# A COMPARISON OF GEORGIA CRITERION REFERENCED COMPETENCY TEST MATH SCORES BETWEEN AT-RISK FIFTH GRADE STUDENTS RECEIVING COMPUTER BASED MATH INSTRUCTION AND AT-RISK STUDENTS NOT RECEIVING COMPUTER BASED MATH INSTRUCTION

by

Katherine Tomaszewski Arnold
Liberty University

A Dissertation Presented in Partial Fulfillment
Of the Requirements for the Degree
Doctor of Education

Liberty University
October, 2013

# A COMPARISON OF GEORGIA CRITERION REFERENCED COMPETENCY TEST MATH SCORES BETWEEN AT-RISK FIFTH GRADE STUDENTS RECEIVING COMPUTER BASED MATH INSTRUCTION AND AT-RISK STUDENTS NOT RECEIVING COMPUTER BASED MATH INSTRUCTION

By
Katherine Tomaszewski Arnold

A Dissertation Presented in Partial Fulfillment
Of the Requirements for the Degree
Doctor of Education

Liberty University, Lynchburg, VA

October, 2013

APPROVED BY:

Rick Bragg, Ed.D.
Committee Chair

Gary Kimball, Ed.D.
Committee Member

Beverly Smith, Ph.D.
Committee Member

Scott Watson, Ph.D.
Associate Dean, Advanced Programs

# ABSTRACT

Katherine Tomaszewski Arnold, A COMPARISON OF GEORGIA CRITERION REFERENCED COMPETENCY TEST MATH SCORES BETWEEN AT-RISK FIFTH GRADE STUDENTS RECEIVING COMPUTER BASED MATH INSTRUCTION AND AT-RISK STUDENTS NOT RECEIVING COMPUTER BASED MATH INSTRUCTION (Under the direction of Rick Bragg, Ed.D.). School of Education, Liberty University, October, 2013.

This causal comparative research study compared the mean score differences from two groups of fifth grade students. Both groups were identified at-risk, based on a curriculum based math measure and their 4th grade performance on the state assessment, for failure to meet math standards on the state assessment. The curriculum based measure in this study was the 2011-12 AIMSweb Mathematic Concepts and Applications (M-CAP). Mathematics achievement was measured by scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test (CRCT). Each group received classroom instruction using Learning Focused Schools (LFS) math strategies. The treatment group received additional computer based math instruction. The computer based math instruction in this study was Individual Prescription for Achieving State Standards (iPASS). A correlation analysis examined the ability of the 2011-12 M-CAP fall cut scores to identify risk status measured by student scores on the 2012 CRCT math subtest. A second correlation analysis examined the predictive ability of the 2011-12 winter M-CAP target scores on the CRCT as indicated by the strength and direction of the relationship. Three research questions were investigated: (a) Is there a difference in the 2012 CRCT math scores between at-risk fifth grade students who participate in computer based math instruction and those who do not? (b) Are the 2011-12 fall M-CAP cut scores able to identify fifth grade students as at-risk to fail the math portion of the 2012 CRCT? (c) Are the 2011-12 winter M-CAP target scores able to predict fifth grade students'

performance on the math subtest of the 2012 CRCT? To address these research questions this study first examined the effectiveness of computer based math concepts and applications instruction on student achievement and its viability as an effective intervention for students identified as at-risk to fail end of year state assessments. Secondly, this study investigated the ability of curriculum based math measures, specifically concepts and applications, to identify students as at-risk for failure. And finally it investigated the ability of target scores to predict achievement outcomes on state assessments. Results of this study will produce additional data to help facilitate informed instructional decisions.

## Dedication

This dissertation is dedicated to my family. My husband Jeff has been very supportive and willing to take on more than his share. Jeff, without your patience and generous spirit this would not be possible. I could not have done this without you. My sister Rubi was always there to offer advice. Sis, thanks for your willingness to listen and your encouragement to persevere. Your pep talks kept me motivated and lifted my spirits.

Acknowledgements

First and above all I wish to thank our Lord, Jesus Christ, for His guidance in this endeavor. He led me to a very supportive and generous dissertation committee. My gratitude goes out to my committee members because I could not have done this without their suggestions and patience. Thank you Dr. Bragg, Dr. Kimball and Dr. Smith; each of you is greatly appreciated. Thank you Dr. Holder for your input and direction; each suggestion made my study stronger. Thank you to the Douglas County School District and the system principals for allowing me to conduct my research and to collect data. Heartfelt thanks to Dr. Besson-Martilotta, she had her own journey but always had time to offer support and guidance.

Table of Contents

List of Tables

List of Figures

## CHAPTER ONE: INTRODUCTION

This research study investigated the ability of a math curriculum based measure (CBM) to identify risk status and to predict performance outcomes on state assessments. Cut and target scores derived from the AIMSweb M-CAP were used to identify students who were at-risk to fail or were on target to pass the math portion of the state assessment respectively. Through the statistical method of receiver operating characteristics or ROC curve analysis, scores were established by test developers to identify those students presenting with low risk, some risk, at-risk, and those students demonstrating math skills of proficiency to meet state standards (AIMSweb 2009b). The focus of this study utilized M-CAP scores from a sample of fifth grade students identified as at-risk because based on this status select schools provided computer based math interventions in an attempt to remediate the math deficits. The fifth grade general education students were identified as at-risk on the fall M-CAP and also did not meet proficiency on the previous year's 2010-2011 CRCT math subtest. The identified students were designated for differentiated instruction in order to address their risk status and to increase their level of math proficiency. The differentiated instruction examined in this study was the use of a computer based math curriculum (iPASS). Student scores on the 2012 math portion of the Georgia Criterion Referenced Competency Test were analyzed from the fifth grade general education students identified as at-risk who participated in the computer based math instruction and at-risk fifth grade students who did not. In two separate analyses fall M-CAP cut scores from the remaining fifth grade general education students identified as at-risk were examined to determine their ability to identify risk for failure on the CRCT, and winter M-CAP target scores from fifth grade general education students

were examined to determine their ability to predict performance on the math subtest of the 2011-12 CRCT.

A review of research literature suggested that the strength of CBM is in its ability to identify skill deficits, to monitor growth, and to provide teachers with information to prescribe academic interventions. Studies have also indicated that CBM has strong criterion validity making it an effective tool to predict outcomes on state assessments (Hintze and Silberglitt, 2005).

The first step in accurate assessment of student achievement is an examination of test-generated data. If interventions are being prescribed based on the data from these measures, it is of utmost importance to determine the ability of CBM scores to identify risk status and to effectively monitor and predict scores on state assessments. Evidence of a positive relationship between the scores on the math curriculum based measure and scores on the end of year state assessment will strengthen the use of CBM as a tool to identify students that are at risk and to provide interventions based on information from these measures.

Chapter one includes background information to provide the framework for the use of curriculum based measures and computer based instruction. It discusses the purpose for this study through problem statements. The significance of this study is: (a) to compare the achievement outcomes made by at-risk fifth grade general education students provided computer based math instruction versus the outcomes made by students not provided computer based instruction, (b) to investigate the ability of fall CBM math cut scores to accurately identify students as at-risk to fail state assessments, (c) to investigate the ability of winter CBM math target scores to predict fifth grade students'

educational outcomes on state assessments, (d) to add to the scant amount of research examining the strength and direction of the relationship of mathematic curriculum based measures to state assessments, and (e) to examine the educational benefits of computer based math instruction as a method to increase math proficiency for at-risk students. Research questions and hypotheses anchor the study and a discussion of the quantitative design is included. The research study is summarized and limitations and assumptions with implications and recommendations for future study are stated.

**Background**

School reform remains an integral and prominent piece of the education landscape. School systems are under federal mandates to provide students with standards based instruction and to reduce the number of students who are not successful with grade level curricula, especially in mathematics and science. With the National Commission on Excellence in Education's publication of the report *A Nation at Risk* (1983), the education standards movement was launched. Over time, school districts across the nation responded by developing state based academic standards which encompassed basic skills that every student should master. These performance standards would later translate into the blueprint that school districts and teachers used to develop state based academic scope and sequence of instruction.

Information from the National Center for Educational Statistics indicated that American students had not demonstrated a competitive knowledge and proficiency of math and science curricula since the 1960s. The Center's statistics indicated that a mere 40% of fourth graders and 35% of eighth graders were scoring at proficiency in math in 2009 (National Center for Education Statistics, 2009). Statistics such as these have

compelled school districts to set guidelines for precisely what students should know at each grade level to increase math and science proficiency. The performance standards would encompass the curricula that would enable American students to compete in a global economy. With performance standards set, measures for alignment and accountability were developed and implemented. Measures of accountability were usually in the form of state assessments which were developed to align with the standards. School districts and teachers were held accountable for delivering the appropriate, differentiated instruction necessary to ensure that students were able to demonstrate competence on summative assessments.

In accordance with No Child Left Behind (NCLB, 2001) mandates, school districts began developing their own standards based criterion assessments to administer at the end of each academic school year. In order to monitor progress toward acquisition of skills that students would need to perform with proficiency on yearly assessments, school districts continued to seek valid instruments that were not time consuming but still had utility and were cost effective as research based formative measures to document the critical increments of student progress (Linn, Baker, & Betetebenner, 2007).

Curriculum-based measures were developed at the University of Minnesota's Institute for Research on Learning Disability (IRLD) over 40 years ago. Stanley Deno worked with his colleagues to develop an accurate and efficient means to document and assess the effects of instruction (Deno, 2003). Validation of his efforts came in 1978 when he and the Minnesota IRLD received a contract from the US Bureau of the Educationally Handicapped to expand the research into the effectiveness of curriculum based measures (Shinn, 2008).

Shinn (1989) asserted that formative assessments, such as CBM, were useful in providing pictures of incremental changes in student skills and progress. These tools were quick and simple "academic thermometers" that could monitor students' acquisition of skills and content that were relevant for positive achievement outcomes. The information provided by these instruments, often in the manner of formative benchmark, target, or cut scores enabled teachers to modify their teaching methods to differentiate instruction for students and to provide baseline and growth measures to chart interventions.

Data from general outcome measures, or curriculum based measures, can also be used to identify students who may be likely to fail summative state assessments (Shapiro, Keller, Lutz, Santoro, & Hintze, 2006). Based on information from their research they concluded that curriculum based measures of math calculation and applications did indeed have strong predictive validity with final state outcome measures. But as also indicated in their study, the research needs to be expanded and more studies should look at the ability of these measures to accurately identify risk status.

Research conducted by Jitendra, Sczesniak and Deatline-Buchman (2005) indicated that accuracy of identifying students who may be at-risk for failure of third grade concepts and applications on state math assessments increased with the addition of information from progress monitoring. This suggests that these tools may be used in combination as well as examined separately to determine the ability of each administration to accurately monitor and provide a skills acquisition picture of the student from which to consider and prescribe interventions.

Research conducted by Jiban and Deno (2007) indicated that simple, one minute curriculum-based measures alone, did not demonstrate the technical aspects necessary to monitor students' progress and academic growth. Third and fifth grade students were given three measures. One consisted of traditional basic facts and another used a cloze procedure requiring greater understanding of math facts since the missing number was placed in various positions. The third CBM was a maze reading measure. Performance on each CBM was analyzed separately and aggregated to determine the relationship of the scores on state assessments. They reported that reliability and criterion validity were not sufficient for the simple, one minute math measures but when combined with the maze reading CBM predicted performance could be improved by as much as 52%. Their research suggests that brief measures alone may not be adequate predictors of performance on state assessments and that combining them with measures of literacy may result in greater predictive validity.

According to Hosp, Hosp and Howell (2007), the value of CBM data lies in their brevity, simplicity, ability to facilitate academic screening, enable progress-monitoring, and aid in formulating diagnostic and outcome decisions. CBM administration takes approximately ten minutes, is quick to score, and interpretation is straightforward. Based on these characteristics many school districts across the nation have adopted these measures finding the scores useful to track student progress, align teacher instruction, and provide information on the predictive relationship of these scores to state assessments. Such information can identify some of the most important factors to consider when developing methodology to increase student achievement and performance on critical state assessments.

The No Child Left Behind (NCLB, 2001) legislation emphasized the need for all students to receive a quality education taught by highly qualified teachers. This was especially critical for those students identified as being at-risk of failure to meet grade level standards as demonstrated on end of year state assessments. Teachers were required to identify students at risk and document Response to Intervention (No Child Left Behind, [RtI], 2001) data through the provision and monitoring of research based interventions and differentiated instruction. These interventions and subsequent data must include progress information derived from formative assessments prior to consideration of special education referrals (No Child Left Behind, [RtI], 2001). Since these data are the critical first step in providing teachers with the information they need to differentiate instruction and to provide opportunities for each of their students to gain mastery on grade level standards, it is imperative that formative, curriculum-based measures are strongly correlated with and predictive of state assessments.

According to Clarke and Shinn (2004), an essential piece in preventing difficulties in math is to identify risk status early and to provide interventions to those students who demonstrate math deficits and are at-risk for later failure. This dissertation looked at the ability of curriculum based measures to identify risk status and through progress monitoring to predict student outcomes on state assessments.

In the current context, Georgia is one of 10 states that have been granted an ESEA Flexibility Waiver. Responding to a call from the United States Department of Education, Georgia applied for this waiver in order to gain more flexibility in methods to measure accountability and ways to assist schools that were struggling (ESEA-Flexibility Waiver, 2011). According to the United States Department of Education (ESEA

Flexibility, 2011) the No Child Left Behind legislation has been enacted as law for the past 10 years. However, the law should have been reviewed and rewritten in 2008 but little progress has been made to reauthorize the legislation. Until NCLB can be reauthorized, the ESEA Flexibility Waiver is a linkage between NCLB mandates and refocused reforms. As a result, it allows state and local education systems more flexibility with some of the requirements mandated in the No Child Left Behind legislation. It is important to note that this flexibility waiver absolutely continues to recognize the importance of annual gains made by students and thereby supports annual testing by summative assessments which are aligned to state standards. In addition, it continues to endorse the significance of making accountability decisions based on the incremental growth and progress made by students and the development of plans for improvement based on methodology which is research based and cost effective (ESEA-Flexibility, 2011). In order to address these two mandates, school systems must be able to accurately identify those students who do not have the skills to meet state standards and to be successful on summative state assessments. Using curriculum based measures that can accurately identify students that are at-risk and can also predict performance on state assessments is the first step in providing timely interventions that students may need to ensure positive educational outcomes.

In light of this new initiative, it remains imperative that state and local districts have efficient and predictive curriculum based measures that are valid and can accurately identify students who may need differentiated instruction. As a result, these measures must be precise in their ability to predict achievement outcomes on summative assessments. Accordingly, the methods for differentiation are critical and should be worth

the investment in instructional time to ensure that students are receiving the most effective interventions necessary to make crucial annual gains.

Differentiated instruction can take many forms. Hall, Strangman, and Meyer (2007) indicated that to differentiate instruction teachers are required to be skilled and flexible enough to individualize the curriculum to the learning needs of the student. According to Hall, et al. (2007) this amounts to meeting students at their learning level and assisting them in the learning process by providing the instruction they need to be successful. This oftentimes includes the ability to recognize the need for supplemental instruction when necessary to access the curriculum. This can be accomplished through the provision of specialized curricula, strategy building, and one method that has been around since the 1960s, computer based instruction.

Lowe (2002) noted that during the 1980s, the computer was viewed as the solution for all educational problems dealing with instruction. Her review of several meta analyses on the effect size of educational outcomes of computer based instruction indicated that, when compared with general classroom instruction, computer based instruction did have a positive effect on student achievement. In addition, with its increased sophistication the use of computers to aid instruction changed from a behaviorist orientation to a cognitive approach. This indicated a belief that students must not only know how to manipulate facts but should also understand underlying concepts of a task first, and this understanding can be facilitated by interacting with the computer.

Kaput and Thompson (1994) discussed the changes in mathematical computer based curriculum and instruction when it shifted from basic procedures in arithmetic to problem solving and reasoning. This may have been the point where more recent

programs and software were being developed to focus on math algorithms taught at specific grade levels. This change may have resulted in a renewed interest in computer based instruction.

In a review of research literature conducted by Gersten, Clarke, and Mazzocco (2007), the teaching of cognitive strategies for math problem solving, which can be used to relate content and to solve a variety of problems, was identified as one of the most essential and effective instructional methods. While fact drill and practice remains important, it should not be the primary focus of math instruction. When students are taught strategies that are applicable in different contexts they may be more successful with unfamiliar measures of math achievement.

According to Gersten, et al. (2007) most prior research on math instruction identified the terms procedural knowledge and conceptual knowledge as distinct. They concluded that at some point, even a simple arithmetic fact requires conceptual understanding of the embedded mathematical principles or procedures and thereby establishes the interconnection of the two. This research study investigated the math achievement of students who were provided computer based math instruction designed to build a framework of conceptual knowledge that they could use and generalize to problem solving on state assessments.

Williams (2000) conducted a thesis study investigating the effectiveness of minute math drills. His results coincided with previous findings that students who used computers to practice scored higher on tests because, it appeared, that math facts became more automatic and were easier to retrieve when problem solving. This ease of retrieval

allowed students to spend more time on the conceptualization of the algorithm needed to solve the problem.

Wilson, Majsterek, and Simmons (1996) conducted a study on the acquisition of multiplication facts using two instructional formats: teacher delivered instruction and computer based instruction. They specifically looked at response opportunities and student success. Results indicated that while teachers provided more response opportunities, the computer was more effective in developing students' skills in automaticity.

Today with the increase in classroom technology and demands on teacher time, the use of effective computer software is a viable addition to classroom math instruction. As a result, many educators are turning to technology and computer software to assist with differentiating instruction for their students. Select schools in this study were using an innovative computer based math curriculum which is built around strategy instruction as a method to differentiate instruction and to increase students' skills in math problem solving.

Fifth grade general education students who did not make a passing score on the 2010-11 math subtest of the state assessment were identified at the beginning of the 2011-12 school year. These students, along with the entire student population, participated in the fall administration of mathematic curriculum based measures (M-CAP). The students who did not pass the 2010-11 CRCT and also scored below average on the M-CAP were identified as at risk to fail the 2012 CRCT. Students in two schools were provided an additional overlay of computer based math instruction. This researcher

investigated the ability of the computer based instruction to increase at-risk students' math achievement on state assessments.

Another type of differentiated math instruction is ROPES, a strategy recommended by the Learning Focused Schools ([LFS], 2007) model. The public school district for this study has adopted the LFS model. All teachers in the district, including those in this study, have attended implementation workshops and professional development classes. In addition, they were expected to implement the strategies outlined in the LFS model which included the use of differentiated instruction and lessons. All students in the district were taught according to the LFS framework.

The theoretical framework on which this study is based are schema theory and transfer of learning. According to Carreher and Schleimann (2002) theories on transfer of learning and use of previous schema to problem solve in novel situations are related to the study of relationships between time-delay testing measures. They examined the effects of prior knowledge and its influence on learning. As a result, the authors proposed the theory that learning does not always need to rely on context, which can change, because when schema is used in new contexts learning continues across situations.

Related research by Chan (2010) investigated the effect of repeated testing on retention and the ease and accuracy of time delay recall of information. Specifically, he investigated the long-term effects of testing on subsequent recall. He found that encoding and delay between first and last test were important considerations when measuring recall of tested information. Overall results indicated that repeated testing and retrieval enhanced and strengthened retention and transfer of learned material.

**Problem Statement**

Educators continue to look for tools that are not time consuming to administer and are still accurate in predicting achievement outcomes for their students. Evidence of alignment with state assessments and predictive ability are critically important if teachers are to differentiate instruction based on information from these measures. A body of research including Cusumano (2007), Clark and Shinn (2004), Helwig, Anderson, and Tindal (2002), Hintz and Silberglitt (2005), and Hosp, Hosp, and Howell (2007) indicated that curriculum based measures (CBMs) were efficient and useful tools that could increase the effects of instruction. They also possessed qualities that made them suitable for monitoring academic progress and interventions.

Deno (1985) proposed that CBM could be an effective tool for identifying risk status, monitoring progress toward acquisition of standards, and predicting outcomes on state tests.

Fuchs and Deno (1991) stated that the tasks developed for curriculum based measures were general outcome measures. These measures encompassed global content rather than sub-skills within the academic domains. Instruments with this measurement capacity can ensure that skills were sampled as part of a year long curriculum scope and sequence as well as providing a sampling of student skills at specific points during the year.

Curriculum based reading measures have a long history of research into alignment and utility as diagnostic tools. Curriculum based measurement of mathematics

(M-CBM), until recently, has not been examined to any great extent (Thurber, Shinn, & Smolkowski, 2002).

One area addressed in the current research study is the paucity of research investigating the relationship of M-CBM, which focuses on concepts and applications, to state assessments. When teachers use designated math cut scores as the basis to identify at-risk students and differentiate mathematics instruction through provision of interventions, there is value in knowing the strength and nature of the predictive relationship between the cut scores and state assessments. Oftentimes, based on this information system-wide interventions for differentiating instruction are prescribed and monitored. As a result, it is educationally and economically critical to know the accuracy of the predictive ability of progress monitoring measures. This researcher investigated the ability of the AIMSweb fall cut scores to identify risk status and the winter target scores to predict achievement outcomes on the CRCT.

Differentiation for math deficits through computerized math instruction, which focused on the use of math strategies to solve problems, was also examined. According to information from the National Council of Teachers of Mathematics ([NCTM], 2012), students should be taught specific strategies for problem solving that can reduce errors and offer greater insight into the process of computations. Accordingly, Lynch (2006) asserted that computers and specialized software can extend learning by encouraging a conceptual approach to teaching (problem solving) rather than a procedural (fact drill) one.

Based on this alignment between curriculum based measures and differentiated instruction, the primary focus of this research study was the comparison of math

achievement on state assessments from fifth grade general education students, identified as at-risk to fail state assessments, who received computer based mathematics instruction and fifth grade general education students, also at risk to fail, who did not receive the computer based math instruction. Secondly, based on fall CBM cut scores students were identified as at-risk. In order to prescribe interventions in a timely manner, it is important to know the ability and accuracy of these scores to identify risk status. And finally, the winter curriculum based target scores were used to monitor student progress. As a result, information concerning the accuracy of winter target scores to predict performance outcomes is another critical element in prescriptive educational planning.

**Purpose Statement**

The purpose of this dissertation was threefold: (a) to test the theory of transfer of learning by investigating two principles from Haskell's theoretical framework for general transfer, (b) to expand contemporary research literature that investigates the ability of computer based math software, specifically in problem solving, to increase math achievement and, (c) to add to the paucity of research that investigates the ability of math curriculum based measures, specifically concepts and applications, to identify risk status and to predict math performance.

This researcher accomplished this by comparing 2011-12 math subtest CRCT scores from 26 at-risk fifth grade general education students who participated in computer based math instruction to the CRCT math subtest scores of 26 at-risk fifth grade students who did not participate. The independent variable was generally defined as Group 1 which was comprised of at-risk fifth grade students who received iPASS computer based math instruction, and Group 2 at-risk fifth grade students who did not

receive computer based math instruction. The dependent variable was defined as scores on the math subtest of the 2012 Georgia Criterion Referenced Competency Test for each of the two groups.

In addition since curriculum based fall cut scores were one criterion used to determine risk status, this study investigated the ability of the 2011-12 AIMSweb M-CAP cut score to identify students as at-risk to meet performance standards on the 2012 math CRCT. The final analysis examined the relationship between the 2011-12 winter M-CAP target scores to the scores on the math subtest of the 2012 CRCT and determined the strength and direction of the relationship.

**Significance of the Study**

The Douglas County School District has incurred a considerable expenditure in funds and personnel to purchase and implement two web-based computer programs. One program was for student assessment and data analysis; the other for supplemental mathematics instruction. Generated data were used to plan and prescribe research based interventions to address academic deficits. What should subsequently be determined is the ability of these formative CBMs to identify risk status and to predict performance outcomes on end of year state assessments. In addition, since these measures provided data used to differentiate instruction, the effectiveness of that instruction should be examined prior to full implementation.

Results from this research study generated data which may assist other school districts to make sound judgments when allocating funds for assessment tools designed to identify students who are at-risk to meet state curriculum standards. This researcher accomplished this by analyzing the relationships between the two CBM scores; the M-

CAP fall cut scores and the M-CAP winter target scores, to the math scores on the 2011-12 CRCT. This attempted to determine the ability of the curriculum based measures to identify risk status and to predict performance outcomes. This study was significant because it added to the research on mathematic CBM which investigates whether it's an effective tool aligned to end of year state assessments and able to identify risk and predict student outcomes on state assessments.

As a corollary, this study provided information concerning the viability and effectiveness of computer based math instruction as an effective and instructionally prudent intervention for students identified as at-risk to fail. This researcher accomplished this by comparing the 2012 CRCT math scores of at-risk students whose math instruction was differentiated through the use of a computer based math program versus those students who were identified as at-risk but did not participate. Both groups received classroom math instruction utilizing LFS strategies. The significance of this portion of the study provided useful information for subsequent analyses into system wide implementation of the computer based math program.

**Research Questions**

The research questions for the study are:

Research Question 1**:** Is there a difference in the math CRCT scores between at-risk fifth grade students who participate in computer based math instruction and those who do not?

Research Question 2: Are the fall M-CAP cut scores able to identify fifth grade students as at-risk to fail the math subtest of the CRCT?

Research Question 3: Are the winter M-CAP progress monitoring target scores able to predict fifth grade students' performance outcomes on the math subtest of the CRCT?

**Hypotheses**

The following are the research hypotheses:

Research Hypothesis-1**:** Fifth grade students identified as at-risk who participate in computer based math instruction compared to at-risk fifth grade students who do not participate in computer based math instruction will have statistically significant different scores as measured by the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

Research Hypothesis-2: At-risk fifth grade students' M-CAP fall cut scores will have a statistically significant relationship to their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

Research Hypothesis-3: Fifth grade students' M-CAP winter progress monitoring target scores will have a statistically significant relationship to their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

The following are the null hypotheses:

Null Hypothesis 1-$H_{o1}$: There is no statistically significant difference in at-risk fifth grade students' achievement scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test between students who participate in computer based math instruction and those who do not.

Null Hypothesis 2-$Ho_2$: There is no statistically significant relationship between at-risk fifth grade students' M-CAP fall cut scores and their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

Null Hypothesis 3- $Ho_3$: There is no statistically significant relationship between fifth grade students' M-CAP winter progress monitoring target scores and their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

**Identification of Variables**

The dependent variable for research question one was math achievement for Group 1 and Group 2 at-risk fifth grade students. The dependent variable for research question two was math achievement for all fifth grade students identified as at-risk. The dependent variable for research question three was math achievement for all fifth grade students identified as on target. The dependent variable for each hypothesis was math achievement as measured by the math subtest scores on the 2012 Georgia Criterion Referenced Competency Test (Georgia Department of Education, [CRCT], 2012).

The independent variable for research question one was group participation in a computer based math program (Group 1) and non-participation (Group 2). Computer based math program was defined as iPASS "a computer-based, web-enabled, math curriculum" (iLearn, 1989 p.3).

The independent variable for research question two was the fall curriculum based math cut scores from all fifth grade students identified as at-risk. The independent variable for research question three was the winter curriculum based math target scores from all fifth grade students identified as on target. Math curriculum based measures (M-CBM) were defined as the fall and winter AIMSweb M-CAP (AIMSweb, 2009a).

M-CAP fall cut scores were defined as less than 8 on math concepts and applications and were used to identify students as at-risk to fail the math portion of the state assessment. M-CAP target scores of 10 or greater were those scores derived at the mid-point which were used to predict performance outcomes on state assessments. Scores of 10 or greater at midpoint were considered on target to meet state standards (AIMSweb, 2009a).

**Definitions**

*AIMSweb*. A research based screening and progress monitoring tool for use by school systems as a part of a Response to Intervention (RtI) solution. (AIMSweb, 2009a).

*Computer Based Instruction (CBI*). Any type of programmed instruction that is presented through computer software. (Ormrod, 1995). Computer based instruction is known by various other terms: computer assisted, computer mediated and computer enabled. For purposes of this study computer based instruction will be used.

*Criterion Referenced CompetencyTest (CRCT).* An end of year achievement test designed to measure state standards in Georgia. (Georgia Department of Education, 2012).

*Curriculum-Based Measurement (CBM).* A progress monitoring measure of individual student growth in reading, spelling, writing, and mathematics. (AIMSweb, 2009a).

*Edformation*. A privately held company offering AIMSweb. (Harcourt, 2006).

*ESEA Flexibility Waiver*. A deferment to states of some No Child Left Behind (NCLB) educational and accountability mandates. (ESEA Flexibility, 2011).

*General Education Students*. Students in Douglas County Public Schools who are not identified as special education.

*iPASS*. A computer-based, web-enabled math curriculum marketed by iLearn. (iLearn, 1989).

*Learning-Focused Schools (LFS)*. A comprehensive school improvement model organized into a framework of teaching strategies to increase student achievement. (Thompson, 2007).

*No Child Left Behind (NCLB)*. Legislation passed by Congress in 2001 to monitor student achievement data. (No Child Left Behind Act of 2001).

*Receiver Operating Characteristics (ROC)*. The diagnostic accuracy or performance of a test to discriminate within a decision threshold.

*Response to Intervention (RtI)*. A system of early screening and progress monitoring of students' growth to improve student outcomes. (No Child Left Behind Act of 2001).

**Research Summary**

This research examined the ability of fall CBM to identify risk status. Students identified as at-risk were prescribed interventions. This study investigated the effectiveness of computer based math interventions. Academic progress was monitored through CBM. This study analyzed the predictive ability of winter CBM to identify students on target to meet performance standards.

This non-experimental study was quantitative and exploratory. It utilized an ex post facto causal comparative analysis as well as correlation analyses.

Archived achievement data from the 2012 CRCT math subtest was compared for Group 1 and Group 2 students using analysis of covariance while controlling for the students' achievement scores on the 2011 CRCT math subtest. This portion of the study

provided additional research that specifically investigated instruction in math strategies, concepts, and applications to the field of research into computer based math instruction.

Correlation analyses were used to determine the ability of the 2011 fall AIMSweb M-CAP cut scores to identify students as at-risk to fail the 2012 math subtest of the CRCT. All fifth grade general education students identified as at-risk were included for this analysis. A separate analysis was used to determine the strength and direction of the predictive relationship of the 2012 AIMSweb winter M-CAP target scores to the 2012 CRCT math scores. Students from the fifth grade general education population were included for this analysis as well. As school districts across the nation continue to rely on these measures to identify students at-risk and to provide rationale for implementation of instructional interventions, there is a great need to expand the research into the ability of math CBM to identify risk status and to accurately predict performance outcomes on state assessments.

**Assumptions and Limitations**

**Assumptions**

Adherence to all test protocols was assumed and test administration was done with fidelity.

Administration procedures were consistent across teachers.

Student scores from the 2010-11 math CRCT, the fall 2011-12 M-CAP, and the 2011-12 math subtest of the CRCT were used in the causal comparative study.

Student scores from the 2011-12 fall and winter AIMSweb M-CAP curriculum based measures and the 2011-12 math subtest of the CRCT were used in the correlation study.

The Learning Focused Schools model has been adopted by the school system and participant teachers utilized the strategies consistently.

Student participation in iPASS was consistent with curriculum implementation protocols.

**Limitations**

External validity: This study used data from fifth grade students from schools at a mid-size metro Atlanta school district. This may affect population validity and generalization to other larger districts.

The sample included a large proportion of students eligible for free and reduced lunch which may affect population validity and limit generalization to other various student characteristics.

Internal validity: Instrumentation threat may have resulted from inconsistency between the two measures despite curriculum based measures purporting to align with state assessments. Participants may have demonstrated testing effects due to similarity of the curriculum based measures across the two administrations.

History may have affected participants due to instructional differences, despite adherence to the LFS model, among teachers and their unique presentation of classroom instruction to the participating groups.

Differential selection may be a threat to validity due to the composition of each group since the groups were already formed.

# CHAPTER TWO: REVIEW OF THE LITERATURE

## Introduction

This chapter contains a review of the seminal literature on the subject of mathematic computer based instruction (CBI). Information concerning the history and research into computer based instruction along with specific components of the iPASS computer based curriculum are reviewed within the context of various ways to differentiate math instruction. A brief review of the Learning Focused Schools model completed the information relevant to the first research question.

To address the second and third questions, a review of the literature on math curriculum based measurement (M-CBM) and its ability to identify risk status and to predict performance outcomes on end of year state assessments is provided. Discussion of quantitative measures implemented in these studies provides information concerning statistical significance of the various studies. Specific information concerning the math portion of AIMSweb is presented within the context of the review.

Information from this literature review focuses the study and attempts (a) to determine the effectiveness of computer based math interventions (iPASS), as measured by student scores on the 2012 math portion of the CRCT, (b) to determine the ability of AIMSweb M-CAP fall cut scores to identify students as at-risk, and (c) to determine the ability of AIMSweb M-CAP winter target scores to predict performance outcomes on the 2012 math portion of the CRCT.

**Theoretical Framework**

Cognitive learning theories on transfer of learning and the use of previous schema in novel situations framed this research study. Shuell (1986) maintained that learning is cumulative and that a significant portion of importance is placed on prior knowledge and how it interacts with the pursuit of and generalization to new knowledge.

According to Anderson (1990) the use of schema or prior knowledge to connect new information was first introduced by Piaget in 1926. Anderson has extended Piaget's ideas on schema and developed a theory based on the use of prior knowledge to conceptualize information within new contexts. His research posits that concepts which are abstract can be made comprehensible when coupled with a foundation of concrete information, usually in the form of prior knowledge.

In the context of this research study, this relationship between fact knowledge and math concepts and problem solving was demonstrated by students in their learning processes as they applied prior knowledge to novel tasks when problem solving. Specifically, students in this study have demonstrated their set of skills on the state assessment and curriculum based measures. These skills and abilities have evolved over the repeated measures during the school year as new knowledge was added to the prior knowledge. As a result, when students were presented with new problem sets they were able to retrieve prior knowledge in the form of facts, concepts or algorithms that were relevant and necessary to complete the new problems.

Experiments conducted by Butler (2010) investigated whether repeated testing was more effective than repeated study as a factor in the transfer of learning. For each experiment the type of learning (repeated study, identical test, or different test) was the

independent variable. The dependent variable in each condition was test performance. Butler was investigating how initial learning conditions could be arranged to promote effective transfer to an expanded base of novel contexts. In the first condition he compared repeated testing to repeated studying and their effects on transfer. Using 3x2 repeated measures ANOVA, results were significant with an effect size of 0.24 for repeated testing which suggested that learning increased from test one to test two over repeated study. In the second condition he used inferential questions, different from test one to test two, and the effect size was also significant at 0.55. The first and second experiments indicated that repeated testing produced better transfer than repeated studying. The third test was a cued recall based on question type, and performance was similar in either same test or variable test; however both conditions, regardless of question type, produced a greater proportion of correct answers than the restudy condition. Overall results indicated that repeated testing proved to be more effective in transfer of knowledge than repeated studying. In his discussion Butler suggested that when facts or information were retrieved more than once, each subsequent retrieval increased the amount of information retained. It appears that the prior knowledge derived from the repeated testing was useful to the students when encountering new questions from various domains.

Research by Jitendra, et al. (1998) examined schema based math strategies versus traditional strategies on the acquisition, maintenance and extension of problem solving skills. Two phases were implemented: phase one assessed the effects of a schema strategy and a traditional strategy on simple one-step word problems. The second phase consisted of maintenance and generalization effects of the two strategies. Thirty four

elementary students, with mild disabilities or identified as at-risk to fail, were taught utilizing the two treatments (schema or traditional). Students in each group had adequate computation skills but were not able to utilize these skills effectively while problem solving. Results indicated that both groups were able to maintain problem solving skills and the ability to transfer the skills to new problems. Of significance were the differences in effect sizes between the two groups. On the immediate posttest (effect size 0.65), delayed posttest (effect size 0.88), and generalization (effect size 0.74) the schema based group outperformed the traditional group. The results suggested that schema based strategies had a significant impact over traditional strategies on the acquisition, maintenance and transfer of mathematic skills to novel word problems.

The current research study, utilized two groups of general education students who were matched on risk status and participated in classroom math instruction provided by their teachers. One group also received an additional instructional session of computer based math instruction using a software program that is designed around schema based strategies for problem solving. This study examined the magnitude of difference in mathematical problem solving skills between the two groups as indicated by their scores on the end of year state assessment.

Bevevino, Dengal and Adams (1999) maintained that students lacked the ability to apply and transfer learning and problem solving skills from one situation to another and that this condition remained a persistent problem in classroom instruction. Despite teachers' best intentions, some students were not always provided with the appropriate strategies or repeated practice and testing they needed to solidify information and make it accessible for retrieval at a later date. This study examined the effectiveness of classroom

math instruction with an additional overlay of computer based math interventions to determine if a significant difference was reflected in student math achievement over those students receiving only classroom math instruction.

Haskell's (2001) theoretical framework for general transfer involves eleven learning and instructional principles. Of relevance to this research study is Haskell's first principle of transfer which posited that primary knowledge is essential. While a knowledge base is crucial for transfer, it is imperative that it is prepared to accommodate new information to add to the prior knowledge (Haskell, 2001).

The AIMSweb M-CAP scores identified the knowledge base of the sample as either having average math skills or skills which were deficient and at-risk. Based on scores from the fall and winter administrations of the M-CAP, this research study attempted to determine the ability of the fall M-CAP cut scores to identify students' prior knowledge and to determine the predictive ability of the winter AIMSweb M-CAP scores on the end of year math portion of the state assessment.

Haskell's sixth principle argued that information retrieval was influenced by method used to encode it. Learners needed to structure a problem and develop a big picture so they might recognize relevant information in situations which were unfamiliar. Prior knowledge was replete with meaning and this meaning transforms information into transferable knowledge (Haskell, 2001).

This study attempted to determine the effectiveness of the method of encoding, based on the use of strategies and algorithms by students participating in the iPASS computer based math program, which was demonstrated by skill transfer to the end of year assessment, the CRCT.

Research by Tournaki (2003) involved teaching 42 second grade students' basic single digit addition. She attempted to determine the effects of teaching addition through drill and practice versus strategy instruction. Tournaki theorized that when math facts were first introduced they were perceived as problems to solve and the teaching of math strategies could lead to better learning and automaticity of fact retrieval and transfer. Two experimental conditions consisted of addend strategy and drill and practice with a control condition. Effectiveness was measured through posttest scores and students' ability to transfer the skill to a later task. The students showed improvement over the control in both strategy and drill and practice conditions. However, of greater significance was the moderate to large effect size (0.49) in the strategy instruction condition over drill and practice. Results from the transfer task indicated that the students' ability to transfer skills to a later task were more accurate only in the strategy condition. These results suggest that strategy instruction is superior to drill and practice and can significantly affect transfer of skills.

In order for instruction to be effective, students must be able to use the knowledge and skills they have acquired over the school year and demonstrate transfer on end of year state assessments. To sum up in a final statement by Perkins and Salomon, "…the ends of education are not achieved unless transfer occurs." (1992, p.3).

**History of Computer Based Instruction**

Research on the use of computer based math instruction in classrooms does not have as extensive a history as research into the effectiveness of computer based instruction in reading. Most rigorous research can be dated back to the 1970's, with most of the math software and technology used for drill and practice (Pollard & Pollard, 2005).

Much of the early research was inconsistent due to a lack of standardized definitions and design flaws. A significant number of studies had mixed results but recently, perhaps due to more stringently designed software, many of the studies indicated that computerized instruction does provide positive outcomes especially in the areas of time on task, cost effectiveness, and increased academic performance (Norris, Smolka, & Soloway, 2000). With the twenty first century's extensive immersion in technology, it is safe to say that computers and instructional software will continue to be integral parts of the instructional landscape.

**Research on Effectiveness of Computer Based Instruction**

Research into the efficacy of computer based instruction has recently had resurgence. Kingston (2009) conducted a meta-analysis of 81 studies from 1997-2007 which compared the effectiveness of computer based instruction in math versus the use of paper and pencil. His results indicated that estimated effect sizes from both modes of instruction were very small. However the majority of students, across all levels, indicated a preference for taking tests on a computer rather than paper and pencil.

Kroesbergen and Van Luit (2003) conducted a meta-analysis of 58 studies from 1985-2000 on the effectiveness of various elementary math interventions and the impact on students who had difficulties learning math. While most of the studies conducted were in basic skills instruction and indicated moderate effects, some studies were conducted on preparatory math and problem solving strategies. Overall results indicated moderate to strong effects with direct instruction (effect size 0.91) indicating a more robust impact than mediated instruction (0.34), or computer assisted instruction (0.51).

Fletcher-Finn and Gravatt (1995) provided a meta- analysis of the effect size of learning outcomes for computer based instruction. Based on their review there was a moderate learning increase in the use of computer based instruction across all subjects in the study. For the years 1987-1992 the effect size was .24 and for more recent years the effect size increased to .33. They suggested that this may have been due to more sophisticated and controlled studies with more comprehensive software.

Kulik and Kulik (1987) reviewed several meta-analyses on education outcomes for computer based instruction for students across all grade levels. A total of 199 studies were reviewed with 32 of the studies conducted on the elementary level. The analysis covered use of computers in (a) computer managed instruction, (b) computer-assisted instruction, and (c) computer-enriched instruction. The studies focusing on the elementary level indicated positive effects of computer based instruction on the achievement of students at this grade level. The authors did find differences in the effect size for those programs offered off-line, such as the four studies of computer managed learning, with an effect size of .07. The average effect size for 28 computer assisted or interactive instruction studies was .47, an increase from the $50^{th}$ to the $68^{th}$ percentile. Overall, some of the benefits listed included a decrease in instructional time of 32% and an increase in positive student attitude by 28%. Clearly these results indicated that computer assisted instruction, which involved the students in interactive lessons, provided robust achievement outcomes for students on the elementary level.

A meta-analysis by Bloc, Oostdam, Otter, and Overmaat (2002), reviewed 42 studies dealing with computer support for beginning reading. These studies comprised 75 experimental conditions and were dated from 1990. Their findings indicated the

overall effect size of computer based support for reading to be .19. They did caution that poor quality of many of the studies may have contributed to the relatively small effect size of the magnitude of difference between the means. However, they concluded that computer assisted instruction for beginning reading deficits was relatively effective and this conclusion corresponded with previous research indicating a positive, albeit, small overall effect.

Another meta-analysis was conducted by Liao (1999) which examined 46 studies comparing the instructional effectiveness of hypermedia, which is student directed computer assisted instruction that allows for access to various media, to nonhypermedia; such as computer based instruction, video, text, and traditional instruction in the elementary grades. The overall grand mean of the weighted effect size for the studies was 0.41. Based on his synthesis, this suggested that hypermedia was more effective with consistent positive results over non-hypermedia. These results added to evidence that the use of technology in classroom instruction had positive and significant effects.

Research conducted by Fuchs, et al. (2006) investigated the effects of computer based instruction on promoting addition and subtraction fact knowledge, as well as transfer of this knowledge to subsequent assessments. Simple fact problems, sums to nine embedded in word problems, were presented to first grade students identified as at risk to fail. No significant effect sizes for subtraction or story problems were identified. Yet for subtraction and word problems, utilizing pictorial number combinations that were developed to pair problem stems with information in short term memory in an effort to encode for long term memory, did result in a statistically significant improvement on subtraction number combinations. The authors suggested that the use of pictorial

representations may aid learning and that more research should be geared toward measuring the effects.

Chang, Sung, and Lin (2004) conducted a study on the use of computer based software, MathCal, designed to diagnose stages in which students' ability to problem solve breaks down. The authors suggested that previous instructional math software systems did not isolate problem solving steps into stages, and thereby made it difficult to diagnose where to intervene. The authors identified four stages, (a) comprehend the problem, (b) devise a plan, (c) engage the plan, and (d) check the solution. By identifying these particular stages, they were able to isolate the stage in which students encountered problems. Using an experimental control group, pre-post test design, the results indicated that MathCal was effective in increasing the problem solving skills of low performing students. For the experimental group, these results indicated that math software designed to recognize student difficulties and to provide interventions and training to students resulted in a statistically significant increase in their problem solving skills at each stage.

Research by Camli and Bintas (2009) investigated the use of computer instruction on solving multiplicative structures. The researchers provided an introductory phase in which students were able to see the solution of the problem step by step, with an explanation of the answer at each step, while interacting with the software. Controlling for entering achievement levels, results of t-test analyses indicated that the difference in performance of the experimental group was statistically significant and greater than the performance of the control group. This result appeared to suggest that students were better able to conceptualize lowest common multiples and greatest common factors when

computation of the problem was broken into discrete steps with accompanying explanations.

Wijekumar, Hitchock, and Turner (2009) investigated the effects of using the Compass Learning Odyssey Math software on the math achievement of fourth grade students in an effort to determine if the widely used instructional program was effective in raising students' math achievement. The study was the first randomized controlled trial conducted to determine a causal relationship with this software. The study did not produce a statistically significant impact on end of year math achievement and only produced a .20 effect size. When comparing the end of year achievement levels for the experimental and control groups, no statistically significant difference was found between the two groups.

**Research Design for iPASS**

A great body of literature exists which identifies instructional practices that have been found to be effective in teaching mathematics. These factors include scaffolding of instruction, use of manipulatives and visual representations, strategy instruction, use of instructional software, structuring of content, and control and sequencing of examples, to list a few (Maccini & Gagnon, 2000; Swanson, 2001; National Research Council, 2001; Baker, Gersten, & Lee, 2002; Kroesbergen & Van Luit, 2003; [cited in iLearn, 2010]).

According to the website (iLearn 1989), Individual Prescription for Achieving State Standards, iPASS is a web-based math curriculum and instruction system that individualizes instruction based on identified student needs. It further states that the research base for the curriculum comes from research in four areas: (a) effective instructional methods, such as schema based strategy instruction (b) successful

methodology for students with math learning difficulties, (c) methods for students with learning difficulties in all areas, and (d) effective methods for instruction using multi-media. The lessons are prescriptive indicating that instruction is matched to specified skill and concept deficits. It is mastery-based allowing students to develop mastery of a skill before moving on to other skills. (iLearn, 1989).

Until recently, most math software was developed to address fact deficits and to increase fact fluency (Lynch, 2006). Today more software is being designed to help students develop strategies which enable them to problem solve quickly and accurately. While students still need to automatically retrieve math facts, they also need to understand how to conceptualize a mathematic problem statement.

The iPASS system purports to deliver instruction which is explicit in the teaching of cognitive strategies for solving math problems. In addition, the strategies taught provide opportunities for generalization as well as being specific enough to cover a range of problems which the student may encounter. An important characteristic of iPASS is its ability to actually provide systematic instruction designed to address the needs of students who exhibit difficulty learning math. The curriculum is designed to teach students the schema based strategies that can assist with retrieval of facts necessary for accurate problem solving in novel situations (iLearn, 2010).

Research reviewed by Geary, Hoard, Nugent and Byrd-Craven (2009), indicated that when children with math difficulties were asked to solve addition problems with fact retrieval only and not the use of counting strategies, they made more errors involving counting strings with addends. These results suggested that when computations are

strengthened through the use of algorithms and strategies, students can be significantly aided in their ability to retrieve facts.

Additional research by Hanich, Jordan, Kaplan, and Dick (2001) indicated that reading difficulties which were co-morbid with math deficits appeared to hamper students' use of strategies to problem solve. They examined the performance of 210 second grade students with, (a) reading and math difficulties, (b) math only difficulties, (c) reading difficulties only, and (c) children not demonstrating reading or math difficulty. Those students with only math difficulties performed better in areas mediated by language than students with reading and math difficulties. However, the math only difficulty group did not out perform the math and reading difficulty group in areas which relied on the use of visuospatial processing and automaticity. These results suggested that the use of visual representations were helpful to students when problem solving.

The design for iPASS lessons uses explicit cognitive strategies in conjunction with scaffolding support from visual representations (iLearn, 1989). In this way, it appears that students are provided additional visual information to help them make sense of mathematical abstractions and provide more support with areas that rely heavily on reading.

Gerston, et al. (2009) posited that the use of heuristic strategies, which were not problem-specific, were very effective instructional techniques because they provided students with an organizational framework or schema. Citing information from a meta-analysis, they stated the mean effect size from four studies using heuristic strategies was significant at 0.56. They further indicated that problem solving interventions which

showed great success were those that used heuristic strategies and were then followed by student discussion of how they came about their solutions.

VanDerHeyden and Burns (2009) indicated that mathematical competence can be construed as the ability to correctly predict, quantify and verify relationships. These are much more complex skills than simple computations and often entail the use of strategies generalized to new contexts. They cite Haughton (1980) as specifying objectives for math instruction that ensured levels of performance which resulted in remembering information over time, endurance and retention of the information, and application of the information to novel situations.

The instruction provided by iPASS for problem solving is schema based. This enables the students to use prior knowledge while being taught the heuristic which will enable them to organize problems based on the type of strategies necessary to solve them (iLearn, 2010).

The design for iPASS methodology is based on explicit instruction with a heavy emphasis on cognitive strategies that allow students to comprehend how math content is related and how to apply strategies when solving a variety of problems. It teaches strategies which are general enough to reduce memory demands but explicit enough to be generalized to a variety of problems. When teaching the use of strategies, it provides models and visual representations outlining each step of the strategy and how it should be used to reach the solution. Students then immediately have opportunities to use those strategies across several sample problems. Based on this strategy instruction, they are taught how to analyze a problem, apply a systematic approach to mapping the problem sentence and then use it to solve the problem (iLearn, 2010).

It appears that many characteristics of the iPASS research design and software are based on sound research studies which have garnered positive results when measured against student acquisition, maintenance, and generalization of academic skills from practice to novel problem solving.

**Learning Focused Schools Model**

The Learning Focused Schools (LFS) model was founded in 1993 by Scott Thompson. The Learning Focused Schools model is a framework that encompasses the US Department of Education's evaluations of strategies which are research based and most effective in raising student achievement (Learning Focused Schools, 2007).
It is based, in part, on the "90/90/90" schools research conducted by Douglas Reeves at the Center for Performance Assessment (Thompson, 2006). Dr. Reeves introduced the "90/90/90" concept to signify schools which had 90% or greater of the student body eligible for free or reduced lunch; 90% or more of the students' members of ethnic minority groups; and 90% or greater meeting achievement standards on at least one state or local academic measure (Reeves, 2003). The Learning Focused Schools model provides research based strategies and practices along with a procedural framework designed to help teachers prepare students' to access state standards.

A body of related research exists on the use of procedures and strategy instruction and its impact on student achievement (Montague, 1997; Pressley, Goodchild, Fleet, Zajchowski, & Evans, 1989; and Goldman, 1989). Research by Marzano, Zeno, and Pollack (1999) determined the percentile gains and effect sizes for the five most effective research based strategies utilized by LFS. According to their research, use of extended thinking strategies had an effect size of 1.61 with a 45 percentile gain. For summarizing,

the effect size was 1.0 with a gain of 34. Use of vocabulary in context yielded 0.85 and a 33 percentile gain. Using advance organizers and non-verbal representations had effect sizes of 0.73 and 0.65 respectively with percentile gains of 28 and 25. These results indicated that the strategies were effective in raising students' achievement. The use of the LFS model encompasses these strategies and ensures that all teachers are consistent in the delivery of instruction which is aligned to state standards, incorporates research based strategies, and utilizes the resources within the school environment.

In the area of mathematics instruction, one of the strategies the LFS model recommends is the use of ROPES as an organizing structure for students as they formulate the steps necessary for mathematic problem solving. Each letter describes an action necessary for comprehending and computing the problems. Learning Focused Schools model strategies in math are based on ROPES. Use of this strategy enables students to engage in problem solving in an organized and effective manner (Cargill, 2009). The letters in ROPES indicate steps necessary for mathematical problem solving:

R Read the problem and underline details

O Omit information which is not necessary

P Plan out the steps

E Efficiently check the steps as you complete the problem

S Study the answer and recheck work (Cargill, 2009).

The school district for this research study has adopted the LFS model and all classroom teachers were trained in model guidelines and implementation procedures.

**History of Curriculum Based Measures**

The provision of differentiated instruction and remediation should be predicated on accurate information concerning students' present expanse of knowledge and skill levels. A pivotal piece of the No Child Left Behind Legislation (NCLB, 2001) and the ESEA Flexibility (2011) is the use of diagnostic tools to identify students' achievement levels prior to and during instruction, monitor progress toward meeting state standards and measure the effectiveness of any prescribed interventions. This response to intervention is most frequently assessed through curriculum based measures (CBMs).

CBM was developed over forty years ago at the University of Minnesota's Institute for Research on Learning Disability (Deno, 2003). The measures were standardized and most frequently used for evaluating the effects of academic interventions across the curriculum (Shinn, 2008).

According to Deno (1992) curriculum based measurement had been used since the late sixties, but it wasn't until the 1970's that the standardization, reliability and validity were deemed technically adequate. During this process a 3-step approach was developed: (a) identifying indicators of basic skills which were amenable to measurement, (b) isolating a variety of measurement parameters, and (c) conducting validity studies to determine which indicators were sufficiently high in criterion validity to be included in the measurement system.

Elliot, Huai, and Roach (2007) have indicated that early and accurate identification of students demonstrating academic deficits was a valuable outcome of screening measures. They cited a number of researchers that had examined the relationship between students' classroom performance and statewide indicators of

performance. Without intervention, students at-risk who did not demonstrate adequate grade level skills continued to struggle and made progress at a much slower rate.

Albers, Glover, and Kratochwill (2007) noted that identified risk can exist on a continuum, and it's critical to track its trajectory through direct skill measurement. This can most effectively be done with progress monitoring tools that are effective and accurate in identifying risk status and can demonstrate predictive validity on state tests. They suggested more research into CBM was necessary to improve predictive ability due to their overly sensitive nature in risk identification.

Gersten et al. (2012) reviewed contemporary research on the use of CBM for diagnostic classification and predictive accuracy. The authors suggested that "when designing screening measures for mathematics, a critical variable to consider is the extent to which performance on those measures relates to later performance in mathematics" (p. 436).

Silberglitt and Hintz (2005) recognized that scarce educational resources would be wasted when CBM incorrectly identified risk levels and that the strength of these measures was based on how accurately they identified the students who needed help and their ability to classify risk levels.

Curriculum based measurement (CBM) has a long history of research, especially in the area of reading. Despite more research time spent on reading measures, Geary (2004) reported that 5% to 8% of children in school were disabled in mathematics. According to Cusumano (2007) during the past 25 years CBMs have been scientifically validated as effective tools to be used as universal screeners, as tools to monitor student progress, to predict student achievement outcomes and ultimately to improve instruction.

He further indicated that students identified as at-risk who were not making progress on grade level standards in reading or math should be identified early and accurately. Students can then be provided with effective interventions and their progress monitored. Using this approach, with respect to curriculum based measurement, the ability to impact achievement outcomes would be greatly enhanced.

Reliability data on CBMs consisted of examinations of internal consistency, test-retest and alternate forms reliability with a traditional benchmark of .80 for progress monitoring and .90 for more stringent applications (Clarke & Shinn, 2004). Most validity research focused on concurrent and predictive validity with mathematics measures yielding coefficients in the .50 to .70 range (Salvia, Yesseldyke, & Bolt, 2007).

Research has indicated that summative measures were not adequate for monitoring students' response to intervention due to time constraints, lack of sensitivity to skill changes, and general lack of applicability for repeated administration (Salvia, Yesseldyke, & Bolt, 2007). As a result, schools began to either develop their own common assessments or used standardized, published curriculum based measures for these purposes.

CBM is a formative, criterion assessment tool with specific characteristics designed to align with the basic skills that students should demonstrate at each grade level. For this reason, many school districts were using them to help identify students at-risk in meeting state standards, to provide specific data for the design of academic interventions, and as predictive measures. Since state curricula were often designed after universal standards and teachers were responsible for providing the precise instruction necessary for students to demonstrate acquisition of grade level standards on state

designed summative assessments, it continues to be imperative that risk status be accurately identified and that progress is monitored (Hosp, Hosp, and Howell, 2007).

Deno (2003) noted that there are nine attributes necessary for adequate development of CBM:

- CBM has alignment with basic skills taught at each grade level which ensures that content is the same, and stimulus materials are consistent with similar responses.
- The measures are designed with technical adequacy, indicating they have established reliability and validity. Thereby ensuring that CBM is not considered to be an informal measure of basic skills, but rather a proven measure of student progress.
- The measures are criterion referenced and require students to demonstrate knowledge of specific skills through performance tasks.
- The standardization of procedures and administration ensure that data can be shared and interpreted by various users. At the least, standard tasks are used for each content area, standard procedures are followed, and standard administration and scoring are ensured.
- Through the use of low-inference measures, correct and incorrect student demonstrations of the skill provide samples of the student's performance.
- The inclusion of decision rules provide interpretive data based performance criteria which have been standardized through sampling procedures.

- CBM emphasized repeated measures over time to track progress which will generate rate of progress as well as level of learning through the use of progress monitoring.
- CBM is efficient to administer because it provides data which is easily converted to performance data rather than percentiles or normal curve equivalents.
- Depending on need, data can be summarized quickly and efficiently using web based management systems.

**Research on Effectiveness of CBM**

Initially, studies on CBM began with early reading skills and while research into the effectiveness of CBM has been extensive, it has not always been in the context of measuring response to intervention or predicting outcomes on summative state assessments. Given the national mandates and requirements to screen for students not making adequate progress, continued investigation into their effectiveness as tools that accurately identify risk status and correlate well with summative assessments should continue to evolve (Cusumano, 2007).

Wallace, Espin, McMaster, Deno, and Foegen (2007) indicated that research varied as to scope. They observed that more research had been conducted for efficacy of reading probes than for mathematics, and that more recent studies had begun to focus more on mathematics as an area in which research needed to be expanded. In addition, the authors noted that while most CBM use has been for progress monitoring, it has recently been proposed for use as a means to predict performance on state defined academic assessments.

Research conducted by Crawford, Tindal, and Stieber (2001), provided data from a two-year period on how well second and third grade students' ability to read from text, modified to resemble a reading probe, could predict performance on state assessments. Results indicated that of the sample of 51 students 65% passed the reading assessment and that a correlation of .84 was calculated between second and third grade reading rates on state assessments. Based on these results, the researchers concluded that there was support in the use of timed readings to predict performance on state assessments.

Additional research by Hintz and Silberglitt (2005) utilized longitudinal analysis of temporal benchmark assessments of reading for a cohort of first through third grade students. Predictive validity was studied against end of year state assessments for each of the grades. Their analysis of the accuracy of these measures was based on probability analysis of (a) specificity, (b) sensitivity, (c) positive and negative predictive power. ROC curve analysis suggested that the R-CBM cut scores were effective on all three measures and that the reading CBM had a significant relationship with the end of year assessments ($r$ = .49 to $r$ = .94) for each of the grades.

Clark and Shinn (2004) argued that criterion measures such as CBM should be able to identify students at-risk to fail and then should also be useful when monitoring their progress. Results from their research into the reliability, validity and sensitivity of math CBM indicated that the measures met reliability criteria for identification of risk status. In addition, the sensitivity of CBM was measured with a repeated measures ANOVA yielding significant F statistics which indicated ability to progress monitor student growth with each measure.

Descriptive research conducted by Tindal, Germann, and Deno (1983) indicated that early math probes were designed to measure early numeracy, single-skill facts, computations, and fact drills. By the early 1990s math application probes were developed and, more recently, other math skills were being assessed to include concepts, various mathematical applications, and problem solving (Hosp, Hosp, & Howell, 2007).

Research by Helwig et al. (2002) investigated the predictive ability of math CBM on computer based tests aligned to state assessments. The CBM was designed to measure higher order conceptualization and eliminated those areas which required procedural and fact knowledge. Eighth grade students participated in the study and their findings indicated a strong correlation ($r$ = .83) between the measures. A discriminant function analysis indicated that CBM was effective in predicting the number of students able to score at a level on the computer test to meet the end of year math standard. The authors concluded that the students who were more accurate on the CBM were able to demonstrate higher order conceptualization and stronger mathematical schemas which would translate into higher scores on math assessments.

Hintz, Christ, and Keller (2002) examined the generalizability of mathematical probes. Their sample consisted of 67 students in first through fifth grades. All students were assessed with single and multiple skill mathematical survey probes. The single skill and multi-skill probes were found to measure distinct constructs. The single-skill probes indicated upwards of 80% of the measurement constructs attributable to variance by student and age which indicated sensitivity to differences in performance. In the case of multi-skill probes, 75% of the variation was explained by student and developmental levels. These results indicated that each type of probe elicited a different type of

behavior within the construct of math achievement. Multi-skill probes assessed general outcomes over a long term, sampling various skills and single skill probes targeted specific subskills. Information concerning both types of outcomes is extremely important for accurate progress monitoring and to ensure utility as a predictor of future achievement on state assessments.

Research by Jitendra, et.al (2005) examined the reliability and predictive validity of CBM which utilized word problems on final state assessments. Internal consistency reliability coefficients were lower than previous studies but their reliability increased when two probes were utilized. They concluded that more than one administration of a measure might be required to obtain an estimate of performance that was reliable. Predictive validity analyses indicated that two administrations of word problem CBM , winter and spring, indicated moderate relationships with final state assessments at $r = .71$ and $r = .54$ respectively.

Shapiro, et al. (2006) noted that questions were raised about the relationship of CBM to state assessments, especially in the area of math since very few studies had investigated this relationship. They investigated the relationship between the fall, winter and spring CBM scores for fifth grade students and end of year performance on state assessments. The computation and concepts and applications probes were examined with the winter concepts and application probes the more accurate predictor ($r = .64$ vs. $r = .53$). Through ROC analysis, cut scores were identified that would classify students' risk status. Overall correct rates were impressive with classification at .65, specificity at .66 and sensitivity .65.

Deno (2003) indicated that CBM can be used to screen for risk status. Since they are usually standardized they are frequently used to compare individuals to performance of a group. He cited a study by Deno, Reschly-Anderson, Lembke, Zorka, and Callender (2002) in which CBM in reading was administered in first to fifth grade at an urban elementary school. Scores were combined within and across grades and based on ROC analysis, students in the lowest 20% of the reading CBM were identified as high risk to fail and in need of progress monitoring. By identifying the proportion of students at-risk individually and within a group, teachers were able to modify instruction and monitor the effects of the intervention. This underscored the value of CBM and its ability to provide valuable information for planning and for using research based interventions to increase student achievement.

There have been studies that have found positive relationships between curriculum based measures and an increase in student achievement (Glover and Albers, 2007; Hosp, Hosp and Howell, 2007). Studies that examined the essential components of early identification and programming for intervention have also shown positive results (Burns, M. K. 2004; Deno, S. 1992; Fuchs, L., 2003).

Helwig, et al. (2002) theorized that measures that could provide a timely estimate of student skills and risk status, coupled with measures that simultaneously monitored a student's progress toward statewide goals would be powerful educational tools. This view makes a compelling argument for early identification of risk status because if predictive validity is demonstrated, interventions can be provided early on to those students identified as at-risk to fail state assessments.

**Research Design for AIMSweb M-CAP**

According to the website, Achievement Improvement Monitoring System, AIMSweb, is a benchmark and progress monitoring system based on direct, frequent and continuous student assessment (AIMSweb, 2009b). AIMSweb has been developed to be fairly independent of standards and curricula which allow it to measure skills regardless of the curricula used by school districts. The Mathematics Concepts and Applications (M-CAP) is the most recent addition to the system.

Within AIMSweb several assessments are designed to assess progress in math skill development. Mathematics Computation (M-COMP) measures basic computation skills. Mathematics Concepts & Applications (M-CAP) is designed to assess the application of concepts in mathematics. Both are designed for universal screening as well as monitoring progress. Both measures provide single and multi-skill probes (AIMSweb, 2009b). The focus of this research study is on the M-CAP.

Generally, technical data for mathematics curriculum based measures are lacking; however, several studies have been conducted and results provide support for their use. Research from Thurber, Shinn, and Smolkowski (2002) yielded an alternate forms reliability coefficient of .91 for computations, with a lesser correlation of .42 for applications on M-CBMs. A study by Foegen and Deno (2001) resulted in internal consistency reliability coefficients ranging from .77 to .93; test-retest correlations from .76 to .88; and parallel forms ranged from .67 to .86

For the development of AIMSweb math probes, extensive psychometric measures were utilized to develop equivalent probes. Pearson's correlation was used to assess consistency with grades. For grade five, the focus of this research study, equivalent

probes yielded a correlation of .84. To measure internal consistency, Cronbach's alpha yielded .84 and split half reliability correlated at .89. Selection of probes was statistically significant at the .05 level. Receiver operating characteristics (ROC) curve analysis was used to develop cut and target scores (Harcourt, 2006; AIMSweb, 2009b).

The National Center on Student Progress Monitoring evaluates curriculum based measures according to criteria set by the Standards for Educational and Psychological Testing. Tools are measured against seven standards, (a) adequate number of forms, (b) improvement rates specified, (c) specific benchmarks, (d) evidence of increased student achievement, (e) sensitivity to improvement, (f) reliability, and (g) validity (NCSPM, 2007). Two AIMSweb reading curriculum based measures met the seven standards. At this time, Edformation is in the process of submitting tools for mathematics computation for review while work continues to ready the math Concepts and Applications probes for review (Edformation, 2002).

Information provided by Fuchs, Fuchs, and Zumeta (2008) examined two approaches for CBM assessments. The first was to measure skills on tasks which were global measures of math competence. The second was a systematic approach that sampled the full math curriculum throughout the year, and thereby, ensured comparable emphasis on alternate forms. They demonstrated that progress monitoring of the relatively complex domains within mathematics concepts and applications was most accurately accomplished using a curriculum sampling approach.

AIMSweb M-CAP probes are designed along the curriculum sampling approach. In addition, each probe has had additional field testing to increase measures of equivalency among the probes. AIMSweb M-CAP is a quick assessment (8 to 10

minutes) that measures general problem-solving skills from the following fifth grade domains: algebra, number sense, operations, patterns and relationships, measurement, geometry, and data and probability (AIMSweb, 2009b).

**Summary**

This literature review is a brief investigation of the research conducted on the efficacy of computer based instruction (CBI) as an instructional method to increase student achievement. It has a long history of instructional use and recently software development has become more sophisticated. As a result, math software is not only used for drill and practice, but it is now used as a total instructional curriculum much like iPASS to address more complex math deficits such as problem solving through the use of strategy building (iLearn, 1989). For this reason, it is critical that information on the effectiveness of research based interventions such as iPASS is accurate and readily available. This will ensure that educators are able to provide student supports that are cost effective and, more importantly, will translate into increased student growth.

The literature has expanded the discussion into the rationale for the use of curriculum based assessment (CBM) as a tool to identify students as at-risk to fail state assessments and for their use to predict performance outcomes. Information from curriculum based measures is heavily relied upon to accurately identify skill levels and based on these measures academic interventions are provided.

The Douglas County School District has contracted with an outside vendor to provide a research based computer program to address mathematical deficits in students identified as at-risk to fail state assessments. As a preliminary measure, this program was implemented in a few select schools since total implementation would be a costly

initiative. If additional data were available to examine the achievement outcomes of students using this program compared to students not using it, critical information to help facilitate discussions for initiating total implementation would be available.

According to Eckert, Dunn, Codding, Begeny, and Kleinmann (2006) the accuracy of the assessment of students' academic abilities has been identified as one of the most crucial variables related to effective instructional planning and positive student outcomes. Therefore it is imperative that these measures identify risk status reliably, are accurate tools for prescribing interventions which address the needs of students identified as at-risk, and are valid predictors of student performance on state assessments.

# CHAPTER THREE: METHODOLOGY

## Introduction

The purpose for this research study was two-fold: To compare the performance outcomes, as measured on the 2012 math subtest of the Georgia Criterion Referenced Competency Test (CRCT), for fifth grade general education students identified as at-risk who participated in a computer based math curriculum versus students identified as at-risk who did not participate. Students were identified as at-risk based on their 2011 fourth grade math CRCT scores and their subsequent 2011-12 fall curriculum based M-CAP cut scores. If students did not meet the performance standard on the 2010-11 CRCT and did not meet the cut score on the M-CAP, they were identified as at-risk and some were provided computer based interventions.

Results of the causal comparative analysis will provide information to educators concerning the impact of additional computer based math instruction to students identified as at-risk to fail state assessments. This will assist education stakeholders to identify programs that are effective in remediating math deficits and are worth the investment in time and resources.

The second purpose for this research study was to investigate the ability of curriculum based measures to identify risk status and to predict performance outcomes on state assessments. Cut scores on the 2011-12 AIMSweb Mathematic Concepts and Applications (M-CAP) fall curriculum based measure were examined from fifth grade general education students to determine their ability to identify students at risk. In addition, target scores on the 2011-12 AIMSweb Mathematic Concepts and Applications (M-CAP) winter curriculum based measure were examined to determine their ability to

predict achievement outcomes on the math subtest of the 2012 CRCT. Tasks on these quick measures were designed to align to grade level curriculum standards and the designated scores on the measures are purported to identify those students who may be at-risk to fail and those that are on target to pass state assessments. Based on the cut scores, students were provided with an additional overlay of computer based math instruction. This study examined the ability of these measures to identify students' risk and to predict performance outcomes on the Georgia Criterion Referenced Competency Test.

The correlation study is relevant because it provides additional information concerning the effectiveness of curriculum based measures to identify students at risk. Furthermore by identifying these students and determining predictive validity, teachers can be proactive early on and implement effective math strategies and interventions that can have a positive impact on student outcomes.

**Research Design**

A causal comparative research design was used to compare achievement scores from two groups of fifth grade students in an attempt to establish a possible causal relationship between computer-based math instruction and meeting standards ( a scaled score of 800 or above) on the math subtest of the CRCT. One group participated in computer based math instruction and the other group did not. This research design was appropriate because the study was non-experimental, ex-post facto, and investigated any possible causal relationships between the independent variables and the dependent variable (Airasian, & Gay, 2002).

The correlation analyses were non-experimental, ex-post facto and investigated the ability of the fall and winter M-CAP scores to identify at-risk fifth grade students and to predict performance outcomes on the 2012 administration of the CRCT respectively. This type of research design was appropriate because archived test data was used to describe in quantitative methods the degree to which the independent variables were related to the dependent variable. (Airasian, & Gay, 2002).

**Questions and Hypotheses**

The research questions for the study were:

Research Question 1: Is there a difference in the math CRCT scores between at-risk fifth grade students who participate in computer based math instruction and those who do not?

Research Question 2: Are the fall curriculum based math M-CAP cut scores able to identify fifth grade students as at-risk to fail the CRCT?

Research Question 3: Are the winter M-CAP progress monitoring target scores able to predict fifth grade students' performance outcomes on the CRCT?

The following were the research hypotheses:

Research Hypothesis 1: Fifth grade students identified as at-risk who participate in computer based math instruction compared to at-risk fifth grade students who do not participate in computer based math instruction will have statistically significant different scores as measured by the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

Research Hypothesis-2: At-risk fifth grade students' fall M-CAP fall cut scores will have a statistically significant relationship to their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

Research Hypothesis-3: Fifth grade students' winter M-CAP progress monitoring target scores will have a statistically significant relationship to their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

The following were the null hypotheses:

Null Hypothesis 1-$H_{o1}$: There is no statistically significant difference in at-risk fifth grade students' achievement scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test between students who participate in computer based math instruction and those who do not.

Null Hypothesis 2-$Ho_2$: There is no statistically significant relationship between at-risk fifth grade students' fall M-CAP cut scores and their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

Null Hypothesis 3- $Ho_{3:}$ There is no statistically significant relationship between fifth grade students' winter M-CAP progress monitoring target scores and their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

**Participants**

Only scores from fifth grade students identified as general education were used in this study. Scores from special education students were not used in this study because modifications and accommodations may have been provided for the assessments. Attempting to control for these variables may have confounded study results. Student scores from each of the twenty elementary schools were used. Scores from the four

schools in the causal comparative study were excluded from the correlation analyses. For research question one student scores of less than 800 on the 2010-2011 fourth grade CRCT state assessment in math and cut scores of less than eight on the 2011-12 fall M-CAP were used. At-risk status was defined as scoring below average (8) on the curriculum based math measure. The sample was formed from the four elementary schools in proximity to the same middle school in a metro Atlanta school district. Two of the schools had computer labs and were implementing the iPASS computer based math program and two of the schools were designated for possible future program implementation. These schools were similar in number of students eligible for free lunch as well as ethnic, racial, and socio-economic background. This district has a total student population of 24, 250 students and is the 17$^{th}$ largest school district in Georgia. There were twenty elementary schools, eight middle schools and five high schools. A majority of the schools had students eligible for free or reduced lunch (NCES, 2012).

Two sets of scores from fifth grade students in the remaining 16 elementary schools were analyzed for research question two; student scores identified as at-risk (less than eight) from the fall administration of the 2011-12 M-CAP and scores from the 2011-12 CRCT math subtest.

Research question three analyzed two sets of scores from fifth grade students. Target scores of 10 or greater were identified as on target to meet the end of year state standard. The target scores from the 2011-12 winter administration of the M-CAP curriculum based measure and the 2011-12 CRCT math subtest were used. This comprised scores from all fifth grade general education students that participated in the 2011-12 fall and winter administrations of the curriculum based measures and the 2011-

12 CRCT math subtest. These scores were gathered from the remaining sixteen elementary schools. The sampling procedure was non-probability sampling since the schools were within proximity to one another, convenient, and the sample population had already been established (Gall, Gall, & Borg, 2007).

The number of participant scores for the first research question was 26 students per group for a total of 52 scores. A sample of 30 scores is considered sufficient and necessary to establish significance at the .05 level (Gall, Gall, & Borg) and this minimum was met for the causal comparative analysis. Criteria for study participation required that each student score was less than 800 on the 2010-2011 fourth grade CRCT math subtest, and that the student M-CAP cut score was classified as at-risk, less than 8. Scores meeting these criteria were further collapsed into a Group 1 set of scores, derived from students who received iPASS computer based math instruction in addition to classroom instruction based on Learning Focused Schools model strategies. Group 2 scores were comprised from students meeting the same criteria but had not participated in iPass computer based instruction.

Participation in the computer based math program consisted of an additional overlay of computer based math instruction for 50 minutes per day, five days per week, from September 2011 to March 2012. The iPASS curriculum was self-paced and students were assessed with a diagnostic placement test. The diagnostic placement test identified the skill set deficiencies and placed students in the appropriate instructional level. The levels were designed to provide strategy based instruction and remediate identified gaps in mathematical foundations. The focus was on mathematical concepts and applications so computations were embedded in word problems. Specific strategies

were provided to help students identify the steps necessary to correctly set up the problem and complete the computations. The strategy instruction was generated by the computer which was based on the student's performance and was thereby individualized according to student needs. iPASS is designed to provide explicit instruction to those students who have math difficulties and have often scored below proficient on state assessments. It provided a comprehensive and prescriptive instructional process which incorporated automatic corrective feedback for any errors. Since iPASS diagnoses, prescribes instruction based on student needs, sets criteria, and monitors student output, it allows teachers in the computer labs to supervise program operation and to support students while they complete their lessons (iLEARN, 2010). This instruction was in addition to the fifty minutes per day of math instruction in their general education classrooms by their teachers. Classroom instruction consisted of the Learning Focused Schools model. Each math lesson was started with an essential question which defined the objectives of the lesson. Lessons were fifty minutes long and students were taught using the ROPES strategies for math instruction.

Group 2 was comprised of scores from students meeting the at-risk criteria and who received 50 minutes of daily classroom instruction but did not have the additional overlay of computer based math instruction.

The sample of scores on the spring 2012 math CRCT subtest from Group 1 (computer participation) were compared with the sample of scores on the spring 2012 math CRCT from Group 2 (non participation) to determine statistical significance of mean score differences between the two achievement measures on the CRCT and to address the first hypothesis.

A set of 2011-12 fall M-CAP cut scores were collected from the fifth grade general education students at the 16 elementary schools who were identified as at-risk to determine the ability of the M-CAP to identify risk status. Another set of 2011-12 winter curriculum based M-CAP target scores were collected from the group of fifth grade general education students at the 16 schools and examined for their ability to predict performance outcomes on the 2012 CRCT. These investigations addressed the second and third hypotheses respectively.

The administration window for the CRCT was two weeks in early April. All data was available as part of the school district's total assessment information and was located in the participating school's databases. All student scores were kept secure and were coded for confidentiality.

**Setting**

The setting was a mid-size, public elementary school district west of Atlanta, Ga. This setting was chosen because many of the students in these schools have struggled in recent years making proficient scores on the math portion of the CRCT. Therefore, the need for tools that identify risk status and predict student achievement outcomes on state assessments are necessary and essential components of accurate diagnostic instruction. In addition information concerning the effectiveness of computer based math interventions was a critical consideration when prescribing interventions and supports that can translate into increased student growth.

**Instrumentation**

The Georgia Criterion Referenced Competency Test (CRCT) is the state mandated end of year assessment designed to measure student achievement of state

standards. The math subtest from the 2011-2012 administration was used as the dependent variable for each of the three research questions. Math ability was measured on three levels: exceeds the standard, meets the standard, and does not meet the standard. Students who did not meet the standard had a scaled score which was below 800 (Georgia Department of Education, 2012). The reliability and validity as a standards-based assessment has been established through the standard error of measurement and Cronbach's alpha (Georgia Department of Education, [Accountability Workbook], 2004).

Individual Prescription for Achieving State Standards (iPASS) is a computer based math curriculum. The program purports to increase students' math achievement through strategy instruction based on use of algorithms and heuristics applicable in novel situations. The research on which it was designed came from numerous studies which examined effective instructional practices (iLearn, 1989).

AIMSweb Mathematics Concepts and Applications (M-CAP) is a curriculum based measure, designed by Edformation, which assesses mathematics problem solving skills (AIMSweb, 2009a). The 2011-12 AIMSweb fall and winter M-CAP scores were the independent variables. Fifth grade students were identified as at-risk if cut scores on the fall administration of the M-CAP were less than 8. Fifth grade student target scores that were 10 or greater on the winter administration were identified as on target to pass the state assessment. At this writing, Edformation was in the process of supplying curriculum based measurement tools in mathematics computation and early numeracy to the National Center on Student Progress Monitoring for technical review (Edformation, 2002).

**Procedures**

Archived test data was used; therefore, an application for expedited status was submitted to the Institutional Review Board (IRB) and approval to conduct the research has been granted. Permission to conduct the study has been granted by the Associate Superintendent of Student Supports and Services for the Douglas County School District.

The principals at the twenty elementary schools were contacted. After explanation of the purpose of the research study and approval, this researcher began to collect the data for the research data set.

All student test data was coded to ensure anonymity. Identification of students participating in iPASS and those not participating was coded for confidentiality and kept secure. No identifying information was a part of the research data set.

Data for both the M-CAP and CRCT was located at the individual schools and kept secure.

Research question 1 compared Group 1 and Group 2 students' mean score differences on the 2011-2012 math CRCT.

Research question 2 correlated cut scores from the 2011-12 fall M-CAP administration to the scores on the math portion of the 2011-2012 CRCT to determine the ability of the curriculum based measure to identify risk status.

Research question 3 correlated target scores from the 2011-12 winter M-CAP administration to the scores on the math portion of the 2011-2012 CRCT to determine the predictive ability of the curriculum based measure.

**Data Analysis**

Data analysis was exploratory and descriptive statistics were computed. To address the first hypothesis the two comparison groups were matched on risk status. Matching is a method to equate groups on variables which may be related to performance on the dependent variable (Gall, Gall, & Borg, 2007). This provided some control over variance in the 2011-2012 CRCT math scores. An analysis of covariance (ANCOVA) was used to attempt to determine if the independent variable, participation in computer based math instruction, caused a difference in the dependent variable while controlling for the test results on the 2010-2011 math subtest of the CRCT. Analysis of covariance is the appropriate test because in-tact groups were used and control over extraneous variables was not possible at the time. The ANCOVA attempted to control for the effects of the 2010-11 CRCT on the independent variable (Airasian, & Gay, 2002).

Data analysis for the second and third research questions utilized correlation analyses. The resulting Pearson Product Moment coefficients identified the strength and direction of the linear relationship of the 2011-12 fall and winter M-CAP scores to student scores on the 2011-2012 math portion of the CRCT. The Pearson correlation is the appropriate test because the covariant statistics were looked at through separate analyses to determine the degree to which the independent variables and the dependent variable varied together. The resulting correlation coefficients indicated the degree of the relationship and predictive ability of the independent variables on the dependent variable (Howell, 2011).

Statistical power increases with sample size and therefore the smaller the effect size necessary to reject the null hypothesis (Gall, Gall, & Borg, 2007). Fifty two student

scores were used to investigate the first hypothesis in this study and this was sufficient to reject the null hypothesis at the .05 significance level. The second and third hypotheses used 590 and 789 scores respectively for the two correlation analyses. The sample sizes were sufficient to reject the null hypotheses at the .01significance levels (Fowler, 1987). The correlation of determination ($r^2$) was used to measure effect size for the correlation analyses (Rosenthal, 1994). Assumptions made on the relationship between the variables were the independence of scores and normal distribution of scores. Normality, linearity, and homoscedasticity were checked through scatterplots and histograms (Howell, 2011).

## CHAPTER FOUR: FINDINGS

The purpose for this study was to determine if there was a statistically significant difference in CRCT math subtest scores between at-risk fifth grade students who participated in a computerized math intervention (I PASS) and at-risk fifth grade students who did not. This study also examined the ability of math curriculum based measures to identify risk status and to predict outcomes on the CRCT.

The research was based, in part, on the Georgia Department of Education's application for the ESEA Flexibility Waiver. This waiver continues to recognize the importance of annual gains made by students but also allows more flexibility in measures of accountability. In this respect it continues to endorse accountability decisions which are made through identifying skill levels, monitoring student growth, and providing research based interventions to students identified as at-risk to meet year end standards (ESEA, 2011). The Douglas County School District addressed components of this waiver by implementing curriculum based measures to progress monitor students' acquisition of basic skills and by providing research-based interventions and instruction to those students identified as at-risk to meet state standards.

The research based intervention in this study was iPASS, a computer based math program purported to increase math concepts and application skills of students performing below grade level and identified as at-risk to meet end of year standards. Based on Lowe's (2002) review of several meta-analyses she concluded that computer based instruction, in various designs, did have a positive effect on student outcomes when compared to general classroom instruction with effect size ranging from .17 to 1.13.

Kroesbergen and Van Luit (2003) conducted a meta-analysis on various math interventions and their impact on students with learning difficulties. A more robust impact was noted for direct instruction than for computer assisted instruction. Mixed results on various studies compelled this researcher to compare the impact of the iPASS computer-based math intervention on achievement outcomes for students identified as at-risk to meet state standards versus at-risk students not receiving the intervention.

Students' skill levels and trajectories of growth were assessed through a curriculum-based measure (CBM) administered in the fall and winter respectively. Deno (1992) reported on the indicators of basic skills, the measurement parameters, and criterion validity of CBM and concluded that the standardization, reliability and validity were technically adequate. Research into the efficacy of CBM was conducted by Wallace, et al. (2007) and varied as to scope with significantly more research focusing on reading than on math. As CBM math probes became more sophisticated in the nineties Hosp, et al. (2007) concluded it was due to movement from single skill facts and drills to applications and problem solving. Helwig, et.al (2002) theorized that any CBM that could estimate skill levels and risk status, as well as monitor progress toward summative assessments would be a very powerful educational tool. By examining the relationships of fall and winter AIMSweb M-CAP scores to the math subtest of the CRCT, this researcher adds to the gap in research that investigates the effectiveness of concept and application CBM to identify risk status and to predict achievement outcomes on state assessments.

The following research questions were generated to address these gaps and to add to the research literature:

Research Question 1: Is there a difference in the math CRCT scores between at-risk fifth grade students who participate in computer based math instruction and those who do not?

Research Question 2: Are the fall M-CAP cut scores able to identify fifth grade students as at-risk to fail the CRCT?

Research Question 3: Are the winter M-CAP progress monitoring target scores able to predict fifth grade students' achievement outcomes on the CRCT?

**Descriptive Statistics and Assessments of Normality**

For Research Question 1 an analysis of covariance was conducted to compare mean score differences on the CRCT between at-risk students who received computer based math instruction versus students who did not while controlling for the 2011 CRCT scores. The 2012 CRCT math subtest scores were compared from 52 students.

Table 1 presents descriptive statistics for the 2012 CRCT math subtest by group participation.

Table 1

*Descriptive Statistics for CRCT by Group*

| | Grade 5-Group 1 | | | | |
|---|---|---|---|---|---|
| CRCT | *n* | M | SD | Skewness | Kurtosis |
| | 26 | 790.19 | 25.82 | .29 | -.12 |
| | **Grade 5-Group 2** | | | | |
| CRCT | *n* | M | SD | Skewness | Kurtosis |
| | 26 | 798.00 | 15.59 | -.36 | -.60 |

Note: Group 1 = iPASS intervention. Group 2 = control-no iPASS intervention. CRCT = Criterion Referenced Competency Test-math subtest. CRCT scores range from 738 to 840. CRCT means and standard deviations are scaled scores.

*Figure 1. Histogram of 2012 CRCT Math Scores for Groups*

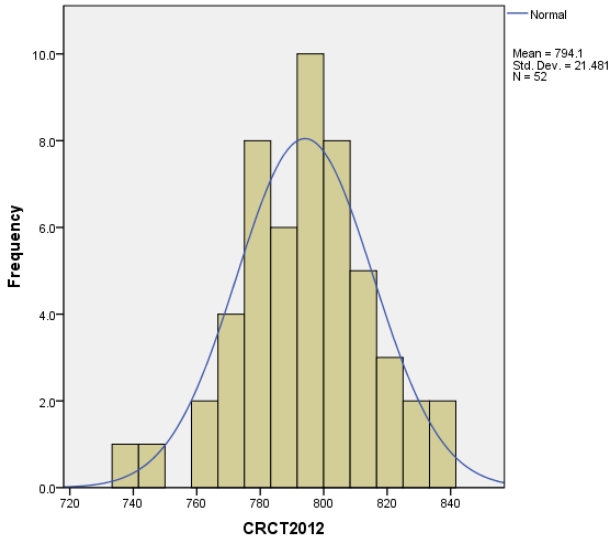
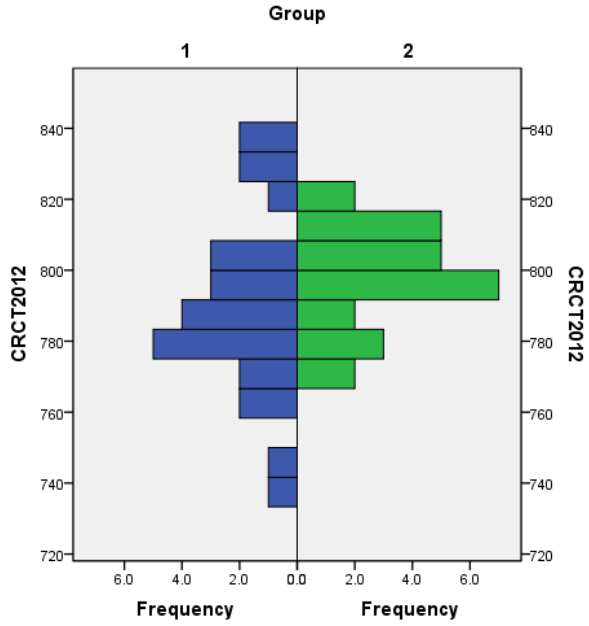



*Figure 2.* Histogram of Group 1 and Group 2 CRCT Scores




Measures of skewness and kurtosis assess normality with values of zero indicating normality and any values within -2 and 2 acceptable indices for normality (Howell, 2011). All measures on both histograms were within range and the assumption of normality was not violated (table 1; Figure 1 and 2). Therefore the ANCOVA was run.

For Research Question 2 a correlation analysis was conducted to examine the relationship between the fall M-CAP cut score and the CRCT math subtest score to determine if the cut score of eight or less was accurate as a predictor of scores of less than 800 on the CRCT. A total of 591 scores were examined.

Table 2 presents descriptive statistics for the fall M-CAP scores and the 2012 CRCT math subtest scores by group participation.

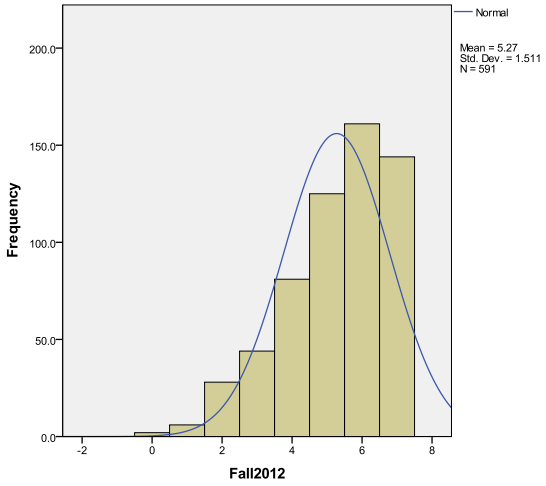
Table 2

*Descriptive Statistics for Fall M-CAP and CRCT*

| | Grade 5-At-Risk Students | | | | |
|---|---|---|---|---|---|
| M-CAP | *n* | M | SD | Skewness | Kurtosis |
| | 591 | 5.27 | 1.51 | -.80 | .09 |
| | Grade 5-At-Risk Students | | | | |
| CRCT | *n* | M | SD | Skewness | Kurtosis |
| | 590 | 811.84 | 29.79 | .60 | .88 |

Note: M-CAP At-Risk scores are cut scores of less than 8. CRCT = Criterion Referenced Competency Test-math subtest. CRCT means and standard deviations are scaled scores. One student took the M-CAP but did not take the CRCT.

*Figure 3*. Histogram of Fall Cut Scores



A negatively skewed distribution indicated that all scores were within range and the assumption of normality was not violated (Table 2; Figure 3).

For Research Question 3 a second correlation analysis was conducted to examine the relationship between the winter M-CAP target score and the CRCT math subtest score to determine if the target score of 10 or greater was a valid predictor of scores of 800 or greater on the CRCT. A total of 789 scores were examined.

Table 3 presents descriptive statistics for the winter M-CAP scores and the 2012 CRCT math subtest scores by group participation.

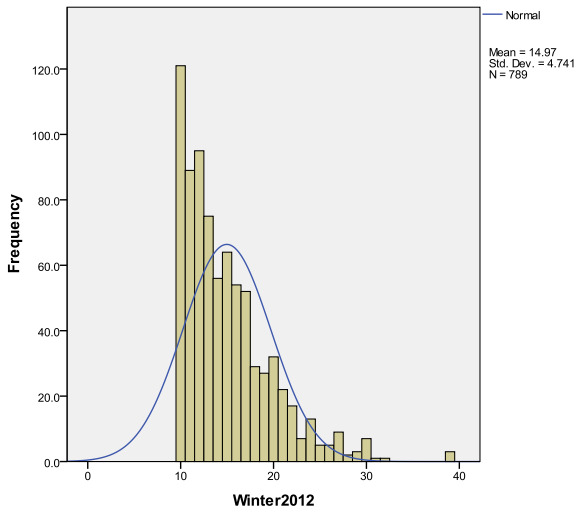
Table 3

*Descriptive Statistics for Winter M-CAP and CRCT*

| | Grade 5-At-Risk Students | | | | |
|---|---|---|---|---|---|
| M-CAP | *n* | M | SD | Skewness | Kurtosis |
| | 789 | 14.97 | 4.74 | 1.46 | 2.85 |
| | Grade 5-At-Risk Students | | | | |
| CRCT | *n* | M | SD | Skewness | Kurtosis |
| | 789 | 852.84 | 39.52 | .60 | 1.50 |

Note: M-CAP Target scores are scores of 10 or greater. CRCT = Criterion Referenced Competency Test-math subtest. CRCT means and standard deviations are scaled scores.

*Figure 4*. Histogram of Winter Target Scores



A positively skewed distribution indicated all measures were relatively within range and the assumption of normality was not violated (Table 3; Figure 4).

**Tests of Hypothesis**

Research Hypothesis 1: Fifth grade students identified as at-risk who participate in computer based math instruction compared to at-risk fifth grade students who do not participate in computer based math instruction will have statistically significant different scores as measured by the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

A test to evaluate the homogeneity-of-slopes assumption was conducted prior to running the ANCOVA. The results were not significant $F(1, 48) = .09$, $p = .77$ and indicated that the relationship between the 2011 CRCT and the 2012 CRCT did not differ significantly as a function of the Group (Appendix A, Table 9). Therefore the ANCOVA was robust to the assumptions of normality and homogeneity of variance.

Table 4

*ANCOVA for Group Scores and CRCT*

Dependent Variable: CRCT 2012

| Source | Type II Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 7109.65 | 2 | 3554.83 | 10.61 | .00 | .30 |
| Intercept | 1158.71 | 1 | 1158.71 | 3.46 | .069 | .07 |
| CRC 2011 | 6317.17 | 1 | 6317.17 | 18.85 | .00 | .28 |
| Group | 5.55 | 1 | 5.55 | .02 | .90 | .00 |
| Error | 16422.87 | 49 | 335.16 | | | |

As shown in Table 4 a one-way analysis of covariance (ANCOVA) was conducted. The independent variable included two levels: Group 1 = iPASS intervention and Group 2 = no intervention. The dependent variable was the math subtest score on the 2012 CRCT and the covariate was the math subtest score on the 2011 CRCT. With alpha set at .05, the ANCOVA was not significant, $F(1, 49) = .02$, $p = .90$.

The results of the ANCOVA indicated that there was not sufficient evidence to reject the null hypothesis, $F(1, 49) = .02$, $p > .05$. There was no statistically significant difference in at-risk fifth grade students' achievement scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test between students who participated in computer based math instruction and those who did not. The relationship between the 2011 CRCT scores and the 2012 CRCT scores was significant, $F(1, 49) = 18.85$, $p < .05$ with the 2011 CRCT accounting for approximately 28% of the variance in the 2012 CRCT when controlling for the intervention.

Research Hypothesis 2: At-risk fifth grade students' fall M-CAP cut scores will have a statistically significant relationship to their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

Table 5

*Correlation Matrix between Fall M-CAP and CRCT*

| | | Fall 2012 | CRCT |
|---|---|---|---|
| Fall 2012 | Pearson Correlation | 1 | .28** |
| CRCT | Pearson Correlation | .28** | 1 |

**p < .01

As shown in Table 5 a bivariate correlation was used to evaluate the null hypothesis that there was no relationship between the fall M-CAP scores and the final

math scores on the 2012 CRCT. Assumptions of linearity, bivariate normality and homoscedasticity were found tenable through scatterplots and due to sample size (n = 591). Results of the statistical testing indicated that there was a positive correlation and statistically significant relationship between the fall 2012 M-CAP (M = 5.27, SD = 1.51) and the math subtest of the 2012 CRCT (M = 811.84, SD = 29.79), r = .28, p<.01. This indicated that fall M-CAP scores of less than 8 were able to identify those students that were at risk to score less than 800 and fail the CRCT.

Table 6

*AIMSweb Fall Scores Cross-Tabulated with CRCT*

| | Failed CRCT | Passed CRCT | Total |
|---|---|---|---|
| AIMSweb <8 | 36.6% | 63.4% | 100% |
| AIMSweb ≥8 | 5.0% | 95.0% | 100% |
| | | | |

Results of the cross-tabulation of scores shown in Table 6 indicated that 37% of the students who scored below 8 failed the CRCT, while 63% of the students who scored below 8 passed the CRCT.

Research Hypothesis 3: Fifth grade students' winter M-CAP progress monitoring target scores will have a statistically significant relationship to their scores on the 2012 math subtest of the Georgia Criterion Referenced Competency Test.

A correlation analysis was conducted to examine the relationship between the winter M-CAP target scores and the CRCT math subtest scores to determine if the target

score of 10 or greater was a valid predictor of passing the CRCT with a score of 800 or greater. A total of 789 scores were examined.

Table 7

*Correlation Matrix between Winter M-CAP Scores and CRCT*

| | | Winter 2012 | CRCT 2012 |
|---|---|---|---|
| Winter 2012 | Pearson Correlation | 1 | .48** |
| CRCT | Pearson Correlation | .48** | 1 |

**p < .01

As shown in Table 7 a bivariate correlation was used to evaluate the null hypothesis that there was no relationship between the winter M-CAP scores and the final math scores on the 2012 CRCT. Assumptions of linearity, bivariate normality and homoscedasticity were found tenable from scatterplots and sample size (n = 789). Results of the statistical testing indicated that there was a positive correlation and statistically significant relationship between the winter 2012 M-CAP (M = 14.97, SD = 4.74) and the math subtest of the 2012 CRCT (M = 811.84, SD = 29.79), r = .48, p < .01. Results indicated that winter M-CAP scores of 10 or greater were able to identify students who would pass the CRCT.

Table 8

*AIMSweb Winter Scores Cross-Tabulated with CRCT*

| | Failed CRCT | Passed CRCT | Total |
|---|---|---|---|
| AIMSweb ≥10 | 5.8% | 94.2% | 100% |
| AIMSweb <10 | 37.2% | 62.8% | 100% |
| | | | |

Results of the cross-tabulation of scores in Table 8 indicated that 94% of the students who scored 10 or greater passed the CRCT and 63% of the students who scored less than10 passed the CRCT.

**Summary of Findings**

Hypothesis 1: Results of the ANCOVA were not significant $p > .05$. Evidence was insufficient and failed to reject the null hypothesis that there was a mean score difference in CRCT scores between the two groups as a result of participation in iPASS.

Hypothesis 2: Results of the Pearson correlation were significant $p < .01$ and evidence was sufficient to reject the null hypothesis that a significant relationship did not exist between the fall AIMSweb cut score and the CRCT. This indicated that the cut scores of less than 8 were able to identify students at-risk to fail the CRCT

Hypothesis 3: Results of the Pearson correlation were significant $p < .01$. Evidence was sufficient to reject the null hypothesis that a significant relationship did not exist between the winter AIMSweb target score and the CRCT which indicated that the target scores of 10 or greater were able to identify students who passed the CRCT.

## Chapter Five: Summary and Discussion

### Introduction

Recent state-wide education statistics indicated that 16% of fifth grade students in Georgia failed the math portion of the CRCT in 2012; the rate was 20% for Douglas County. This continues to be an alarming rate (Georgia Department of Education, 2012b). Mandates from the No Child Left Behind Act and the recent ESEA Flexibility Waiver endorsed concerted efforts by educators to identify students at risk to fail state assessments, to provide interventions to identified risk areas and to monitor student progress. Research indicates that screening students is the first step in identifying students who were at risk for learning difficulties (Clark, B. & Shinn, M. 2004; Petscher, et al. 2011; Safer, N. & Fleischman, S. 2005). Curriculum based measures have a long history of use and research into their practicality and accuracy in identifying students at risk and in predicting student outcomes (Deno,1992; Hosp, et al., 2007). When students are identified at-risk, steps must be taken to provide them with interventions that are research based and can remediate their deficits. Computer based instruction in math is experiencing resurgence as an effective tool that can be used by teachers to supplement classroom instruction, provide opportunities for drill and practice and recently to address deficits in concepts and applications. Programs such as iPASS, Compass Learning Odyssey Math, Key Math, and Accelerated Math are just a few of the ones used on the elementary level in schools across the nation (What Works Clearinghouse, 2010).

This research study was conducted to add to the research which investigated the relationship between curriculum based measures and state assessments and which examined the impact of computer based instruction on student outcomes. This study

tested the ability of the AIMSweb M-CAP curriculum based measure to identify students who may be at risk to fail the Georgia CRCT. Based on fall cut scores in the at-risk range, students were provided with additional math support through the iPASS program. This computer based program was developed to train students in the use of mathematical strategies so they can be successful with math reasoning and problem solving on assessments. This research tested the impact of this program on students' CRCT scores. This study also tested the ability of the AIMSweb winter M-CAP target scores to progress monitor skills and identify students who will pass the CRCT.

**Summary of the Research Results**

An initial step before addressing the hypotheses was to calculate descriptive statistics and to assess assumptions of normality. The mean, standard deviation, skewness and kurtosis were computed using SPSS 1.9 for the test data. Values of zero indicated normality and any values within -2 and 2 were acceptable indices for normality (Howell, 2011). Assumptions of normality were verified by examining scatterplots, histograms, skewness and kurtosis with no significant heteroscedasticity indicated; thus assumptions of normality and linearity were tenable. Sample sizes for the correlation analyses were sufficient to determine the Pearson Product Moment robust (Fowler, 1987). A test for the homogeneity of slopes was conducted for the ANCOVA which indicated that a linear relationship existed between the 2011 CRCT and the 2012 CRCT. The homogeneity of slopes assumption was not violated; therefore, no serious impact to the robustness of the statistical test of the ANCOVA (Poremba & Rowell, 1997).

Results of the ANCOVA statistical testing answered Research Question 1 and indicated that a statistically significant relationship did not exist between Group 1

students' CRCT scores compared to Group 2 CRCT scores. This resulted in sufficient evidence not to reject the null hypothesis and to conclude that no statistically significant mean score difference on the CRCT was detected between the two groups. Mean score differences on the CRCT indicated that despite participation in iPASS both groups failed to meet the end of year standard on the CRCT (Group 1 M = 790.19, Group 2 M = 798.00).

Results of the Pearson Correlation answered Research Question 2 and indicated that there was a statistically significant relationship between the fall M-CAP cut score and the CRCT score. A cut score of 8 was noted to be sensitive and able to identify students who failed the math portion of the end of year CRCT. From a total of 591 students, 215 students or 37% failed the 2012 CRCT.

Results of the Pearson Correlation answered Research Question 3 and indicated that there was a statistically significant relationship between the winter M-CAP target score and the CRCT. A target score of 10 was accurate in identifying 94 % of students who passed the 2012 CRCT.

**Discussion of Research Results**

One of the important goals of the ESEA Flexibility Waiver and education policymakers is the identification and use of innovative and effective methods and approaches to improve math achievement (ESEA Flexibility, 2011). The use of computer based software designed to increase math skills specifically in the area of concepts and applications has recently increased as a result of these formulated policies. While many research studies which focused on various aspects of math instruction indicated small to moderate effect size for the use of computers in increasing math achievement (Fletcher-

Finn, et al., 1995; Fuchs, et al., 2006; Kingston, 2009; Pollard, 2005); Wijekekumar et al. (2009) research on the effect of the Compass Learning Odyssey Math did not yield a statistically significant impact on end of year assessments. The researcher Mevarech, et al. (1991) reported that when computer based instruction was used on an individual basis without teacher support, it was found to have a negative impact on mathematics achievement. Linn, Podell, and Tournaki-Rein (1994) found no significant difference between CAI and computations using paper and pencil. Wilson, et al. (1996) examined CAI multiplication fact instruction against teacher directed and found greater effects for teacher instruction. As with most relevant research, the literature will consist of studies with significant results as well as those which are not. The statistical result for Hypothesis 1 was not significant and did not find that the use of iPASS made a difference between the groups on the CRCT. These results are similar to the results found by Wijekekumar et al. (2009). This indicated that the use of this intervention was not effective in remediating the math deficits that were identified on the fall M-CAP. Carnine (1997) suggested that one cause of continued math difficulties may be a poor fit between the learning style of the student and the type of instruction. Educational technology can be a powerful tool, but some interventions may be more effective in one skill area than another. Teachers need to be aware of the strengths and weaknesses as well as the learning characteristics of their students, especially those that are at risk. With this knowledge, teachers are better equipped to provide interventions that are effective for certain domains and motivating for the student. Doing otherwise is an unnecessary expenditure of time and resources.

Statistical results for Hypotheses 2 indicated that a significant relationship existed between the fall M-CAP cut score and the CRCT math subtest scores. The overall results of the correlation analysis aligned with previous research that endorsed the use of CBM as a research based tool that had demonstrated to be useful in the identification of student outcomes on state assessments (Deno, 1992; Fuchs, et al., 2008; Glover & Albers, 2007; Helwig, et al., 2002; Hosp, et al., 2007; Jiban & Deno, 2007). The statistical test results indicated a positive correlation between the fall M-CAP and the CRCT. It showed that scores of less than 8 on the fall M-CAP were not sufficient to earn a score of at least 800 on the CRCT and pass. The M-CAP fall cut score was accurate in identifying 37% of students who failed the CRCT while false positives were at 63%. The effect size of .08, explains 8% of the common variance between the M-CAP and the CRCT. Despite the magnitude of the effect size (Cohen, 1988; Rosenthal, 1994), over one third of the students identified failed the CRCT and this information can be critical to teachers. Hosp, et al. (2007), stated that the use of CBM evolved from intervention methods that were research-based and useful to teachers when making decisions about what and how to teach. If students who showed consistent and pervasive learning difficulties were identified, teachers could begin to develop preventive instructional programs. Jenkins, Hudson, and Johnson (2007) stressed the importance of early identification of students at risk as the key to proper placement into interventions. Therefore, information received at the beginning of the school year was critical if teachers hoped to provide the remediation necessary to reverse the student's course.

Statistical results for Hypothesis 3 indicated that a significant relationship existed between the M-CAP winter target score and the CRCT math subtest score. These results

extended the work that endorsed use of math CBM as a valid measure to predict student outcomes on state assessments (Hintz & Silberglitt, 2005; Shapiro, et al. 2006). Results from the statistical testing indicated that the winter target score of 10 or greater was accurate in identifying 94% of the students as on target and who passed the CRCT. The effect size explains 23% of the common variance between the M-CAP and the CRCT. Overall, the winter M-CAP target score was a better predictor of student outcomes than the fall M-CAP cut score which expanded research results from Jitandra et al. (2007) and Shapiro (2006). Knowing this, teachers can use the target score as a benchmark for monitoring student gains. By recognizing that they have to be at a specific level by the end of the first semester to pass the CRCT they can adjust and differentiate instruction. Hosp, et al. (2007) posited that progress rates were not ceilings and that students can always make greater progress as long as the instruction meets the needs of the students. Since the M-CAP target score was an accurate predictor of the CRCT, teachers can know at the beginning of the school year the amount of growth their students must make in order to meet the standard and pass it.

Overall results of the correlation analyses indicated that the M-CAP fall and winter scores were useful indicators of students who were at risk to fail and those who were on target to pass state assessments. While the correlations were not as robust as in previous studies (Foegen & Deno, 2001; Wallace, et al. 2007; Keller-Margulis, et al. 2008) they demonstrated to be useful measures that teachers can use to identify and monitor student progress and drive instruction.

The researchers Silberglitt and Hintz (2005) examined four methods used to establish binary classifiers useful in academic assessments. According to their research,

receiver operating characteristic (ROC) analysis was most accurate and flexible when setting levels. Their analysis examined the trade-off between the sensitivity or true positive and specificity the true negative of set scores. ROC analysis was conducted by the developers of the AIMSweb M-CAP to identify scores of 8 and 10 as classifiers of risk and on target to pass state assessments (Edformation, 2002). Results of the cross-tabulations in this study showed that the fall and winter scores of the M-CAP had a high percentage of false positives. Fuchs, et al. (2003) determined that a measure was most effective when procedures for identifying risk were implemented which produced a high percentage of true positives while minimizing false positives. One way to reduce the number of false positives may be through teachers' consistent progress monitoring over intervals and use of instructional strategies that prove effective in remediating deficits.

At this time, the state of Georgia is debating whether or not to adopt the Common Core Standards (CCS) and re-develop the end of year assessment. The Douglas County School District is in the process of developing a bank of formative Common District Assessments (CDA) that will align with the CCS. Information from this research will be timely and useful as they look to align the various formative assessments with the state assessment and develop a progress monitoring system that can be linked with effective interventions and which is designed for student success. Having this information, teachers can ensure that the CDA will supplement the information gathered from the M-CAP. This will eliminate duplication and unnecessary testing of students and additional constraints on teacher time. If teachers see that the information is useful and not just duplication they may be more receptive to administer the formative assessments

consistently and with fidelity, keep running records of student data, and monitor the effectiveness of prescribed interventions.

**Limitations**

This study was conducted in a small district outside Atlanta and may represent very different demographics not representative of districts throughout Georgia and elsewhere. Scores from fifth grade students were used in this study with a limited age group and results may not be generalizable to other grades or age groups. States have various end-of-year criterion measures; use of the Georgia Criterion Referenced Competency Test may affect generalization to other states using different assessments.

Subgroups were not disaggregated and examined separately. Students with Individual Education Plans often receive various and specific accommodations for testing; therefore, special education students who received accommodations in math on the CRCT were excluded from the analyses.

One school using the iPASS program was not able to access the software for 2 weeks in February due to computer maintenance. Students rotated computers but instructional time was reduced. One fifth grade classroom in the non-iPASS group had a long-term substitute for the second semester of the school year; this may have had an impact on the methods of instruction during math class.

Non-homeroom teachers administered the M-CAP; this may have affected student participation in the testing.

AIMSweb M-CAP scores had a high rate of false positives.

**Implications for Practice**

Despite limitations to this study, there were important implications. Lynch (2006) asserted that technology use in the classrooms could transform learning. While results from the ANCOVA on computer based math instruction were not significant, teachers should continue to integrate use of math software into daily lessons. Research by Jaspers and Van Lietshout (1994) found that use of computer software for math problem solving and representations was best facilitiated with teacher instruction.

In the current study, iPASS instruction was monitored by the computer lab teacher and not the participants' math teachers. Best practice would be for the math teachers of the students to continuously monitor their performance in iPASS by examining their iPASS score reports. This would ensure that classroom instruction was tailored to address any areas of persistent difficulty.

Results from the correlation analyses were promising. Teachers have additional information on how students will perform on the CRCT. Teachers will know within the first month which students lack the skills necessary to be successful on the CRCT. They will also know at the midpoint of the year which students are showing positive academic growth and on target to pass the CRCT. The researcher Cusumano (2007) indicated that CBM can provide systematic data concerning student skill development. This data can be used by teachers to provide interventions or enrichment. Results of this study identified 37% of students in the first month with scores of 8 or less as at-risk. At midpoint, scores of 10 or greater were accurate in predicting outcomes on the CRCT for 94% of the students. Additional administrations of the M-CAP would provide teachers with a running record of student performance. Researcher Fuchs (2003) reported when teachers

used consistent progress monitoring to track students they were more accurate identifying students who required extra or different instruction, they were able to develop more effective programs, and as a result their students were able to achieve more. Curriculum based measures were designed to be formative assessments and the M-CAP should be used frequently to monitor students who have persistent difficulties with math. VanDerHayden & Burns (2009) recommended use of CBM to monitor student skills after use of interventions. Frequent use of the M-CAP to monitor effects of iPASS would be prudent. The ability to predict student performance and the alignment of teacher instruction and provision of interventions based on this information is the key to student success. When teachers know they have a measure like the M-CAP that is able to identify risk and predict performance they can use that information to modify their instructional strategies to better address student needs. As Cusumano (2007) stated, " CBM has gained a strong footing as a metric for monitoring student academic progress…..it is critical that all individuals that work with children understand its use in the school as both a screening instrument that identifies risk…..and as a metric that monitors students' acquisition of skills (p. 30).

**Recommendations for Future Research**

Although an abundance of research exists on student motivation, studies should be conducted to look at student motivation and the use of math software. Since technology use is so prevalent, it may be that students are satiated with its use or that the software is not motivating enough to keep their interests. Additional experimental research conducted in a controlled setting should be conducted to determine if iPASS instruction can, in fact, remediate math deficits.

As indicated by the test developers, work continues to ready the M-CAP probes for technical review by the National Center on Student Progress Monitoring (Edformation, 2012). Among Schatschneider, et al. (2008) considerations for a CBM screening measure is the need to identify what is meant by at-risk, to establish goals for the screening at each interval, and to increase instances of positive predictive power while decreasing negative power. In light of the number of false positives with the cut and target scores, additional research into the sensitivity/specificity of these designated scores and their relationship to the criterion validity, constructs of achievement and content of skills on the CRCT should be undertaken to determine the best trade off between true positives and false positives.

The local district's Department of Professional Development should provide additional workshops into the use of M-CAP as a risk identification and prediction tool. Qualitative studies should be undertaken to determine how best to align data from M-CAP with classroom instruction, how to use M-CAP data to prescribe interventions, and how to provide systematic procedures to progress monitor their effectiveness. Researchers Yell, et al. (1992) found that teachers overwhelmingly viewed lack of time as a barrier to frequent administration of CBM; despite teachers stating that it took less than 10% of time from instruction. If teachers have comprehensive instruction into the capabilities of CBM, it may lessen their reluctance to consistently monitor their students' progress and may reduce the negative view that some teachers have concerning the amount of time needed to complete all the testing required during the school year. Their input could prove to be the most valuable when deciding on the most efficient and cost

effective instructional model which eliminates duplication and provides the greatest opportunities for student success.

**Conclusion**

The results of this research are promising because data from the M-CAP can provide teachers with on-going information concerning their students' status in relation to the CRCT. This information is critical if teachers are to address the mandates of NCLB and provide their students with a quality education. Schools must continue to identify those students who are likely to fail the state assessment. By the same token they should take precautions to not over or under identify risk or proficiency; both would result in a waste of scarce resources.

Teachers should be receptive to using all tools at their disposal to provide instruction that is designed for each student's needs. Research has shown that student achievement improves in direct relation to effectiveness of a teacher's instructional design (Foegen, A., Jiban, C., & Deno, S. 2007; Fuchs, L., Deno, S., & Mirkin, P. 1984; Fuchs, L., Fuchs, D., Hamlett, C., & Stecker, P. 1991).

Results of this study will provide teachers with additional information to align data from the M-CAP to state curriculum standards and the CRCT. This will ensure that each student's skill level is recognized and that steps are taken to increase every student's chance to excel.

REFERENCES

AIMSweb (2009a). *AIMSweb progress monitoring and improvement system*. Retrieved from http://www.aimsweb.com/

AIMSweb (2009b). *Mathematics Concepts and Applications: Administration and Technical Manual*. San Antonio, TX: Pearson, Inc.

Airasian, P.W., & Gay, L. R. (2002). *Educational Research: Competencies for Analysis and Application* (7th ed.). New York, NY: Prentice-Hall

Albers, C., Glover, T., & Kratochwill, T. (2007). Where are we, and where do we go now? Universal screening for enhanced educational and mental health outcomes. *Journal of School Psychology, 45*, 257-263.

Anderson, J. R. (1990). *Cognitive psychology and its implications* (3rd ed.). New York, NY: W. H. Freeman.

Bevevino, M. M., Dengal, J., & Adams, K. (May-June, 1999). Constructivistic theory in the classroom: Internalizing concepts through inquiry learning. *The Clearing House, 72*(5), 275-278.

Bloc, H., Oostdam, R., Otter, M., & Overmaat, M. (2002). Computer assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72*(1), 101-130. doi: 10.3102/00346543072001101

Burns, M. K. (2002). Utilizing a comprehensive system of assessment to intervention using curriculum based assessments. *Intervention in School and Clinic, 38*, 8-13.

Butler, A. C. (2010). Repeated testing produces superior transfer relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118-1133.

Camli, H., & Bintas, J. (2009). Mathematical problem solving and computers: Investigation of the effect of computer aided instruction in solving lowest common multiple and greatest common factor problems. *International Journal of Human Sciences, 6*(2), 348-356.

Cargill, D. (2009). Know the ROPES. Retrieved from: http://www.learning focused.com/index.php/resources/newletter/know-the-ropes

Carnine, D. (1997). Instructional design in mathematics for students with learning disabilities. *Journal of Learning Disabilities,* 30, 130-141.

Carraher, D. W., & Schleimann, A. D. (2002). The transfer dilemma. *The Journal of the Learning Sciences, 11*(1), 1-24.

Chan, J. C. K. (2010). Long-term effects of testing on the recall of nontested materials. *Memory, 18*(1), 49-57. doi: 10.1080/09658210903405737

Chang, K-E., Sung, Y-T., & Lin, S-F. (2004). Computer-assisted learning for mathematical problem solving. *Computers & Education, 46,* 140-151.

Clark, B., & Shinn, M. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33*(2), 234-248.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hilldale, NJ: Erlbaum Associates

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

Cusumano, D. L. (2007). Is it working? An overview of curriculum based measurement and its uses for assessing instructional, intervention, or program effectiveness. *The Behavior Analyst Today, 8*, 24-34.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219-232. [as cited in Deno, 2003].

Deno, S. L. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure: Alternative Education for Children and Youth, 36*(2), 5-10. doi: 10.1090/1045988X.1992.9944262

Deno, S. L., Reschly-Anderson, A., Lembke, E., Zorka, H., & Callender, S. (2002). *A model for school wide implementation: A case example*. Paper presented at the annual meeting of the National Association of School Psychology, Chicago, IL. [as cited in Deno, 2003].

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37,* 184-192.

Eckert, T. L., Dunn, E. K., Codding, R. S., Begeny, J. C., & Kleinmann, A. E. (2006). Assessment of mathematics and reading performance: An examination of the correspondence between direct assessment of student performance and teacher report. *Psychology in the Schools, 43*(3), 247-265. doi: 10.1002/pits.20147

Edformation. (2012). *AIMSweb Progress monitoring and improvement system*. Retrieved from http://www.aimsweb.com/

Elliot, S. N., Huai, N., & Roach, A. T. (2007). Universal and early screening for educational difficulties: Current and future approaches. *Journal of School Psychology, 45,* 137-161.

ESEA Flexibility. (2011). *What ESEA flexibility means for students, teachers, and parents: Answering the public's questions.* Retrieved from: http://www.ed.gov/esea/flexibility

ESEA Flexibility Waiver. (2011). *Elementary and Secondary Education Act (ESEA) of 1965*. Retrieved from http://www.doe.k12.ga.us/External-Affairs-and Policy/communications/pages/NCLB-Waiver.uspx

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861-974.

Fletcher-Flinn, C. M., & Gravatt, B. (1995). The efficacy of computer assisted instruction (CAI): A meta-analysis. *Journal of Educational Computing Research, 12*(3), 219-241.

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

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

Fowler, R. L. (1987). Power and robustness in product-moment correlation. *Applied Psychological Measurement, 11*(4), 419-428.

Fuchs, L. S., Deno, S., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research, 21*(2), 449-460.

Fuchs, L. S., & Deno, S. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, 57, 488-500.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research, 28*(3), 617-641.

Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice, 18*(3), 172-186.

Fuchs, L., Fuchs, D., Hamlet, C., Powell, S., Capizzi, A., & Seethaler, P. (2006). The effects of computer-assisted instruction on number combination skill in at-risk first graders. *Journal of Learning Disabilities, 39*(5), 467-475.

Fuchs, L., Fuchs, D., Compton, D., Hamlett, C., & Seethaler, P. (2007). Mathematics screening and progress monitoring at first grade: Implications for responsiveness to intervention. *Exceptional Children*, 73, 311-330.

Fuchs, L., Fuchs, D., & Zumeta, R. (2008). A curricular-sampling approach to progress monitoring: Mathematics concepts and applications. *Assessment for Effective Intervention, 33*(4), 225-233. doi:10.1177/1534508407313484

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational Research* (8th ed.). Boston, MA: Pearson

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

Geary, D., Hoard, M., Nugent, L., & Byrd-Craven, J. (2009). Strategy use, long-term memory, and working memory capacity. In D. B. Berch, and M. M. M. Mazzocco (Eds.). *Why is Math so Hard for Some Children? The Nature and Origins of*

*Mathematical Learning Difficulties and Disabilities* (pp. 83-105). Baltimore, MD: Brooks.

Georgia Department of Education (2004). Georgia Accountability Workbook. Retrieved from http://public.doe.k12.ga.us/Curriculum-Instruction-and-Assessment/Accountability

Georgia Department of Education (2012a). Criterion Referenced Competency Tests. Retrieved from http://www.gadoe.org./Curriculum-Instruction-and-Assessment/Pages/CRCT-Content-Descriptions.aspx

Georgia Department of Education (2012b). Criterion Referenced Competency Tests State Statistics. Retrieved from htt://www.doe.k12.ga.us/External-Affairs-and-Policy/communications/Pages/PressReleaseDetails.aspx?PressView=default&pid=46

Gersten, R., Clarke, B., & Mazzocco, M. (2007). Historical and contemporary perspectives on mathematical learning disabilities. In D. B. Berch & M. M. M. Mazzocco (Eds.). *Why is Math so Hard for Some Children? The Nature and Origins of Mathematical Learning Difficulties and Disabilities* (pp.7-27). Baltimore, MD: Brooks.

Gersten, R., Chard, D. J., Madhavi, J., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*(3), 1202-1242.

Gersten, R., Clarke, B., Newman-Gonchar, R., Haymond, K., & Wilkins, C. (2012). Universal screening in mathematics for the primary grades: beginnings of a research base. *Exceptional Children, 78*(4), 423-445.

Glover, T., & Albers, C. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135.

Goldman, S. (1989). Strategy instruction in mathematics. *Learning Disabilities Quarterly, 12*(1), 43-55.

Hall, T., Strangman, N., & Meyer, A. (2011). *Differentiated instruction and implications for UDL implementation*. Wakefield, MA: National Center on Accessible Instructional Materials. Retrieved from: http://aim.cast.org/learn/historyarchive/backgroundpapers/differentiated_instruction_udl

Hanich, L., Jordan, N., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning disabilities. *Journal of Educational Psychology, 93*(3), 615-626.

Harcourt (2006). *Harcourt assessment acquires Edformation*. Retrieved from: http://www.pearsonassessment.com

Haskell, E. H. (2001). *Transfer of Learning: Cognition, instruction, and reasoning*. New York, NY: Academic Press.

Haughton, E. C. (1980). Practicing practices. *Journal of Precision Teaching, 1*, 3- 20. [Cited in Van Der Heyden & Burns, 2009].

Helwig, R., Anderson, L., & Tindal, G. (2002). Using a concept-grounded, curriculum-based measure in mathematics to predict statewide test scores for middle school students with LD. *The Journal of Special Education, 36*(2), 102-112. doi: 10:1177/00224669020360020501

Hintz, J. M., Christ, T. J., & Keller, L. A. (2002). The generalizability of CBM survey level mathematics assessments: Just how many samples do we need? *School Psychology Review*, *31*, 514-528.

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

Howell, D. (2011). *Fundamental Statistics for the Behavioral Sciences* (7th ed.). Belmont, Ca: Wadsworth.

iLearn: (1989). Retrieved from website: http://www.ilearn.com/AboutiPass.asp

iLearn: (2010). *Research base for the design of iPASS*. Retrieved from website: www://iLearn.com/pdf/iPASS_ResearchBase.pdf

Jaspers, M. W. & Van Lieshout, E. D. (1994). A CAI program for instructing text analysis and modeling of word problems to educable mentally retarded children. *Instructional Science*, 22, 115-136.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for service delivery in an RTI framework: Candidate measures. *School Psychology Review*, 36, 582-599.

Jiban, C., & Deno, S. (2007). Using math and reading curriculum-based measurements to predict state mathematics test performance: Are simple one-minute measures technically adequate? *Assessment for Effective Intervention, 32*(2), 78-108. doi: 10.1177/1534508407032002050

Jitendra, A., Griffin, C., McGoey, K., Gardill, M., Bhat, P., & Riley, T. (1998). Effects of mathematical word problem solving by students at risk or with mild disabilities. *Journal of Educational Research, 91*(6), 345-355.

Jitendra, A. K., 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.

Kaput, J. J., & Thompson, P. W., (1994). Technology in mathematics research: The first 25 years in the JRME. *Journal for Research in Mathematics and Education, 25*(6), 676-684.

Keller-Margulis, M., Shapiro, E., & Hintz, J. (2008). Long term diagnostic accuracy of curriculum based measures in reading and mathematics. *School Psychology Review*, 37, 374-390.

Kingston, N. M. (2009). Comparability of computer-and paper-administered multiple-choice tests for k-12 populations: A synthesis. *Applied Measurement in Education*, *22*, 22-37.

Kroesbergen, E. H., & Van Luit, J. (2003). Mathematics interventions for children with special needs: A meta-analysis. *Remedial and Special Education, 24*(2), 97-114.

Kulik, J. A., & Kulik, C. L. C. (1987). Review of recent research literature on computer-based instruction. *Contemporary Educational Psychology, 12*, 222-230.

Learning Focused Schools. (2007). *Learning focused solutions*. Retrieved from: http://www.learningfocused.com

Liao, Y-K. C. (1999). Effects of hypermedia on students' achievement: A meta-analysis. *Journal of Educational Multimedia and Hypermedia, 8*(3), 255-277.

Linn, A., Podell, D., & Tournakie-Rein, N. (1994). CAI and the development of automaticity in mathematics skills in students with and without mental handicaps. *Computers in the Schools*, 11, 43-58.

Linn, R. L., Baker, E. L., & Betetebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31*(6), 3-16.

Lowe, J. (2002). Computer-based instruction: Is it a panacea? *Journal of Research on Technology in Education, 34*(2), 163-171.

Lynch, J. (2006). Assessing effects of technology usage on mathematics learning. *Mathematics Education Research Journal*, *18*, 29-43.

Mevarech, Z. R., Silber, O., & Fine, O. (1991). Learning with computers in small groups: Cognitive and affective outcomes. *Journal of Computing Research*, 2, 233-243.

Marzano, R., Zeno, B., & Pollack, J. (1999). *Research into practice: Assessment, grading and record keeping in the classroom*. Aurora, CO: McREL.

Montague, M. (1997). Cognitive strategy instruction in mathematics for students with learning disabilities. *Journal of Learning Disabilities, 30*(2), 164-177.

National Center for Education Statistics. (2009). *The Nations Report Card: Fast Facts*. Retrieved from http://nces.ed.gov/fastfacts/display.asp?id=514

National Center for Education Statistics. (2012). School Demographics Report. Retrieved from: http://nces.ed.gov/survey/sdds/index/aspx

National Center on Student Progress Monitoring (NCSPM, 2007). Retrieved from:

http://www.studentprogress.org/weblibrary.asp

National Commission on Excellence in Education. (1983). *A Nation at Risk: The imperative for educational reform*. Washington, DC: Government Printing Office. Retrieved from http://www2.ed.gov/pubs/NatAtRisk/risk.html

National Council of Teachers of Mathematics. (2012). *Differentiated instruction resources*. Retrieved from: http://www.nctm.org/resources/content.aspx?22753

No Child Left Behind Act of 2001, Pub. L. No. 107-110. (2001). Retrieved from http://www2ed.gov/policy/elsec/leg/esea02/107-110.pdf

Norris, C., Smolka, J., & Soloway, E. (2000). Extracting value from research: A guide for the perplexed. *Technology and Learning, 20*(11), 45-48.

Ormond, J. E. (1995). *Human Learning* (2nd ed.). Engelwood Cliffs, NJ: Merrill

Perkins, D. N., & Salomon, G. (1992). *Transfer of Learning*. Contribution to the International Encyclopedia of Education (2nd ed.). Oxford, England: Pergamon Press.

Petscher, Y., Young-Suk, K., & Foorman, B. R. (2012). The importance of predictive power in early screening assessments: Implications for placement in the response to intervention framework. *Assessment for Effective Intervention, 36*(3), 158-166.

Pollard, C., & Pollard, R. (2005). Research priorities in educational technology: A delphi study. *Journal of Research on Technology and Education, 37*(2), 145-160.

Poremba, K., & Rowell, R. (1997). *Testing for the homogeneity of slopes analysis: A tutorial.* Paper presented at the annual meeting of the Southwest Educational Research Association. Austin, Tx. Retrieved from ERIC database ED409319.

Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., & Evans, E. (1989). The challenges of classroom strategy instruction. *The Elementary School Journal, 89*(3), 301-342.

Reeves, D. (2003). *High performance in high poverty schools: 90/90/90 and beyond*. Retrieved from: http://www.sjboces.org/nisl/high%20performance%2090%2090%2090%20and%20beyond.pdf

Rosenthal, R. (1994). Parametric measures of effect size. In Cooper, H., & Hedges, L. (eds.). *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation [Cited in Howell, 2011].

Safer, N., & Fleischman, S. (2005). Research matters/How student progress monitoring improves instruction. *How Schools Improve, 62*(5), 81-83.

Salvia, J., Yesseldyke, J. E., & Bolt, S. (2007). *Assessment in special and inclusive education (*10th ed.). Boston, MA: Houghton Mifflin.

Schatschneider, C., Petscher, Y., & Williams, K. (2008). How to evaluate a screening process: The vocabulary of screening and what educators need to know. In L. Justice & C. Vukelich (Eds.) *Achieving excellence in preschool literacy instruction*. (pp. 304-316). [Cited in Petscher, et al., 2011).

Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintz, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests. *Journal of Psychoeducational Assessment, 24*(1), 19-35.

Shinn, M. R. (1998). *Advanced applications of curriculum-based measurement*. New York, NY: Guilford Press.

Shinn, M. R. (2008). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology V* (p. 243-261). Bethesda, MD: National Association of School Psychologists

Silberglitt, B., & Hintz, 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*(4), 304-325.

Shuell, T. J. (1986). Cognitive concepts of learning. *Review of Educational Research*, *56*(4), 411-436.

Thompson, M. (2006). Learning-focused research. Retrieved from: http://learningfocused.com

Thompson, M. (2007). *What is learning-focused?* Retrieved from: http://learningfocused.com/index.php/resources/

Tindal, G., Germann, G., & Deno, S. (1983). *Descriptive research on the Pine County norms: A compilation of findings*. (Research Report No. 132). Institute for Research on Learning Disabilities. [Cited in AIMSweb Technical Manual, 2009).

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review*, *30*, 363-382. doi: 10.1177/073724770503000203

Tournaki, N. (2003). The differential effects of teaching addition through strategy instruction versus drill and practice to students with and without learning disabilities. *Journal of Learning Disabilities, 36*, 449-458.

VanDerHeyden, A., & Burns, M., (2009). Performance indicators in math: Implications for brief experimental analysis of academic performance. *Journal of Behavior Education, 18,* 71-91. doi: 10.1007/sl0864-0009-908-x

Wallace, T., Espin, C. A., McMaster, K., Deno, S. L., & Foegen, A. (2007). CBM progress monitoring within a standards-based system: Introduction to the special series. *Journal of Special Education, 4*(2), 66-67.

What Works Clearninghouse, Eds., (2010). What works clearinghouse intervention report. Retrieved from: http://ies.ed.gov/ncee/wwc/Publications_Reviews.aspx?F=All%20Publication%20and%20Review%20Types,1,#pubsearch

Wijekumar, K., Hitchcock, J., & Turner, H. (2009). *A multisite cluster randomized trial of the effects of CompassLearning Odyssey[R] Math on the math achievement of selected grade 4 students in the mid Atlantic region*. Final Report. National Center for Education Evaluation and Regional Assistance. NCEE 2009-4068. Retrieved from ERIC database. (ED507314)

Williams, L. P. (2000). *The effects of drill and practice software on multiplication skills*: "*Multiplication Puzzles" versus "The Mad Minute*." (Masters thesis. Johnson Bible College, 2000). Retrieved from ERIC database. (ED443706)

Wilson, R., Majsterek, D., and Simmons, D. (1996). The effects of computer-assisted versus teacher-directed instruction on the multiplication performance of elementary students with learning disabilities. *Journal of Learning Disabilities, 29*(4), 382-390.

Yell, M., Deno, S., & Marston, D. (1992). Barriers to implementing curriculum-based measurement. *Diagnostique*, 18, 99-105.

APPENDIX A

Table 9
ANCOVA: Test of Homogeneity of Slopes Assumption

Dependent Variable: CRCT 2012

| Source | Type II Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 7139.568[a] | 3 | 2379.856 | 6.968 | .001 | .303 |
| Intercept | 577.952 | 1 | 577.952 | 1.692 | .200 | .034 |
| Group | 29.478 | 1 | 29.478 | .086 | .770 | .002 |
| CRCT 2011 | 4521.253 | 1 | 4521.253 | 13.239 | .001 | .216 |
| Group*CRCT 2011 | 29.915 | 1 | 29.915 | .088 | .769 | .002 |
| Error | 16392.952 | 48 | 341.520 | | | |

a. R Squared = .303 (Adjusted R Squared = .260)

The interaction is not significant, $F(1, 48) = .088$, $p = .77$. Partial $\eta^2$ of .002 is of small size indicating that in the sample the mean differences in the 2012 CRCT test scores did not vary as a function of the 2011 CRCT test scores. The assumption of homogeneity of slopes is tenable.

May 2, 2013

Katherine Tomaszewski Arnold
IRB Exemption 1591.050213: A Comparison of Georgia Criterion Referenced Competency Test Math Scores Between At-Risk Fifth Grade Students Receiving Computer Based Math Instruction and At-Risk Students Not Receiving Computer Based Math Instruction

Dear Katherine,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and that no further IRB oversight is required.

Your study falls under exemption category 46.101 (b)(4), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:

> (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Please note that this exemption only applies to your current research application, and that any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.

If you have any questions about this exemption, or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at irb@liberty.edu.

Sincerely,

**Fernando Garzon, Psy.D.**
*Professor, IRB Chair*
**Counseling**

**(434) 592-4054**

LIBERTY UNIVERSITY.

*Liberty University | Training Champions for Christ since 1971*

# Douglas County School System

P.O. Box 1077 ~ Douglasville, GA 30133 ~ 770-651-2000 ~ www.douglas.k12.ga.us

Dr. Gordon Pritz, Superintendent


November 26, 2012

Katherine Arnold
ESES/BUES

Dear Ms. Arnold:

Permission is granted for you to conduct your research study. The title of your study is "*A Comparison of Georgia Criterion Referenced Competency Test Math Scores Between At-Risk Fifth Grade Students Receiving Computer Based Math Instruction and At-Risk Students Not Receiving Computer Based Math Instruction.*" All information to be gathered will be done in a confidential and appropriate manner. The Douglas County School System is to receive a copy of all completed research findings.

Sincerely,

Mrs. Pam Nail
Associate Superintendent for Student Achievement & Leadership