teachers then scored the assessments. In scoring the mathematics assessments, students received one point for each correct number that was correctly positioned. This process is called scoring for digits correct (Witt & Chun, 2005). Scoring for digits correct rather than merely counting the number of correct problems provides a way to give students credit for solving more complex problems which may involve multiple steps. To score mathematics problems, teachers simply counted number of digits correct during the two-minute test. Digits were correct if the correct number was located in the proper place value for that problem. After all the M-CBM were scored, results were recorded in the data management software provided by STEEP.

Elementary benchmark assessment levels for the sixth-grade M-CBM on computation fluency by STEEP identifies three levels based on performance. Those levels are frustrational, instructional, and mastery. The frustrational level is the lowest level and ranges from a score of 0 - 39. The instructional level represents the grade level equivalent and ranges from a score of 40 - 79. The level representing mastery ranges from a score of 80 or higher (Witt & Chun, 2005).

Mississippi Curriculum Test, 2nd Edition (MCT2)

The MCT2 is a testing program aligned with the Mississippi Frameworks and is used in the Mississippi State Accountability System. The MCT2 consists of criterion-referenced assessments in reading/language arts, writing, science, and mathematics (MDE, 2008). These assessments are aligned with the Mississippi Frameworks that were revised for language arts in 2006 and for mathematics in 2007. Students in grades three through eight, including special education students whose Individualized Education Plan (IEP) specify the same goals, are required to take the MCT2.

The section of the MCT2 that was used in the present study is the sixth-grade mathematics assessment, which is untimed and administered by two people following standard directions and procedures. The mathematics portion of the MCT2 is administered over a one-day period. The assessment is made up of 60 multiple-choice questions that vary according to depth of knowledge and benchmarks.

According to MDE (2010):

> All Mathematics MCT2 assessments measure grade and content specific curriculum found in the 2007 Mississippi Mathematics Framework-Revised. The student mastery of grade-level curriculum for seventh grade students is measured based upon the following competencies: (1) 15 number and operations items, (2) 10 Algebra items, (3) 9 Geometry items, (4) 8 Measurements items, and (5) 8 Data Analysis and Probability items. There are 10 experimental items and 50 core items, which makes the total 60 items for the Mathematics segment of the MCT2. Scores for students, schools, districts, and the state of Mississippi are based on the core items only. (p. 8)

Different test forms are used for each administration of the MCT2; however, the number of core test items is identical on every form. According to the MDE (2010), multiple measures were used to account for the reliability and validity of the MCT2. Regardless of the form, the results of the assessment will be reported in one of four categories' performance descriptors (MDE, 2011). The performance descriptors for the sixth-grade mathematics portion of the MCT2 are listed in Table 1.

Table 1 Performance Level Descriptors for Sixth-Grade Mathematics MCT2

Proficiency Level	Scale Score Ranges
Advanced	The advanced scores for grade 6 mathematics range from 164 and above.
Proficient	The proficient scores for grade 6 mathematics range from 150-163.
Basic	The basic scores for grade 6 mathematics range from 142-149.
Minimal	The minimal scores for grade 6 mathematics range from 141 and below.

Note. Students who score advanced perform beyond the requirements for success in that content area (mathematics) for the 6th grade.
Students who score proficient demonstrate solid academic performance in that content area (mathematics) for the 6th grade.
Students who score basic demonstrate partial mastery in that content area (mathematics) for the 6th grade.
Students who score minimal inconsistently demonstrate the knowledge or skills for mastery for that content area (mathematics) for the 6th grade.

Procedure

Prior to requesting approval to conduct the study from the Institutional Review Board for the Protection of Human Subjects in Research (IRB) at Mississippi State University, permission was requested from the school district. After obtaining approval from district personnel to conduct the study, an application was submitted to the IRB of Mississippi State University. Upon IRB approval (see Appendix A), August, December, and April M-CBM scores and MCT2 scores were retrieved from the EZ Test Tracker system and exported to Microsoft Excel. The EZ Test Tracker is a web-based application provided by Educational Leadership Solutions (Shelly & Baer, 2003) and used by the school district to store all collected test data. Once the data were exported to Microsoft Excel, students who did not have scores for all four of the assessments were eliminated, then student names and MSIS numbers were removed. The data were then transferred to the Statistical Package for the Social Sciences (SPSS) for analysis.

EZ Test Tracker

EZ Test Tracker is a web–based application that is used by school districts to collect all test data and schedule information necessary for analyzing student and teacher performance on criterion- and norm-referenced test (Shelly & Baer, 2003). EZ Test Tracker is a product that is provided by ELS. The product is used by this particular district to navigate through test data by allowing the use of filters to view data based on subgroups, teachers, and student performance. The data base provides educators with reports, graphs, multiple years of state, district, and school level data, and information regarding the passing of state tests. The district uses EZ Test Tracker to store results from the CBM provided by STEEP, common assessments created by the district, and results from both the reading/language arts and mathematics MCT2 for grades three through eight and the SATP in U. S. History, English II, Algebra I, and biology for the state of Mississippi. With the school district's permission, EZ Test Tracker was used by the researcher to obtain data.

Data Analysis

After data from all three M-CBM and data from the mathematics MCT2 were gathered from EZ Test Tracker, the data were analyzed using SPSS 12.0. Pearson correlations were calculated as a preliminary analysis. These correlations examined the nature of the relationships between the M-CBM scores from August, December, and April of sixth-grade students during the 2009-2010 school year and scores from the mathematics portion of MCT2 administered in May of 2010.

Descriptive statistical analysis using means, frequencies, and percentages were used to describe the students. Descriptive statistics were also used to identify cases with missing data, to examine the dispersion of the data, and to assess for normality. If cases existed with missing data, these particular cases were eliminated from the analysis.

Summary

Chapter 3 described the methodology that was used to conduct the study. Three null hypotheses guided this study to determine if relationships existed between M-CBM scores and mathematics MCT2 scores of sixth-grade students who attended a school located in rural east central Mississippi during the 2009-2010 academic year. During the 2009-2010 school year, 310 students were enrolled in this school; however, there were 271 participants who were assessed all four times. Archived data from three M-CBM (August and December of 2009 and April of 2010) and the results from the 2010 mathematics MCT2 were retrieved from EZ Test Tracker. Using a correlational research design, data were analyzed at the .05 alpha level using the Pearson *r* correlation coefficient to test each hypothesis.

CHAPTER IV

RESULTS

Progress monitoring systems are used as formative assessments to guide classroom instruction and to monitor students' progress toward meeting academic standards. The existing scores from the M-CBM progress monitoring system provided by STEEP were used. Using this system, school personnel assess students in August, December, and April to gauge their progress toward meeting the standards measured by the mathematics MCT2. Although considerable time is allocated to administer the three M-CBM, little is known about the relationship between the scores on the M-CBM and the scores on the mathematics MCT2. The purpose of this study was to determine the relationships between the scores of three M-CBM and the scores from the mathematics MCT2 for sixth grade students. To fulfill the purpose of this study, three hypotheses were tested. This chapter first presents a descriptive summary of the scores on the measures (M-CBM scores and mathematics MCT2 scores) that provided the data for this study, then details the results of the analyses of data that were used to test each hypothesis. The chapter concludes with a summary of the major findings of this study.

Descriptive Summary of Measures

Data used in this study represent the M-CBM scores and the mathematics MCT2 scores of the 271 sixth grade students that attended the school. Scores on M-CBM are categorized into assessment levels and scores on the mathematics MCT2 proficiency levels. The assessment levels for the M-CBM are frustrational, instructional and mastery.

With each administration of the M-CBM (August, December, and April), the majority of students' scores were in the range of scores that are categorized as instructional. During the first testing cycle (August), 162 of the 271 students' scores corresponded to the instructional level category indicating that those students were performing as expected. Students' scores in the frustrational level indicated that students were not performing at the expected level. With each testing cycle, the number of students performing at that frustrational level decreased from 40 during the first cycle (August) to 22 at the last testing cycle (April). An inverse trend was observed with the number of students scoring at the mastery level (surpassed expectations). With each testing cycle, the number of students scoring at the mastery level increased from 69 at the first testing cycle to 112 at the last testing cycle. Table 2 displays the number and percentage of students scoring at each.

Table 2 M-CBM Assessment Levels

Assessment Level	**Testing Cycle**					
	August		**December**		**April**	
	N	**%**	**N**	**%**	**N**	**%**
Frustrational	40	15%	36	13%	22	8%
Instructional	162	60%	136	50%	137	51%
Mastery	69	25%	99	37%	112	41%
Total	**271**	**100%**	**271**	**100%**	**271**	**100%**

Likewise, each MCT2 mathematics score corresponds to one of the following proficiency levels: minimal, basic, proficient, and advanced. Of the 271 students' mathematics MCT2 scores, slightly over half (52%) of the scores were in the proficient or advanced proficiency levels, indicating that these students either met or exceeded expectations. The scores of the other 48% of students corresponded to either the minimal or basic level, indicating that they had not met expectations. Table 3 displays the number and percent of students' scores corresponding to each proficiency level.

Table 3 2010 Mathematics MCT2 Proficiency Levels

Proficiency Level	N	%
Minimal	74	27%
Basic	58	21%
Proficient	88	33%
Advanced	51	19%
Total	**271**	**100%**

Data Analysis

This section of Chapter 4 presents the results of the data analysis that were used to test the three hypotheses that guided this study. In addition to the results that were used to test the hypotheses, supplemental analyses were performed to facilitate a more thorough understanding of the relationships between M-CBM scores and mathematics MCT2 scores. Following the results of data analysis for each hypothesis are the results of the supplemental analysis by mathematics proficiency level.

Hypothesis 1

There is no relationship between the August 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.

A Pearson correlation coefficient was calculated to determine the relationship between the August M-CBM scores and the mathematics MCT2 scores. A moderate positive correlation was found (r (269) = .40, p =.00), indicating a significant linear relationship between the two variables. Students with high scores on the August M-CBM tended to have high scores on the mathematics MCT2. As a result of this finding, Hypothesis 1 is rejected. There is a positive, moderate relationship between the August M-CBM scores and the mathematics MCT2 scores.

To understand more thoroughly the relationship between the August M-CBM scores and the MCT2 scores, a Pearson correlation coefficient was calculated to determine if the relationship between the scores continued to exist when examined by mathematics MCT2 proficiency levels. When examined by proficiency level (minimal, basic, proficient, and advanced) , the only statistically significant relationship found was a weak relationship between the August M-CBM scores and the mathematics MCT2 scores of students scoring in the proficient range, (r (86) = .21, p =.04). No statistically significant relationships were found between the August M-CBM scores and the mathematics MCT2 scores of students in the advanced range (r (49) = .16, p =.26), basic range (r (56) = -.07, p = .58), or the minimal range (r (72) = .05, p =.66). Consequently, it appears that the relationship found between the August M-CBM scores and the mathematics MCT2 scores is mainly attributed to the scores of students who scored in the proficient range. Table 4 displays the results of this analysis.

Table 4 August M-CBM and Mathematics MCT2 Correlations by Proficiency Levels

	N	*df*	*r*	r^2	*P*
MCT2 Scores	271	269	.40	.16	.00*
Minimal	74	72	.05	.00	.66
Basic	58	56	-.07	.00	.58
Proficient	88	86	.21	.04	.04*
Advanced	51	49	.16	.03	.26

*p<.05

Hypothesis 2

There is no relationship between the December 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.

A Pearson correlation coefficient was calculated to determine the relationship between the December M-CBM scores and the mathematics MCT2 scores. A moderate positive correlation was found (r (269) = .43, p =.00), indicating a significant linear relationship between the two variables. Students with high scores on the December M-CBM tended to have high scores on the MCT2. As a result of this finding, Hypothesis 2 is rejected. There is a positive, moderate relationship between the December M-CBM scores and the mathematics MCT2 scores.

Further analysis was conducted examining the relationship between the December M-CBM scores and the mathematics MCT2 scores by proficiency level. No statistically significant relationships were found between the December administration of the M-CBM and the mathematics MCT2 scores of students in the advanced range (r (49) = .19, p =.18), proficient range (r (86) =.20, p =.07), basic range (r (56) = -.003, p = .98), or the minimal range (r (72) = .02, p =.90). A moderate, statistically significant relationship was found between the December M-CBM scores and the mathematics MCT2 scores for the total group (N = 271). When relationships between December M-CBM scores and mathematics MCT2 scores by proficiency level were computed, none of the relationships were found to be statistically significant. Therefore, it appears that the relationship between the December M-CBM scores and the mathematics MCT2 scores was not associated with any specific mathematics MCT2 proficiency level. Table 5 displays the results of the analysis of data examining relationships between December M-CBM scores and mathematics MCT2 scores.

Table 5 December M-CBM and Mathematics MCT2 Correlations by Proficiency Levels

	N	*df*	*r*	r^2	*P*
MCT2 Scores	271	269	.43	.18	.00*
Minimal	74	72	.02	.00	.90
Basic	58	56	-.003	.00	.98
Proficient	88	86	.20	.04	.07
Advanced	51	49	.19	.04	.18

*$p<.05$

Hypothesis 3

There is no relationship between the April 2010 M-CBM scores and the May 2010 mathematics MCT2 scores.

A Pearson correlation coefficient was calculated to determine the relationship between the April M-CBM scores and the MCT2 scores. A moderate positive correlation was found (r (269) = .47, p = .00), indicating a significant linear relationship between the two variables. Students with high scores on the April M-CBM tended to have high scores on the MCT2. As a result of this finding, Hypothesis 3 is rejected. There is a positive, moderate relationship between the April M-CBM scores and the mathematics MCT2 scores.

To explore this relationship further, Pearson r correlations were computed to determine if there were relationships between the April M-CBM and mathematics MCT2 scores by proficiency level. The only statistically significant relationship found was a weak relationship between the April M-CBM scores and the MCT2 scores of students scoring in the proficient range, (r (86) = .22, p = .04). No statistically significant relationships were found between the August M-CBM scores and the MCT2 scores of students in the advanced range (r (49) = .05, p = .71), basic range (r (56) = .04, p = .78), or the minimal range (r (72) = .01, p = .92). Consequently, it appears that the relationship found between the April M-CBM scores and the mathematics MCT2 scores was mainly attributed to the scores of students who scored in the proficient range. Table 6 displays the results of this analysis.

Table 6 April M-CBM and Mathematics MCT2 Correlations by Proficiency Levels

	N	*df*	*r*	r^2	*P*
MCT2 Scores	271	269	.47	.22	.00*
Minimal	74	72	.01	.00	.92
Basic	58	56	.04	.00	.78
Proficient	88	86	.22	.05	.04*
Advanced	51	49	.05	.00	.71

$^*p<.05$

While the M-CBM scores of all three assessments (August, December, and April) were related to the mathematics MCT2 scores, the April scores had the strongest relationship to the mathematics MCT2 scores ($r = .47$). To follow up on this finding, a multi-linear regression was calculated to determine if a significant prediction model could be developed that could be used to predict students' MCT2 scores based on their M-CBM scores from August, December, and April. A significant regression equation was found (F (3, 267) = 25.24, $p < .001$), with an r^2 of .22. In other words, 22% of the variation in mathematics MCT2 scores is explained by the variation in the M-CBM scores. According to the prediction model, students' predicted MCT2 score is equal to 123.41 + .01(August M-CBM scores) + .05(December M-CBM scores) + .32(April M-CBM scores). Students' MCT2 scores increased .01 for each point on the August M-CBM, .05 for each point on the December M-CBM, and .32 for each point on the April M-CBM. However, only the April M-CBM scores were significant predictors of MCT2 scores.

Summary

In determining if relationships existed between M-CBM administered three times during the 2009-2010 school year and the mathematics MCT2 administered in 2010, three hypotheses were tested. The results from all three hypotheses indicated that there was a moderate positive correlation between each administration of the M-CBM and the mathematics portion of the MCT2, with the strongest relationship found between the April M-CBM scores and the mathematics MCT2.

Further analysis was done to determine if the relationship between M-CBM scores and mathematics MCT2 scores continued to exist when examined by mathematics MCT2 proficiency levels. In the follow-up analysis examining Hypothesis 1, a weak relationship was found between the August M-CBM and the mathematics MCT2 scores of students scoring in the proficient range. No significant relationships were found at the minimal, basic, or advanced levels. The results of the analysis of data for Hypothesis 2 revealed that while a moderate, statistically significant relationship was found between the December M-CBM scores and the mathematics MCT2 scores for the total group ($N = 271$), when relationships between those variables by proficiency level were computed, none were found to be significant. Results of the analysis of data for Hypothesis 3 also resulted in a rejection of the null hypothesis. Further analysis revealed a weak statistically significant relationship between the April M-CBM scores and the MCT2 scores of students scoring in the proficient range. No statistically significant relationships were found between the April M-CBM scores and the MCT2 scores of students in the advanced, basic, or minimal ranges. The final supplemental analysis, multi-linear regression, resulted in a significant linear regression model that could be used to predict students' mathematics MCT2 scores. Table 7 displays the major findings of all data analyses.

Table 7 Major Findings of the Study

	Hypothesis 1 August M-CBM		Hypothesis 2 December M-CBM		Hypothesis 3 April M-CBM	
	r	*P*	*r*	*p*	*r*	*P*
All Mathematics MCT2 Scores	.40	.00*	.43	.00*	.47	.00*
Minimal MCT2 Scores	.05	.66	.02	.90	.01	.92
Basic MCT2 Scores	-.07	.58	-.003	.98	.04	.78
Proficient MCT2 Scores	.21	.04*	.20	.07	.22	.04*
Advanced MCT2 Scores	.16	.26	.19	.18	.05	.71

*$p<.05$

CHAPTER V

SUMMARY, CONCLUSIONS, RECOMMENDATIONS

This chapter presents the summary, conclusions, and recommendations for the study. The first section provides a brief summary of the previous chapters (the introduction, the literature review, the methodology, and the data analysis), including the research hypotheses. The second section focuses on the findings of the study and provides the conclusions according to the research hypotheses. The second section also discusses the limitations of the study. The final section discusses implications for practice and recommendations for practice and for future research.

Summary

School districts across the United States spend thousands of federal dollars on researched-based strategies, materials, and programs that have demonstrated effectiveness in increasing student achievement (Simpson et al., 2004). The deadline for the goal of having all students become proficient, as measured by standardized testing, in reading, language arts, and mathematics as set by NCLB is approaching the 2014 mark (Katsiyannis et al., 2007; Robelen, 2002; Simpson et al., 2004). With high-stakes tests being given at the end of the year, schools must utilize a student progress monitoring system (US DOE, 2001). Though progress monitoring systems are required by NCLB, the time that it takes to screen students and the cost associated with them are concerns.

CBM are being used as one method of monitoring students' progress toward proficiency (Helwig et al., 2002; Hintze & Silberglitt, 2005; Jiban & Deno, 2007;

Jitendra et al., 2005). The data from CBM are immediately available to teachers and should be used to guide instruction. While several studies support the use of CBM to predict outcomes on state reading assessments (Hintze & Silberglitt, 2005; McGlinchey & Hixon, 2004; Shapiro et al., 2006; Stage & Jacobsen, 2001; Wiley & Deno, 2005), there is very little research that examines the relationship between mathematics CBM and state assessments in mathematics.

The purpose of this study was to determine the relationships between the scores of three assessments of the M-CBM and the scores from the mathematics portion of the MCT2 for sixth-grade students. Of the 310 students enrolled at the sixth-grade school in rural Mississippi, scores for 271 students were used in the study. Results from the August, December, and April scores of the M-CBM during the 2009-2010 school year and scores from the mathematics portion of the MCT2 that was administered in May 2010 were retrieved from the ELS data management system. A correlational research design was used to fulfill the purpose of the study and to test the following three research hypotheses:

1. There is no relationship between the August 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.
2. There is no relationship between the December 2009 M-CBM scores and the May 2010 mathematics MCT2 scores.
3. There is no relationship between the April 2010 M-CBM scores and the May 2010 mathematics MCT2 scores.

Pearson correlation coefficients were calculated to determine the relationships between the MCT2 scores from May 2010 and the three scores of the M-CBM from August, December, and April of the 2009-2010 school year. The results of the analyses indicated that scores on the mathematics MCT2 were positively related to scores on each of the M-CBM. Moreover, the results also indicated that the strength of the relationship increased with each M-CBM. However, the differences in the strength of the relationships were small. The relationships between the mathematics MCT2 and the August M-CBM, the December M-CBM, and the April M-CBM were .40, .43, and .47 respectively.

A moderate positive correlation was found for all three null hypotheses, indicating a significant linear relationship between the August, December, and April M-CBM scores and the mathematics MCT2 scores. The scores on the August M-CBM were found to have a .40 correlation with the mathematics MCT2 scores. The December M-CBM scores were found to have a .43 correlation with the mathematics MCT2. Scores from the April M-CBM were found to have a .47 correlation with the mathematics MCT2 scores. Students with high scores on the M-CBM tended to have high scores on the mathematics portion of the MCT2. Of the August, December, and April M-CBM scores, the April scores had the strongest relationship with scores on the mathematics portion of the MCT2.

Further analysis was conducted examining the relationships between M-CBM scores and mathematics MCT2 scores by proficiency levels. No significant relationships were found between the measures for students who scored in the minimal, basic, or advanced levels on any of the three assessments. A significant but weak relationship between the mathematics MCT2 scores and the August and April M-CBM scores of students scoring at the proficient level was found. There was no significant relationship found with students scoring proficient on the MCT2 and the December M-CBM scores.

Conclusions

The results of this study indicate that the scores on the M-CBM of the sixth-grade students in this district were related to the scores on their mathematics MCT2. The results also indicate that with each M-CBM, the magnitude of the relationship increased. The correlation coefficients found in this study ranged from .40 to .47, indicating moderate relationships. The findings of this study are somewhat consistent with the findings reported by Crawford and colleagues (2001), Foegen and Deno (2001), and Stage and Jacobsen (2001).

Crawford et al. (2001) found that scores on a reading CBM were related to the scores from the California state mathematics assessment for a group of second- and third-grade students. In their analysis, they also found moderate relationships (.53 for second graders and .46 for third graders) between a reading CBM and the mathematics state assessment. The reading CBM used was a measure of fluency in that scores were determined by the number of words read correctly per minute. On the surface, it would appear that the relationships found between reading fluency and mathematics were chance relationships or relationships that just happened to appear. However, when the totality of skills measured by the state mathematics assessment is considered, the relationships discovered were meaningful. In addition to basic computational skills, the state assessment included math word problems. Surely, being a fluent reader would have a positive effect on students comprehending the word problems prior to determining the correct computational skill to apply. Unlike the present study, Crawford et al. (2001) only examined relationships between a reading CBM and the state mathematics assessment. But even when examining just one relationship, the correlation coefficient was just as high or higher than the coefficients of the present study. Also unlike the present study, reading fluency in relation to mathematics was examined. The M-CBM used in this study is a measure of computational fluency, as scores represent the number of digits correct in two minutes. The M-CBM fails to account for measures of reading fluency which impact a student's ability to comprehend the word problems that are on the mathematics MCT2. Therefore, while the findings for the present study and the study conducted by Crawford et al. (2001) are consistent, the results of both studies signify the need to use a more comprehensive CBM which includes measures of both reading fluency and computational fluency.

Foegen and Deno (2001) found a correlation coefficient of .55 when they examined the relationship between an M-CBM and the California Achievement Test scores. The CBM used by the authors measured both mathematics and reading fluency for a group of students in Grades 6 through 8. With a coefficient of .55, the percentage of variance in the state assessment scores explained by the variance in the CBM scores was 30%, which is higher than the percentage of variance in the present study. The percentage of explained variance in the study was only 22%, and that was for the (April) M-CBM. If the M-CBM are being used appropriately, that is to monitor learning and modify instruction, it would seem logical that by the third score of the M-CBM, more than 22% of the variance would be explained. With a coefficient of determination of .22, 78% of the variance in MCT2 scores is unexplained by the variance in M-CBM.

Stage and Jacobsen (2001) administered a CBM assessing for oral reading fluency in September, January, and May to 173 fourth graders and used a growth curve analysis to determine the relationship to the Washington Assessment of Student Learning (WASL). Like the present study, the CBM was given during three periods in the school year and the relationships being examined included scores on the CBM and state assessments. The difference was that the present study assessed for computational fluency instead of oral reading fluency. Stage and Jacobsen (2001) predicted May WASL reading performance using slope in oral reading fluency probes that were administered during three months. Unlike the present the study, the results of their study indicated that of the three reading CBM-R the first assessment (September) was the most related to the scores on the WASL and could be used to predict whether or not fourth-grade students would pass the state assessment. The present study found that the third administration (April) had the strongest relationship (.47) to the mathematics MCT2. Stage and Jacobsen (2001) found that the positive predictive power that the first CBM-R (September) for low scores predicted WASL failure was .41, and the negative predictive power for the first administration for high scores predicted WASL success was .90.

The studies conducted by Foegen and Deno (2001), Crawford et al. (2001), and Stage and Jacobsen (2001) resulted in findings very similar to the present study, in that all of the studies found relationships between CBM scores and state assessment scores. Although the studies had similarities, each study shared various differences when compared to the present study.

Limitations

The study does have some limitations that require careful consideration of the findings. First, the study was conducted in one school district in Mississippi. Scores for sixth-grade students were used. There is a possibility that the scores of this district would not be fully representative of other districts within the state of Mississippi. Additional replication of the findings of this district's data across other districts in Mississippi is needed.

Secondly, only students who had all data sets (M-CBM for August, December, April, and MCT2) were included in the analysis. Some attrition from the August sample of M-CBM data to the sample of MCT2 data was inevitable. Such attrition can potentially result in a final-analysis sample that is somewhat different from the original sample and presents limitation to interpreting data.

Finally, this study offers findings related to performance in mathematics fluency. Since this is one of few studies to report these types of data, it is unknown whether the findings here are similar to what would be found in other types of studies like this. Further studies examining M-CBM performance and state-wide assessment measures are certainly needed.

Implications of the Study

This study contributes to the literature related to M-CBM and state tests. This study also contributes to the evidence regarding the positive correlation between CBM in general and state tests.

CBM are reliable, valid, and flexible. School districts should consider the continuation of administering the M-CBM three times a school year, if not more frequently. A potential implication of this study is the possible use that M-CBM may have in serving as an effective screening device that predicts outcomes on statewide assessment measures. The findings of this study demonstrated that a two-minute mathematics probe assessing mathematics fluency in multiplication and division when administered three times during the 2009-2010 school year (August, December, and April) had a moderate positive significant relationship to the mathematics portion of the MCT2 that was administered in May of 2010. The August, December, and April schedule of the M-CBM in this study represents a likely scenario for gathering benchmark data that identifies students who may need interventions. Because the third scores of the M-CBM were from April and had the strongest relationship to the MCT2 of the three sets of scores, teachers have approximately one month remaining in the school year before the MCT2 is given at the end of the school year. The remaining time could be used to guide classroom instruction as well as provide interventions to at-risk students. Not only can CBM be used as a means to respond to intervention, but they can also be used to monitor the effects of enrichment being provided to high achieving students. This could reach students who are "on the bubble," or points away from scoring in the next level. With the demands of increasing the number of students scoring at proficient levels, teachers also have data that could pinpoint students who could possibly score minimal, basic, proficient, or advanced.

The moderate positive relationships that were found in the present study provide great information to school districts. Students who scored low on the M-CBM tended to score low on the MCT and could possibly score at the minimal or basic levels, which in turn could lower the Quality Distribution Index. Students who scored high on the M-CBM tended to score high on the MCT and could possibly score at the proficient and/or advanced levels, which could raise the Quality Distribution Index for the school. Those students scoring in the middle could possibly be on the border and could end up scoring either in the minimal category or the proficient category. Again, this could result in lowering or raising the Quality Distribution Index for the school. Although the use of CBM can be costly and time consuming, the use of results could be what schools need to earn higher accountability levels.

General Recommendations

NCLB requires school districts to have progress monitors in place; however, there is a variety of progress monitoring systems available. Progress monitors may come in the form of CBM, common assessments that measure benchmarks, or both. These tools can be purchased from a company or created by districts, schools, or teachers. These tools can assess a single skill, multiple skills, concepts, or applications. Some districts choose to utilize these tools more frequently than others. A moderate positive significant relationship was found between M-CBM and MCT2 scores when the M-CBM assessed mathematics fluency in multiplication and division. On the MCT2, 30% of the questions found in numbers and operation could possibly cover multiplication and division. School districts should consider the type of CBM that may be best for what they are assessing in the end on the state test. The district used in this study had the choice of using an M-CBM that tested concepts and applications. The M-CBM that assesses concepts and applications may be a better predictor for this district and grade level.

This district chose to administer the M-CBM three times during the school year. Because the relationship continued to get stronger with scores from each administration, giving the assessment more frequently could produce a stronger relationship. Because time and money are already issues, the district may chose to only increase the administration times of students who score in the bottom 16th percentile.

The most important aspect of CBM is how a teacher uses the data. Schools hoping to increase the number of students performing at the proficient or advanced levels should also consider requiring teachers to create an action plan for all students based on the results of CBM scores. This plan should not only address individual interventions, but it should also target class-wide instructional modifications.

Recommendations for Further Studies

Based on the findings of this study, the following are recommendations for further research. First, this study involved scores from one sixth-grade school in one district with 310 students in Mississippi. Future research should be broadened to include more grade levels and possibly more school districts. Further research should also consider the various types of CBM and the frequency of their administration. The M-CBM used at the school which generated data for this study was purchased from the STEEP system which offered a different type to sixth-grade schools in concepts and applications. Districts have begun to create their own formative assessments that are used to monitor progress toward benchmark assessments. If formative assessments that are district-created can predict outcomes on state tests, it is possible that thousands of dollars could be saved each year.

REFERENCES

Algozzine, R., Ysseldyke, J., & Elliott, J. (1997). *Strategies and tactics for effective instruction* (ed.). Longmont, CO: Sopris West.

Allsopp, D., Kyger, M., Lovin, L., Gerretson, H., Carson, K., & Ray, S. (2008). Mathematics dynamic assessment: Informal assessment that responds to the needs of struggling learners in mathematics. *Teaching Exceptional Children, 40*, 6–16.

Batsche, G., Elliot, J., Graden, J. L., Gimes, J., Kovaleski, J.F., Prasse, D., … Tilly III, W. D. (2005). *Response to intervention: Policy considerations and implementation*. Alexandria, VA: National Association of State Directors of Special Education, Inc.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-73.

Black, P. & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability* (pp. 20-50). Chicago: National Society for the Study of Education.

Black, P. & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 92*(1), 81-90.

Brooks, J. G., & Brooks, M. G. (1999*). In Search of Understanding: The Case for Constructivist Classrooms*. Alexandria, VA: ASCD - Association for Supervision and Curriculum Development.

Bruner, J. S. (1966). *Toward a theory of instruction.* Cambridge, MA: Belknap Press.

Buchanan, J. (2007). *The impact of the high-stakes Mississippi Curriculum Test on teachers' instructional practices*. Buchanan, J. ProQuest Dissertations and Theses 2007. Section 0132, Part 0288 121 pages; [Ph. D. dissertation]. United States – Mississippi: Mississippi State University; 2007. (Publication Number: AAT 3255995)

Burch, P. (2006). The new educational privatization: Educational contracting in the era of high stakes accountability. *Teachers College Record, 88*(2), 129-135.

Burns, M., VanDerHeyden, A., & Jiban, C. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review, 35*, 401-418.

Cabrera. C. (2008). *Curriculum-based measurement as a predictor for California standards tests.* ProQuest Dissertations and Theses [M.S. dissertation]. California State University, Fresno. (Publication Number: AAT 1460337)

Christ, T. & Schanding, G. (2007). Curriculum-based measures of computational skills: A comparison of group performance in novel, reward, and neutral conditions. *School Psychology Review, 36*(1), 147-158.

Clarke, S. (2005). *Formative Assessment in the Secondary Classroom*. London: Hodder Murray.

Clarke, S. (2009). Using curriculum-based measurement to improve achievement. *Principal, January/February 2009.* Retrieved from http://www.naesp.org/resources/2/Principal/2009/J-F_p30.pdf

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide tests. *Educational Assessment, 7*, 303-323.

Deno, S. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232.

Deno, S. (2003). Curriculum-based measures: Development and perspectives. *Assessment for Effective Intervention, 28*(3-4), 3-12.

Educational Testing Service. (2006, May 30). Balanced assessment systems: Redefining excellence in assessments. Portland, OR: R. J. Stiggins. Retrieved March 29, 2011, from http://www.edzone.net/~mascd/Events/06BootCamp/Cohort1-2007/Stiggins_Balanced_Assessment_Systems.pdf

Evans-Hampton, T., Skinner, C., Henington, C. & McDaniel, C. (2002). Effects of inconspicuous timing during curriculum-based measurement of mathematics: An investigation of situational bias. *School Psychology Quarterly, 31*, 529-539.

Foegen, A., & Deno, S. (2001). Identifying growth indicators for low achieving of students in middle school mathematics. *The Journal of Special Education, 35*(1), 4-16.

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41*(2), 121-139.

Frohbieter, G., Greenwald, E., Stecher, B. & Schwartz, H. (2011). *Knowing and doing: What teachers learn from formative assessment and how they use the information.* (CRESST Report 802). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*(2), 188-192.

Fuchs, L., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*, 199-208.

Fuchs, L.S., Fuchs, D., Hamlett, Thompson, A., Roberts, P.H., Kubek, P., Stecker, P.M. (1994). Technical features of a mathematics concepts and applications curriculum-based measurement system. *Diagnostique, 19*(4), 23-49.

Gall, M. D., Gall, J. P. & Borg, W. R. (2003). *Educational research: An introduction* (5th ed.). Boston: Allyn and Bacon.

Geary, D. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*, 4-15.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance of decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high stakes outcomes. *Scientific Studies of Reading*, 5(3), 257-288.

Guthrie, J.W. & Springer, M.G. (2004). A Nation at Risk revisited: "Wrong" reasoning sometimes results in "right" rules? *Peabody Journal of Education*, 79(1), 7–35.

Hadley, K. (2005). *Mathematics anxiety of elementary teachers and its effect on student mathematics achievement.* Ph.D. dissertation, Utah State University, United States—Utah. Retrieved April 14, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3200363)

Helwig, R., Anderson, L., & Tindal, G. (2002). Using a concept-grounded, curriculum-based measurement in mathematics to predict statewide test scores for middle school students with LD. *The Journal of Special Education, 36*(2), 102-112.

Heritage, M., Kim, J., Vendlinski, J., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practices, 28*(3), 24-31.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review*, *34*(3), 372-386.

Hosp, M., & Hosp, J. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure, 48*(1), 10-17.

Hosp, M., Hosp, J., & Howell, K. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement.* New York: Guilford.

Ikeda, M. J., Neesen, E. & Witt, J. C. (2008). Best practices in universal screening. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (pp.721-734). Bethesda, MD: National Association of School Psychologists.

Jiban, C. L., & Deno, S. L. (2007). Using mathematics and reading curriculum-based measurements to predict state mathematics test performance: Are simple one-minute measures technically adequate? *Assessment for Effective Intervention, 32*(2), 78-89.

Jitendra, A., Sczesniak, E., & Deatline-Buchman, A. (2005). An exploratory validation of curriculum-based mathematical word problem-solving tasks as indicators of mathematics' proficiency for third graders. *School Psychology Review, 34*(3), 358-371.

Johnson, B., & Christensen, L. (2008). *Educational Research: Quantitative, Qualitative and Mixed Approaches*. Los Angeles, CA: Sage Publications.

Johnson, D. D., & Johnson, B. (2002). *High stakes: Children, testing, and failure in American schools.* Lanham, MD: Rowan & Littlefield.

Katsiyannis, A., Zhang, D., Ryan, J. B., & Jones, J. (2007). High-stakes testing and students with disabilities. *Journal of Disability Policy Studies*, *18*(3), 160-167.

Ketterlin-Geller, L., McCoy, J., Twyman, T., & Tindal, G. (2003). How do critical thinking measures fit within standards-based reform? *Assessment for Effective Intervention, 28*(3&4), 37-48.

Kung, S. (2007). *Predicting the success on a state standards test for culturally and linguistically diverse students using curriculum-based oral reading measures.* Kung, S. ProQuest Dissertations and Theses [Ph.D. dissertation]. University of Minnesota. (Publication Number: AAT 3275053)

Marston, D. (1989). A curriculum-based measurement approach to assessing academic performance: What is it and why do it. In M. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford.

McGlinchey, M. & Hixson, M. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*(2), 193-202.

Mississippi Department of Education (MDE). (2002). *Mississippi's plan for student achievement: Assessment, achievement, accountability.* Retrieved September 25, 2008, from http://www.mde.k12.ms.us/extrel/pub/saread.pdf

Mississippi Department of Education (MDE). (2008). *Mississippi curriculum test, second edition: Technical manual for 2007-2008 test administration.* Retrieved July 17, 2008, from http://www.mde.k12.ms.us/acad/osa/technical/MCT2_TechReport_2008.pdf

Mississippi Department of Education (MDE). (2009). *Mississippi curriculum test, second edition: Technical manual for 2008-2009 test administration.* Retrieved September 12, 2010. from http://www.mde.k12.ms.us/acad/osa/technical/MAAECF/MAAECF%20TechReport%20Manual%202008-2009.pdf

Mississippi Department of Education (MDE). (2010). *Mississippi curriculum test, second edition: Interpretive guide for grades3-8 language arts and mathematics mct2.* Retrieved August 3, 2010, from http://www.mde.k12.ms.us/acad/osa/mct2/2010_IG.pdf

Mississippi Department of Education (MDE). (2011). Mississippi assessment and accountability reporting system. *Office of Research and Statistics.* Retrieved February 7, 2011, from http://orasp.mde.k12.ms.us/MAARS/maarsprintableMCT2.jsp?usersessionId=48

National Assessment of Educational Progress (NAEP) mathematics report for Mississippi. Retrieved September 25, 2008 from http://www.mde.k12.ms.us/OSA/NAEP/Reports/Math_Snapshot_Mississippi_4.pdf

National Center for Education Statistics. (2009). *The national report card: Mathematics 2009 (NCES 2010-451).* Institute of Education Sciences, U.S. Department of Education, Washington, DC.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

No Child Left Behind Act of 2001. Pub. L. No. 107-110.

Olson, L. (2002). Testing systems in most states not ESEA-ready. *Education Week, 1*, 26-27.

Pemberton, J.B. (2003). Tips for communicating academic progress. *Teaching Exceptional Children, 35*(4), 16-22.

Piaget, J. (1965). *The moral judgment of the child.* NY: Free Press.

Pinchok, N., & Brandt, W. C. (2009). *Connecting formative assessment research to practice*. Washington, DC: Learning Point Associates. Retrieved from http://www.learningpt.org/pdfs/FormativeAssessment.pdf

Popham, W.J. (2008).*Transformative assessment.* Alexandria VA: ASCD.

Powell-Smith, K. A. (2004, February). *Individual differences in FCAT performance: A national context for our results.* Paper presented at the annual meeting of the Pacific Coast Research Conference, Coronado, CA.

Robelen, E. (2002). An ESEA primer. *Education Week on the Web, 21*(16), 28-29.

Russell, S.J. (2000). Developing computational fluency with whole numbers. *Teaching Children Mathematics, 7*, 154-158.

Shapiro, E., Keller, M., Lutz, J., Santoro, L., & Hintze, J. (2006). Curriculum based measures and performance on state assessment and standardized tests: Reading and mathematics performance in Pennsylvania. *Journal of Psychoeducational Assessment, 24*(1), 19-35.

Shinn, M. R. (2002). Best practices in using curriculum based measurement in a problem-solving model. In A, Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 671-698) Bethesda, MD: National Association of School Psychologists.

Shelly, P., & Baer, K. (2003). EZ Test Tracker. [Data management system]. Ridgeland, MS: Educational Leadership Solutions, Inc.

Shriberg, D. & Kruger, L. J. (2007). Introduction and overview high stakes testing. *Journal of Applied School Psychology, 23*(2), 1-6.

Silberglitt, B., Burns, M. K., Madyun, N. H., & Lail, K. E. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *Psychology in the Schools*, *43*(5), 527-535.

Silberglitt, B. & Hintze, J. (2005). Formative assessment using CBM-R cut scores to track progress toward success on state-mandated achievement tests: A comparison of methods. *Journal of Psychoeducational Assessment, 23*, 304-325.

Simmons, D. C., Kame'enui, E. J., Good, R. H. III, Harn, B. A., Cole, C., & Braun, D. (2002). Building, implementing, and sustaining a beginning reading improvement model: Lessons learned school by school. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 537-569). Bethesda, MD: National Association of School Psychologists.

Simpson, R. L., LaCava, P. G., & Graner, P. S. (2004). The No Child Left Behind Act: Challenges and implications for educators. *Intervention in School & Clinic*, *40* (2), 67-75.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 20*, 407-419.

Stecker, P., Fuchs, L., & Fuchs, D. (2005). *Using curriculum-based measurement oral reading fluency performance standards to predict success on local and state test of reading achievement.* Paper presented at the annual conference of the National Association of School Psychologists, Washington, DC.

System to Enhance Educational Performance. (2007). *System to Enhance Educational Performance (STEEP): About us.* Retrieved August 17, 2009, from http://www.isteep.com/about_us.html

System to Enhance Educational Performance. (2011). *System to Enhance Educational Performance (STEEP): News and events.* Retrieved August 17, 2009, from http://isteep.com/news_and_events.html

U.S. Department of Education. (2001). *The Elementary and Secondary Education Act (The No Child Left Behind Act of 2001)*. Retrieved September 28, 2008, from http://www.ed.gov/legislation/ESEA02

U.S. Department of Education. (2009). *Mathematics 2009: National Assessment of Educational Progress at Grades 4 and 8.* Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2009 Mathematics Assessments.

VanDerHeyden, A. M., & Burns, M. K. (2005). Using curriculum-based assessment and curriculum-based measurement to guide elementary mathematics instruction: Effect on individual and group accountability scores. *Assessment for Effective Intervention, 30*, 15-31.

VanDerHeyden, A. M., Witt, J., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology, 45*, 225-256.

VanDerHeyden, A. M., Witt, J., & Naquin, G. (2003). Development and validation of a process for screening referrals to special education. *School Psychology Review, 32*, 204-227.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children, 69*, 391-409.

Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes.* Cambridge, MA: Harvard University Press.

Webb, M. (2007). *The functional outcomes of curriculum-based measurement and its relation to high-stakes testing.* [Ph. D. dissertation]. ProQuest Dissertations and Theses.University of Houston. (Publication Number: AAT 3263299)

White House. (2001). *No Child Left Behind: Transforming the federal role in education so that no child is left behind.* Washington, DC: The White House.

Wiley, H., & Deno, S. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26*, 207-214.

Wiliam, D. (2005). *Keeping learning on track: Formative assessment and the regulation of learning in making mathematics vital: Proceedings of the twentieth biennial conference of the Australian Association of Mathematics Teachers*, (pp. 20-34). Sydney, Australia: Australian Association of Mathematics Teachers.

Wiliam, D. & Thompson, M. (2008). "Integrating assessment with learning": What will it take to make it work? In C. A. Dwyer, Ed., *The Future of Assessment: Shaping Teaching and Learning,* 53-82. New York: Lawrence Erlbaum Associates.

Witt, J., & Chun, M. (2006). *STEEP assessment manual*. Miami, FL: iSTEEP Learning.

Ysseldyke, J. E., & Algozzine, B. (2006). *Assessment for students with special needs*. Thousands Oaks, CA: Corwin Press.

Ysseldyke, J. E. & McLeod, S. (2007). Using technology tools to monitor response to intervention. *Handbook of Response to Intervention*, IV, 396-407, doi: 10.1007/978-0-387-49053-3_29

APPENDIX A

NOTIFICATION OF APPROVAL TO CONDUCT RESEARCH

December 19, 2011

Eutrophia Lenora Hogan-Samuel
1384 Louisville St. Lot 12
Starkville, MS 39759

RE: IRB Study #11-353: Using Math Curriculum Based Measures to Predict Outcomes on the Math Portion of the Mississippi Curriculum Test, 2nd Edition

Dear Ms. Hogan-Samuel:

This email serves as official documentation that the above referenced project was reviewed and approved via administrative review on 12/19/2011 in accordance with 45 CFR 46.101(b)(4). Continuing review is not necessary for this project. However, any modification to the project must be reviewed and approved by the IRB prior to implementation. Any failure to adhere to the approved protocol could result in suspension or termination of your project. The IRB reserves the right, at anytime during the project period, to observe you and the additional researchers on this project.

Please note that the MSU IRB is in the process of seeking accreditation for our human subjects protection program. As a result of these efforts, you will likely notice many changes in the IRB's policies and procedures in the coming months. These changes will be posted online at http://www.orc.msstate.edu/human/aahrpp.php.

Please refer to your IRB number (#11-353) when contacting our office regarding this application.

Thank you for your cooperation and good luck to you in conducting this research project. If you have questions or concerns, please contact me at nmorse@research.msstate.edu or call 662-325-3994.

Sincerely,

Nicole Morse
Assistant Compliance Administrator

cc: Debra Prince (Advisor)