8-2013

Predicting Performance on the Tennessee Comprehensive Assessment for Reading for Third Graders using Reading Curriculum Based Measures

Robert S. Kirkham
East Tennessee State University

Predicting Performance on the Tennessee Comprehensive Assessment for Reading for Third Graders using Reading Curriculum Based Measures

_____

A dissertation

Presented to

The faculty of the Department of Educational Leadership and Policy Analysis

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Doctor in Education in Educational Leadership

_____

by

Robert Scott Kirkham

August 2013

_____

Dr. James Lampley, Chair

Dr. Cecil Blankenship

Dr. Virginia Foley

Dr. Don Good

Keywords: Reading Curriculum Based Measure, Response to Intervention, Tennessee Comprehensive Assessment Program, Reading First, Reading

ABSTRACT

Predicting Performance on the Tennessee Comprehensive Assessment for Reading for Third Graders using Reading Curriculum Based Measures

by

Robert S. Kirkham

Despite flexibility waivers granted to states by the United States Department of Education from some provisions of the No Child Left Behind Act, our nation's public schools continue to struggle to improve reading proficiency as measured by high stakes assessments. To reach state targets for reading proficiency schools must use data at the earliest point possible to inform instructional strategies and identify students at risk of failure. The response to intervention model holds promise for improving reading outcomes particularly for early elementary students.

The effective use of reading curriculum based measures (R-CBM) to determine if instruction is adequate to produce students who score proficient or advanced on state mandated reading assessments is critical to achieving the goals for student learning. The population selected for this study included all third grade students from an East Tennessee school district. The third graders attended 13 schools and included 911 third grade students of which 770 students participated in the study. This included 372 male and 398 female students. Approximately 47% of the students were economically disadvantaged as determined by qualifying for free and reduced priced meals.

The purpose of this study was to investigate the relationship between 4 predictor variables (fall R-CBM, winter R-CBM, spring R-CBM, and median R-CBM) and the Tennessee Comprehensive Assessment Program (TCAP) third grade reading and language arts assessment. Each data set included 4 R-CBM scores expressed in words read correctly and TCAP reading language arts scale scores. Gender and free and reduced price meals eligibility information for all third graders from the 2010-2011 school year were also collected. Results reflected a strong predictive relationship between the AIMSweb R-CBM and TCAP reading and language arts measure for third grade students. Zero order correlations in the multiple regression analysis ranged from .70 to .74 for the 4 predictor variables. A linear equation was developed to predict TCAP scores from a single R-CBM score (fall, winter, spring, and median). Based on this study practitioners may be able to establish goals for student reading that are strongly correlated with achieving proficiency on the TCAP reading and language arts assessment.

DEDICATION

This dissertation is dedicated to my family and the exceptional students, teachers, mentors, and friends who have guided me throughout my life and career.

"Keep away from those who try to belittle your ambitions. Small people always do that, but the really great make you believe that you too can become great."-- Mark Twain

ACKNOWLEDGMENTS

I have been blessed to have so many incredible people in my life. Without you I would not be writing this acknowledgment.

I would like to thank my Mom and Dad for instilling in me the value of a great education and a responsibility to help others. You always believed that David and I would do great things with our lives. Thank you Nicole for insisting I finish this dissertation and allowing me the time I needed to make progress. I look forward to spending more time with my family with this capstone project complete. Thank you Collin, Adam, Katelyn, and Emily for all the time you have given me to work on this dissertation, when you would have rather been playing with me. If it wasn't for you guys I likely would not have finished this program. I did not want you to see me give up on something important. Please never give up on your dreams even if it seems like it will never happen.

Thank you Dr. James Lampley, without your guidance I would not have made it this far. I appreciate the great spans of time you were able to allocate to me at critical points in this project. I am thankful that you agreed to oversee this study. Thank you to my committee members, Dr. Cecil Blankenship, Dr. Virginia Foley, and Dr. Don Good. Your thoughtful guidance and mentorship will always be appreciated. I would like to thank BettyAnn Proffitt who was the first person I spoke to at ETSU and you made me feel at home instantly.

Thank you, Dr. Susan Lee, Dr. Martha Murray, Dr. Alisa Teffeteller, Dr. Stephanie Walker, and the Sevierville Cohort for your continuous concern and encouragement throughout this endeavor. We will always have a special bond.

Thank you to Rob Britt, Tammy Cash, Judy Coppenger, Susan Elrod, Alvin Hord, Sylvia Kerr, Dr. Jane Morton, David Murrell, and Buffy Wyrosdick. You are all exceptional professionals and valued friends.

TABLE OF CONTENTS

Page

ABSTRACT 2

DEDICATION 5

ACKNOWLEDGMENTS 6

LIST OF TABLES 11

LIST OF FIGURES 12

Chapter

1. INTRODUCTION 13

Statement of the Problem 16

Research Questions 17

Significance of the Study 18

Definitions of Terms 20

Delimitations and Limitations 21

Overview of the Study 21

2. REVIEW OF LITERATURE 23

Foundation for Effective Reading Instruction 23

Educational Reform and High Stakes Testing 29

Basis of Curriculum-Based Measurement 32

Development of Oral Reading Fluency 36

Response to Intervention 38

High Stakes Assessment 43

Construct Validity of Curriculum-Based Measurement of Reading 44

Predictive Validity 45

Chapter Page

3. RESEARCH METHOD 49

Introduction 49

Research Questions and Null Hypotheses 50

Population 52

Instrumentation 53

Data Collection 54

Data Analysis 55

Summery 56

4. ANALYSIS OF DATA 57

Analysis of Research Questions 58

Research Question 1 58

Research Question 2 60

Research Question 3 62

Research Question 4 63

Research Question 5 65

Research Question 6 66

Research Question 7 68

Research Question 8 70

Research Question 9 72

Summary 74

5. SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS FOR PRACTICE AND FUTURE RESEARCH 75

Introduction 75

Summary of Study 75

Conclusions 78

Recommendations for Practice 80

Page

Recommendations for Future Research 82

REFERENCES 84

APPENDICES 92

Appendix A: Letter to Director of Schools 92

Appendix B: Letter from Director of Schools 94

VITA 95

LIST OF TABLES

Table		Page
1.	Demographics	58
2.	The Bivariate and Partial Correlations of Predictors with TCAP Reading Language Arts Score	60
3.	The Bivariate and Partial Correlations of Predictors for Males with TCAP Reading Language Arts Score	61
4.	The Bivariate and Partial Correlations of Predictors for Female with TCAP Reading Language Arts Score	63
5.	The Bivariate and Partial Correlations of Predictors for Economically Disadvantaged with TCAP Reading Language Arts Score	64
6.	The Bivariate and Partial Correlations of Predictors for Non Economically Disadvantaged with TCAP Reading Language Arts Score	66

LIST OF FIGURES

Figure Page

1. Scatterplot Depicting the Relationship Between Standardized Fall R-CBM Scores and TCAP Reading Scores 68

2. Scatterplot Depicting the Relationship Between Standardized Winter R-CBM Scores and TCAP Reading Scores 70

3. Scatterplot Depicting the Relationship Between Standardized Spring R-CBM Scores and TCAP Reading Scores 72

4. Scatterplot Depicting the Relationship Between Standardized Median R-CBM Scores and TCAP Reading Scores 74

CHAPTER 1

INTRODUCTION

Educators, researchers, politicians, and the public have debated the reform of our nation's public education system for several decades. Several reform models have been implemented in public schools to address this concern (Shores & Chester, 2009). Our nation has set the lofty goal in the reauthorization of the Elementary and Secondary Education Act (ESEA) of 2001 that all students will read on grade level by 2014 (No Child Left Behind [NCLB], 2002). In September of 2011 the United State Department of Education with the looming requirement that all students read on grade level by 2014 offered state educational agencies the opportunity to propose alternative reform models. This was intended to be a temporary measure until Congress passed a reauthorization of the ESEA that could address weaknesses of the previous legislation. In exchange for adopting educational reforms states were permitted relief from the three key provisions of the NCLB Act. A significant problem with the continued implementation of NCLB was the number of states, districts, and schools that faced penalties because of the rigid pass or fail nature of performance benchmarks that were set to reach 100% proficiency of all students in the 2014 school year. This led to the realization that almost all schools and districts would fail to meet adequate yearly progress (AYP). This was coupled with the progressively rigid penalties for failing to move a sufficient number of students in each identified subgroup to proficiency in the areas of reading and language arts, mathematics, and graduation rate. State flexibility waivers were designed to address what are viewed as shortcomings of the 2001 reauthorization of the ESEA.

In exchange for relief from NCLB requirements states were required to raise expectations in their academic standards and were encouraged to adopt the rigorous Common Core State Standards (CCSS) published by the National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). If the rigorous CCSS English and language arts standards are to be realized, schools will need to change current practices to achieve different results. A National Research Council study cited by Heller, Holtzman, and Messick (1982) was credited as the source for the response to treatment model. This model evolved into what is commonly known as Response to Intervention (RTI) or Response to Instruction (Gerber, 2005). The National Reading Panel Report (2000) recommended a three-tiered intervention model based on the principals of RTI. The Commission on Excellence in Special Education (2002) published a report that was critical of the current model employed in the United States to identify learning disabilities and recommended the RTI Model as a method by which to identify students as learning disabled. Reschly and Hosp (2004) reported that "significant variability between states also continues to exist in SLD prevalence, conceptual definitions, and classification criteria" (p.200). Based on 2001 – 2002 special education census data, the prevalence of a specific learning disability for student populations varied widely from state to state, with Kentucky reporting a 2.96% rate to a 9.49% rate in Rhode Island (Reschly & Hosp, 2004).

In this study the participating East Tennessee school district used a three-tiered intervention model similar to that used in Tennessee Reading First Schools (Tennessee Department of Education, Tennessee Reading First., 2006). This intervention model followed the recommendation of the National Reading Panel and operated as a general education intervention model that required 90 minutes of high quality research based instruction was provided to all students. The classroom teachers were given universal screening or benchmark

R-CBM data three times per year on each student assigned to their classroom. If a student scored below the pre-established cut score measured in words read correctly (WRC), weekly individual R-CBM progress monitoring probes would be given for 6 weeks and differentiated instruction was provided by the classroom teacher in a small group during reading instruction. If insufficient progress was made during the 6-week period, the student would enter tier two intervention. The student would continue the weekly progress monitoring established in tier one, as well as an additional 30 minutes of reading intervention provided by the classroom teacher or a trained paraprofessional. If the student was not successful with this level of support as measured by R-CBM progress monitoring data, after 12 weeks the student would progress to tier three. In tier three the student would be provide an additional 60 minutes of small group instruction in combination with the 90-minute differentiated core reading curriculum.

Vaughn and Roberts (2007) identified effective leadership as an essential factor in RTI implementation. According to Vaughn and Roberts leaders must be "committed to prevention – oriented practices" and "curriculum leaders who are willing to assure that scientifically based practices are implemented" (p.45). Sindelar, Shearer, Yendol-Hoppey, and Liebert (2006) researched school reform and found that teachers were more likely to embrace school reform when instructional changes targeted that hardest to educate students, and teachers received ongoing support and training. For teachers and administrators to become more data driven the authors cited a need for timely evaluation that was then used to inform instruction. It would also greatly benefit schools to have a meaningful predictor of student achievement as measured by the NCLB mandated high stakes summative assessment (Sindelar et al., 2006).

Statement of the Problem

The purpose of the study was to investigate the relationship between formative reading curriculum based measure (R-CBM) and the Tennessee Comprehensive Assessment Program third grade reading and language arts assessment. If a correlation between the two assessments is found, it would provide teachers and administrators an assessment that requires little time to track students' progress in the critical skills of reading. The RTI process could be more closely tied to student proficiency and the adequate yearly progress (AYP) students and schools must show in annual high stakes assessment. NCLB requires schools and school districts to meet AYP goals for all students as well as all subgroups. This was a concern because some groups have been historically low performing on state-wide assessments. It was important to assess how well the RTI (three tiered intervention model and R-CBM scores) identify the students at risk of not achieving state proficiency standards. Identifying how well R-CBM scores identified students who are at risk of not meeting AYP as measured by TCAP reading and language arts scores was the goal of this study. While similar studies have been conducted in other states, the relationship between Tennessee's TCAP reading and language arts assessment and R-CBM has not yet been established. It was informative to compare the 13 elementary schools in one East Tennessee school district to determine if students' rates of progress are significantly different from the national aggregate of words read correctly (WRC).

Research Questions

This study was guided by the following research questions:

Research Question 1

Is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM scores) and the criterion variable TCAP reading proficiency scores?

Research Question 2:

For males, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM scores) and the criterion variable TCAP reading proficiency scores?

Research Question 3

For females, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM scores) and the criterion variable TCAP reading proficiency scores?

Research Question 4:

For economically disadvantaged students, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM scores) and the criterion variable TCAP reading proficiency scores?

Research Question 5:

For students not economically disadvantaged, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM scores) and the criterion variable TCAP reading proficiency scores?

Research Question 6:

What is the linear equation that predicts TCAP reading scores from fall R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Research Question 7:

What is the linear equation that predicts TCAP reading scores from winter R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Research Question 8:

What is the linear equation that predicts TCAP reading scores from spring R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Research Question 9:

What is the linear equation that predicts TCAP reading scores from median R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Significance of the Study

The No Child Left Behind Act (NCLB) required that students demonstrate proficiency on state academic standards as measured by annual criterion-referenced state assessment in reading (Standerfer, 2006). Tennessee was granted a flexibility wavier by the U.S. Secretary of Education (2012) that granted relief from key provisions of the NCLB act in exchange for creating ambitious goals for students in achievement on state mandated assessments in reading and language arts and mathematics and adopting more rigorous CCSS and corresponding assessments. Many public school teachers feel pressure to increase students' academic growth in

reading and mathematics and the number of students achieving proficiency as measured by students' proficiency on high stake assessment. As the required proficiency level increases each year, it becomes more important to identify students who are not progressing as soon as possible to maximize students' outcomes. This researcher seeks to build on the body of knowledge in reading assessment by providing educators a better understanding of critical information regarding the predictive values of a widely used, commercially available standardized reading assessment (R-CBM). A review of the literature did not yield any widely published studies that examined the predictive value of R-CBM to the T-CAP reading and language arts assessments. The population was significantly robust to offer utility in predicting what students are at-risk for nonproficient scores on the state mandated reading and language arts assessment (Baker et al., 2008; Cummings, Atkins, Allison & Cole, 2008; Foorman et al., 2006; Fuchs, 2004; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Stage & Jacobsen, 2001; Wood, 2006; Zimmerman & Dibenedetto, 2008). This study may provide valuable information regarding the relationship between the R-CBM measure and the TCAP reading and language arts subtest used in the state of Tennessee. The R-CBM will determine what level of reading fluency, measured in words read per minute, a third grade student in the participating school district would need to score to likely score a proficient or advanced scores on the TCAP reading and language arts assessment. This study might shed light on the ability of one standardized measure of oral reading fluency to predict student achievement on the third grade TCAP reading and language arts assessment.

Definitions of Terms

1. Adequate Yearly Progress (AYP) – The NCLB law requires states to use a single accountability system for public schools to determine whether all students, as well as subgroups of students, are making progress toward meeting state academic content standards. The goal is to have all students reaching proficient levels in reading and math by 2014 as measured by performance on state tests. Progress on those standards must be tested yearly in grades 3 through 8 and in one grade in high school. The results are then compared to prior years, and, based on state-determined AYP standards, used to determine if the school has made adequate progress towards the proficiency goal (NCLB, 2002).
2. Curriculum based measurement (CBM) – CBM was developed to incorporate data-based decision-making into instructional planning by providing direct and continuous measurement of student progress toward specific instructional objectives. This measure is then used to determine the effectiveness of the instruction (Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008).
3. Oral Reading Fluency (ORF) –ORF is an assessment in which a student reads a preselected grade level passage, typically for 1 minute, and an evaluator records the number of words read correctly and the number of errors exhibited (Deno, Marston, Shinn, & Tindal, 1983).
4. Reading Curriculum Based Measurement (R-CBM) – The term currently is used synonymously to describe curriculum-based measure of oral reading (CBM-R) and oral reading fluency (ORF) in literature (Reschly, Busch, Betts, Deno, & Long, 2009).

5. Response to Intervention (RTI) – RTI is a student centered assessment models that use problem-solving and research based methods to identify and address learning difficulties in children. Core components of RTI include high-quality classroom instruction, universal screening, continuous progress monitoring, research based interventions, and high fidelity of instinctual intervention (Johnson, Mellard, Fucks, & McKnight, 2006).

Delimitations and Limitations

This research was conducted in a school district located in East Tennessee. The participating school district has a student population of 10,761 and 712 teachers (Tennessee Department of Education, 2011). This largely rural, growing suburban, district included 55.1% of students identified as economically disadvantaged and 13.3% of students identified as students with disabilities. This study was delimited to all third grade students in 13 elementary schools. Several limitations were associated with this study. First, the ability to generalize findings are limited because only third grade students participated in the study and all participants were members of the same East Tennessee school district. Second, the primary researcher was an administrator in the district studied. However the researcher did not participate directly in the gathering of assessment data used in this study. The researcher also actively participated in the development of the district's response to intervention reading plan.

Overview of the Study

This study was organized into five chapters. Chapter 1 is an introduction to the study, statement of the problem, research questions, significance of the study, definition of terms, delimitations, and limitations. Chapter 2 presents a review of the related literature that addresses

foundations for effective reading instruction, educational reform and high stakes testing, basis of curriculum-based measurement, development of oral reading fluency, response to intervention, high stakes assessment, and construct validity of curriculum-based measurement of reading. Chapter 3 addresses the research design and methodology. Chapter 4 presents the results of the data analyses. Chapter 5 is a detailed data analysis summary, conclusion, and recommendations for practice and further research.

CHAPTER 2

REVIEW OF LITERATURE

Foundation for Effective Reading Instruction

In 2000 the National Reading Panel (NRP) issued a report that was intended to provide educators with a formula for evaluating an effective early reading literacy program. This was not the first time the teaching of reading was part of the national political dialog. To understand the significance of the NRP's contribution to the teaching of reading it is important to understand the historical development of both written communication and the dominate theories of reading instruction in the United States.

In the infancy of written language, pictures were the conduit for written thought. This form of primitive, picture driven, communication was logographic. Logographic communication was self-explanatory for simple commutation of concrete words (Stahl, Osborn, & Lehr, 1990). However, as people found a need to communicate more complex thoughts, pictures became difficult to interpret and too prone to misunderstanding. As written symbols were refined, an important transition evolved from symbols representing pictures to symbols representing sounds. This phonological use of symbols has evolved into the syllabic writing systems used by many groups of people. The great advantage of a syllabic system is that it requires individuals to learn far fewer syllables than actual words. According to Rozin and Gleitman (1977), English is comprised of about 5,000 syllables. This is far less than the number of words in the English language; it is still a formative task to learn to decode this written system. Symbols or letters began to represent syllables, initial consonant sound, until the Phoenicians introduced the vowel to the existing 19 constants. According to Stahl et al. (1990) the development of a written alphabet is one of the most important contributions in the history of the world.

The development of the alphabet created the opportunity for people to communicate complex ideas with relatively few symbols and were relatively easy to reproduce. The advantage of a formal alphabet was not without cost. The modern English alphabetic word system is abstract and conceptually complex. This alone makes the mastery of reading and writing a complex cognitive task. The complexity of the English language is that it is not perfectly alphabetic. In the English language the phonemic significance of a letter is sometimes modified by the letter next to it creating the need for cumbersome rules for spelling including words that do not conform to phonic rules. The complexity of the English language has raised controversy over how best to teach reading (Stahl et al., 1990).

In the United States during Colonial times, the common process for teaching students to read was an uncomplicated two-step process: Teach students the code, then have them read available literature most commonly the Bible and political pamphlets (Stahl et al., 1990). The letters and their phonemic significance were taught by "…presentation of key words (for example, G is for glass), practice in reading simple syllables, and exercise in spelling" (p.17). This format for reading instruction continued until the mid 1800s when the drill and practice of letter-sound correspondence was questioned. In a highly political manner the Secretary of the Massachusetts Board of Education, Horace Mann, described the letters in current reading books as,…"skeletons shaped, bloodless, ghostly apparitions, it is no wonder that the children look and feel so death-like, when compelled to face them" (p.18). Mann advocated the use of leveled readers in which he theorized children learn to read whole, meaningful words first. While not well received by the teaching community at the time Mann's, Meaning First Curriculum gradually gained momentum.

By the 1920s Meaning First Curriculum became the primary method for reading instruction in the county. In the 1930s and 1940s reading instruction focused on reading comprehension and developing higher order thinking skills of reading for content. The meaning first concept expected children to recognize words by sight or holistically. When students failed to recognize words, they were taught to use context and pictures to derive meaning. Phonics instruction, if used, was subordinate to leveled reading and use of context clues.

Flesch (1955) authored, "Why Johnny Can't Read" a book written to illustrate the need for a phonics approach to reading instruction in schools. Flesch argued that only some children spontaneously gained the skills required to be fluent readers. Many more required a more natural system that taught the relationship between letters and groups of letters and phonemes they represent. The discussion regarding phonics was very political… "There is a connection between phonics and democracy – a fundamental connection. Equal opportunity for all is one of the inalienable rights, and the word method interferes with that right" (p.130). In a similar manner to Mann's introduction of the meaning first concept, Flesch advocated phonics instruction. Flesch's ideas prompted research into the best approach to basic reading instruction. However, because his position was viewed as extreme by some experts, schools, and publishers, the incorporation of phonics was delayed. Profit motive, intellectual honesty, and allegiance to the United States became part of the public debate over reading methods.

In 1959 Chall (1967) began, a longitudinal study of English reading instructional methods, review, and analysis of previous research in the field of reading. This comprehensive analysis, sponsored by the Carnegie Foundation, reported that phonics instruction was a valuable component of reading instruction for beginning readers. Reading instruction was most

successful when phonics instruction was linked to reading of meaningful text that aided the development of comprehension skills.

Bond and Dykstra (1967), as part of a US Office of Education research project investigating first grade reading instruction, found that all studied programs used some degree of letter recognition and decoding. All researched methods produced students who could both read and comprehend text. Children provided systematic phonics instruction exceeded the basic reading skills of students instructed from basal instruction that used leveled reading passages. Interestingly, programs that provided systematic phonics and connected reading and meaning surpassed the basal only approach in basic reading, reading fluency, and comprehension. The research refined the question from is phonics or whole word instruction to one of timeliness, degree, and duration of instruction. This comprehensive long-term study found that comprehensive programs produced better outcomes for all readers despite varied readiness skills (Bond & Dykstra, 1967). This was particularly important in dispelling the belief that systematic phonics instruction was only useful to more cognitively adept learners (Spach & Spache, 1973). Gersten and Keating (1987) conducted a longitudinal study of student reading outcomes. They found that children who received direct systematic phonics instruction with the opportunity to participate in consequential connected reading in primary elementary grades yielded significantly better outcomes in standardized measures of reading, reduced incidence of dropping out of high school, and increased postsecondary education attendance.

Educational research since the early 1970s has established the impact teachers have on how much students learn (Rupley, Blair, & Nichols, 2009). According to the National Reading Panel (NRP) reading instruction is most effective when teachers use a framework of five essential components of effective reading instruction. The NRP's five components of effective

reading instruction are phonemic awareness, phonics, fluency, vocabulary, and comprehension (National Institute of Child Health and Human Development, 2000). According Villaume and Brabham (2003) teachers who provide explicit balanced instruction while supporting students who struggle to acquire basic reading skill consistently produce students who outperform their peers. Researchers at the Florida Center for Reading Research (2007) studied characteristics of Reading First schools whose rate of reading growth exceeded schools with similar demographical populations. The highest performing schools exhibited the following seven common characteristics: strong leadership, positive belief and teacher dedication, data use and analysis, effective scheduling, professional development, scientifically based intervention programs, and parent involvement (Rupley et al., 2009; Torgesen, Houston, Rissman, & Kosanovich, 2007).

The role of the teacher in the educational process can be a determinate factor in students' acquisition of basic reading skills including fluency and comprehension (Moats, 2009). The ability of a teacher to differentiate instruction for diverse learners may be essential to learners' success. This is particularly true of populations at higher risk of reading difficulty including economically disadvantaged students, English Language Learners, and students with disabilities. Research suggests that multicomponent interventions require teacher knowledge in phonology, phoneme-grapheme correspondence, morphology, semantic organization, syntax, discourse, and pragmatics. According to Moats (2009) teachers must become expert in the structure of language at the sublexical level. This coupled with the skills to communicate information to develop a student's insight into words orthography and their association to each other in connected text. Researchers in reading have suggested that a highly competent teacher of reading may reduce the likelihood of reading problems and reduce the severity of reading

problems in children at risk of reading failure (Blachman et al., 2004; Denton, Foorman, & Mathes, 2003; Foorman et al., 2006; Mathes et al., 2005; Vellutino, Tunmer, Jaccard, & Chen, 2007).

Rupley, Blair, and Nichols (2009) asserted that providing effective instruction, especially to populations at risk of reading problems, was effective in preventing most students' reading difficulties. Reading First was designed to provide our nation's schools a means to organize instruction in schools with a tiered approach. This model used universal benchmarking of critical reading skills to inform appropriate intervention for students who need additional attention. This model differed from previous attempts to identify students at risk of reading difficulty in that it did not seek to identify and remove students from traditional classroom reading instruction. Reading First instead stressed that all teachers use best practices in teaching reading. In Reading First schools every teacher in grades K-3 was provided high quality professional development in the area of reading. This model required students to be evaluated for potential weakness in reading periodically and the results to be used to guide instructional decisions in the school, grade level, general classroom, and, if indicated, more intensive targeted intervention by a specialized reading instructor. The Reading First model is a collaborative model in which all teachers need to have a command of the speech sound system of phoneme structure and the meaning of fluency to the overall task of reading and comprehending printed text (Roehrig, Duggar, Moats, Glover, & Mincey, 2008).

Educational Reform and High Stakes Testing

In 1965 the Elementary and Secondary Education Act (ESEA) became law in the United States. One component of this law was to provide economic support to schools that served economically disadvantaged students. The intent of this federal legislation was to lessen the achievement gap between students of different economic backgrounds. As federal funding increased for public schools nationally, accountability for student progress became a focus of educational policy makers.

In 1960 the National Assessment of Educational Progress (NAEP) was charged with the development of a system to assess student learning in reading and mathematics. Because of the diversity in states' educational expectations, the NAEP standards became the national standard. The NAEP reported scores by region structured in such a way that states and schools could not be compared. Despite increased federal spending for economically disadvantaged students, the achievement gap remained a problem. The National Commission on Excellence in Education's Report titled; A Nation at Risk (1983) concluded that if educational reform was necessary for the US, to remain economically competitive in the global marketplace. After the passage of the 2001 reauthorization of the Elementary and Secondary Education Act (ESEA), renamed No Child Left Behind (NCLB), most schools focused on high stakes state assessment (Ardoin & Christ, 2008).

The United States Congress was expected to reauthorize the ESEA in 2007 to address what are viewed, as shortcomings of the 2001 reauthorization know as NCLB. While waiting for Congress to take up this legislation the United States Department of Education (USDOE) offered states flexibility from some prescriptive provisions in the law. The most significant difficulty of states, districts, and schools was the pass-fail nature of performance benchmarks that had led to

the over identification of schools and districts as failing to meet adequate yearly progress (AYP). This coupled with the progressively rigid penalties for failing to move a sufficient number of students in each identified subgroup to proficiency within a school year in the areas of reading and language arts, mathematics, and graduation rate. The USDOE created a process for states to request relief from many provisions of the NCLB Act. States that chose to submit an application to the USDOE were evaluated as to how well the proposed accountability system met three general principles: 1. College and career ready expectations for all students; 2. state developed differentiated recognition, accountability, and support; and 3. supporting effective instruction and leadership. The State of Tennessee submitted an initial ESEA flexibility request on November 14, 2011, and agreed to meet all principles of ESEA Flexibility (Tennessee Department of Education, 2011, November). After addressing the concerns of the review panel Tennessee's request for ESEA flexibility was approved (Duncan, 2012).

The Tennessee ESEA Flexibility Wavier granted to Tennessee by the USDOE waived portions of the NCLB Act. The most important change was to the accountability system from the "unreasonably high proficiency targets that resulted in the majority of schools labeled as failing" (Tennessee Department of Education, 2012, March p.8). The new accountability system was based on growth for all students and closing the achievement gap. Annual Measurable Objectives (AMOs) were set for each school district in Tennessee for both achievement and gap closure. Districts were classified into three categories: Exemplary (meeting the majority of both achievement and gap closure AMOs), Intermediate (meeting the majority of either achievement or gap closure measure AMO), and Needs Improvement (failing to meet the majority of either achievement or gap closure AMOs). Unlike the NCLB accountably model that examined each schools progress with students in multiple subgroups, the new accountability model identified

some schools that based on student performance and value-added growth on the TCAP Assessment. Some Tennessee schools are identified in to one of three categories; reward schools (top 10% of schools based on absolute performance and value-added growth), focus schools (the 10% of schools with the largest achievement gaps), and priority schools (bottom 5% of schools, based on absolute performance). Students reaching proficiency in reading and language arts continue to be a major part of school and district success. Achievement in the area of reading as measured by the TCAP will be measured in third and seventh grades, as well as a combined measure of students reading proficiency grades 3-8 will be critical in districts meeting AMOs (Tennessee Department of Education, 2012, March).

Final regulations for the reauthorized Individuals with Disabilities Education Act of 2004 (IDEA) were published in the Federal Register on August 14, 2006, and became effective on October 13, 2006. This reauthorization of IDEA adopted language from the NCLB Act to align reform efforts across both statutes. Both laws called for the use of scientifically based research as defined in the NCLB Act and permitted the use of up to 15% of a school systems IDEA funds for early intervening services, which mirror tiered interventions found in both the Reading First Program and RTI. It also allowed states flexibility in the use of ESEA and IDEA funds for school-wide programs, which benefit students at-risk of academic difficulty (United States Department of Education, 2007). The reauthorization of IDEA also made changes to the definition of specific learning disabilities (SLD) that required schools to provide high fidelity interventions to students while monitoring the students progress as a means of validating instructional interventions before a student could be identified as a student with a disability. The Tennessee Department of Education, Division of Special Education (2010) adopted new guidelines for the identification of a SLD. These guidelines mirrored the IDEA as reauthorized

in 2004 while also adding some language from NCLB including the necessity to provide evidence that underachieving children were getting scientifically validated instruction in reading and math (Tennessee Department of Education, 2010). The TDOE required the use of "data-based documentation of repeated formal assessment of student progress during instruction (progress monitoring data) that has been collected and recorded frequently (a minimum of one data point per week in each area of academic concern)" (p. 21). The TDOE advocated for the use of a response to intervention model created though a U.S. Department of Education Office of Special Education Programs State Improvement Grant administered by the IRIS Center for Training Enhancements.

Basis of Curriculum-Based Measurement

The origins of curriculum-based measures (CBM) are traceable to the pioneering work of Deno (1985) while a researcher at the Institute for Research on Learning Disabilities (IRLD). Deno and Mirkin (1977) developed an intervention model for students with disabilities known as Data-Based Program Modification System (DBPMS). DBPMS was founded on the idea of repeated formative assessment (Deno, 2003). DBPMS also applied the principles of criterion-referenced tests and mastery measurement (Fuchs 2004). Mastery measurement was a system of learning objectives that are organized into a hierarchy of discreet skills. Mastery was established when the specific criteria were met. When a specific criterion was mastered, an instructor would begin teaching and assessing the next learning objective in the hierarchy of skills (2004). The use of mastery measurement provided frequent formative assessment based in part on the exiting research of Carroll (1963) that explored the use of the criterion mastery model. This model focused on individual student growth and was sensitive to a student's mastery of specific objectives (Zimmerman & Dibenedetto, 2008). Recognizing that teachers' expectations differed

a practical set of procedures for evaluating a student's progression in reading, mathematics, and writing was needed in order to validate the effectiveness of a teacher's instruction (Deno, Wallace, & Espin, 2005).

Researchers at the Institute for Research on Learning Disabilities (IRLD) identified critical concerns associated with existing methods of monitoring student progress, including mastery measurement. One basic, yet critical, criticism of mastery measurement was that mastery measurement evaluated each skill in the form of a criterion-based test of a specific skill. This isolation of skills indicated progress in isolation that allowed students to show mastery of skills without generalizing skills to more complex tasks that did not fit easily into one discreet skill. In reading this would allow students to demonstrate mastery on like words such as consent-vowel-consent (c-v-c) words, which all fit a specific pattern. However, students could no longer show mastery of c-v-c words when mixed with other letter combinations. Researchers at IRLD questioned the basic premise of mastery measurement that a series of short-term objectives accumulate to broader competence. To overcome the observed problems associated with mastery measurement, an alternative was theorized that would prove to be reliable, valid, and efficient enough to be used to contribute to daily instructional decision-making (Deno, 1985; Deno, 2003; Fuchs, 2004; Zimmerman & Dibenedetto, 2008).

Deno (1985) made the case that none of the current systems of CBA (curriculum-based assessment) were suitable to guide and inform teacher decision-making. The need to establish a new standardized functional formative assessment to guide and inform teacher decision-making led researchers at IRLD to develop the assessment and procedures that has become to be known as CBM (Deno et al., 2005).

Deno (1985) acknowledged the need to develop a more effective system of CBA. His research would result in the development of the CBM. The first of two objectives was to develop an evaluation process and instrument that teachers could frequently make decisions regarding a student's progress and differentiate instruction accordingly. The second goal was to determine if teachers using this new process of student evaluation would be more effective, thus producing higher student achievement. CBM assessments were developed to be used as formative assessments administered weekly, monthly, or periodically throughout a school year. Each CBM assessment included items that are representative of the entire curriculum correcting a major criticism of master measurement and other forms of CBA.

A major goal of the IRLD was to develop a measurement and evaluation procedure that teachers could use to make decision about whether to modify a student's instructional program (Deno, 1985, 2003). The curriculum-based measure was developed to routinely monitor student progress within a curriculum. The following four characteristics were developed to guide the development of this new assessment:

1. Reliability and valid if results of their use were to be accepted as evidence regarding student achievement and the basis for making instructional decisions.
2. Simple and efficient if teachers were going to use them, or teach others to use them, to frequently monitor student achievement.
3. Easily understood so that the result could be clearly and correctly communicated to parents, teachers, and students.
4. Inexpensive since multiple forms were to be required for repeated measurement (Deno, 1985 p. 221).

The mastery of reading is generally accepted as the comprehension of printed text (Fuchs, Fuchs, Hosp, & Jenkins, 2001). However, development of an assessment that measured a student's reading comprehension of text proved difficult. Many instruments were high in content validity but lacked the simplicity, efficiency, and economy required to produce an assessment

that had utility in a school setting. Three methods emerged that assessed reading skills and meet the initial criteria of reliability, simplicity, efficiency, and economy. The cloze procedure, word meaning, and read-aloud tasks became the basis for early standardized reading CBM assessment. According to Graney and Shinn (2005) all CBMs in reading were correlated with performances on more traditional standardized assessments.

Deno et al. (1983) found that reading aloud from text, known commonly as oral reading fluency (ORF), could reliably discriminate between students who received special education services in reading and students who were not in need of special education services in reading. It was expected that ORF would highly correlate with standardized measures of basic letter or word reading. The strong relationship between ORF and basic letter or word reading is well established. It was also found that a relationship existed between ORF and standardized assessments in reading comprehension and global reading measures such as high stakes state assessment in reading (Reschly, Busch, Betts, Deno, & Long, 2009). The most widely researched and used from of CBM in the United States is oral reading or similar read-aloud measures known as R-CBM (Reschly et al., 2009).

The Research Institute of Progress Monitoring (RIPM) at the University of Minnesota explored the utility of ORF as a screening instrument for reading. Elementary students in grades 1 through 6 were asked to read preselected grade level text for 1 minute after which the total number of words read correctly (WRC) were recorded. The results of this national study revealed patterns in reading fluency in each grade level that mirrored existing developmental growth patterns in height and weight. The result of this study solidified the ability of ORF to meet the original goal of the Institute that was to produce a reliable, efficient, easily understood, and inexpensive measure of students' achievement (Deno, 1985, Fuchs, & Deno, 1991). Deno

discovered that the read-aloud tasks when compared to standardized comprehension measures yielded Pearson correlation coefficients from .70 to .95 (p.223).

Development of Oral Reading Fluency

According to Deno (2003) previous research had established that the speed a student is able to translate text into spoken language is a valid measure of reading ability. The ability to orally decode appropriate grade level text under time constraint is the hallmark of curriculum-based measurement of oral reading (CBM-R). According to Shinn and Bamonto (1998) the use of ORF as a measure of reading progress was largely relegated to university researchers during the 1990s. CBM-R methods continued to be validated and refined during the 1990s. However, CBM-R methods were not adopted in many schools. A primary concern of educators was the time that it took to assess students in core content areas such as reading. Yell, Deno, and Marston (1992) found that teachers who were currently using CBM in reading listed time as the most important obstacle to implementing ORF progress monitoring within the classroom. Ironically respondents indicated that the time used to administer CBM was less than 10% of their instruction time. It also was difficult to convince practitioners, teachers and administrators, that passage reading for speed and accuracy could function as a valid indicator of a student's overall reading skill (Deno, 2003; Reschly et al., 2009).

The implementation of CBM-R in public schools grew rapidly after the publication of the NRP Report in 2000. The Reading First Assessment Committee (RFAC) developed criteria to evaluate reading programs in kindergarten through third grade as a required component of the Reading First initiative. Reading First was the reading focused component of the broad education reform legislation entitled No Child Left Behind (2002). RFAC was assigned the task of selecting an assessment that would assist state and local educational agencies in selecting

screening, diagnostic, and classroom based instructional assessments in reading as a required component of receiving Reading First grants. The RFAC recommended that a student's progress in reading be measured at least three times per school year using a CBM (Kame'enui et al., 2006). The versatility of CBM allowed it to produce scores that served four key areas: screening, diagnosing, progress monitoring, and outcome evaluations. The RFAC also found the consistency reliability of fluency to be (.70), which was greater than measures of phonics (.64), comprehension (.57), vocabulary (.52), and phonemic awareness (.41). The use of proxy measures such as 1 minute ORF tests were thought to be useful in monitoring progress in comprehension and vocabulary, allowing educational decision makers such as teachers and administrators to make data informed instructional decisions on the effectiveness of instruction for individual students, classrooms, grade levels, and schools.

The Reading First initiative required tri-annual universal screening of all students in second and third grade using a measure of ORF known as a benchmark assessment. This literacy framework was widely implemented in the Reading First initiative that provided grant funds to states, districts, and individual schools to implement the core principles of the NRP Report. Because of subsequent availability of timely high quality reading data and increased accountability based on state mandated criterion-based assessments, educators began to use R-CBM to predict performance of on state assessments and benchmarks for predicting student proficiency on high stakes state assessment. Research has established that R-CBM has a Pearson correlation coefficient with national assessments of .74 and a significantly lower correlation with state assessments of .65 (Reschly et al., 2009). The mounting research base spanning almost 30 years caused researcher Fuchs (2004), "…to ask how much corroborating evidence is needed

before school leaders and professional organizations feel compelled to require research-based forms of CBM as a specific approach to progress monitoring" (p. 192).

Response to Intervention

Shores and Chester (2009) described Response to Intervention (RTI) as a legitimate, effective school improvement model when taken seriously and implemented completely. The potential to transform classrooms into highly effective, motivating places for all students is real. This RTI process involved identifying students with core academic deficits as early as possible so that effective strategies and specialized instruction can be implemented with fidelity before students fall significantly behind their peers. Shores and Chester stated that the RTI process will help teachers and administrators to focus resources on enabling students to be successful.

The RTI process as defined by the National Research Center on Learning Disabilities (NRCLD) is a student centered assessment model that uses problem solving and research based methods to identify and address learning differences in students (Bender, Berkeley, Peaster, & Saunders, 2009; Fuchs, Johnson, McKnight, & Mellard, 2006). A core component of RTI includes high-quality classroom instruction, universal screening, continuous progress monitoring, research based interventions, and high fidelity of instructional intervention (Fuchs et al., 2006). CBM was initially developed to provide timely instructional feedback to teachers of students with exceptionalities. However, it became an assessment instrument to evaluate the quality of instruction in reading and math instruction for all students. After the reauthorization of the IDEA in 2004, RTI became a method to identify a student as failing to make acceptable progress in the general curriculum and indicate a need for intervention of increased intensity including consideration of the possibility of a specific learning disability.

Despite criticism within the field of learning disabilities that an RTI process is not able to reliably differentiate SLD from other conditions such as intellectual disabilities, emotional disturbance, or attention-deficit hyperactivity disorder, it has become a method that schools could consider in identifying a student as a student with a learning disability. The final regulations for the reauthorized IDEA were published in the Federal Register on August 14, 2006, and became effective on October 13, 2006. The reauthorization allowed local educational agencies to identify students with SLD using an RTI process.

> With regard to identifying children with SLD, the regulations: (1) allow a local educational agency (LEA) to consider a child's response to scientific, research based intervention as part of the SLD determination process; (2) allow States to use other alternative research-based procedures for determining whether a child has a SLD; (3) provided that States may not require the use of a severe discrepancy between intellectual ability and achievement to determine whether a child has a SLD; and (4) requires a public agency to use the State criteria in determining whether a child has a SLD and discuss the role that response to scientific research-based intervention plays in a comprehensive evaluation process. (U.S. Department of Education, 2007, p. 1)

Before the 2004 reauthorization of IDEA, the IQ-achievement discrepancy model had been used exclusively by states educational agencies (SEA) and local educational agencies (LEA) as the process to identify a specific learning disability. The aptitude – achievement discrepancy model for the identification of a SLD was criticized as a wait to fail model, as struggling students would receive no specialized instruction or support until a discrepancy was measured and special education services could begin (Stecker, Fuchs, & Fuchs, 2008).

Because of changes in the IDEA, states adopted policies to allow LEAs to identify students as having an SLD without the use of a discrepancy model (Fuchs & Fuchs, 2006). The Tennessee Department of Education, Division of Special Education (2010) adopted new guidelines for the identification of a SLD. These guidelines mirrored the IDEA as reauthorized in 2004 while also adding some language from NCLB. The TDOE (2010) required the use of

data-based documentation of repeated formal assessment of students' progress during instruction (progress monitoring data) that has been collected and recorded frequently, a minimum of one data point per week in each area of academic concern. The TDOE advocated for the use of a RTI model created though a U.S. Department of Education, Office of Special Education Programs State Improvement Grant administered by the IRIS Center for Training Enhancements. While not requiring a specific process, the intention was to require schools to adopt the use of CBM to meet the new requirement. In Tennessee and many other states the use of an IQ-achievement discrepancy model was still permissible (Berkely et al., 2009). Requirements for the identification of a SLD required an intervention structure that allowed LEAs to comply with the new standards. Many districts used the three-tiered model similar to the one adapted by the Tennessee Department of Education for use in Reading First schools. The use of CBM to provide formal assessment of students became a practice in Tennessee Schools.

In RTI students' progress is periodically monitored in conjunction with high-quality instruction to validate the students benefit from the instruction. RTI was used to identify both academic and behavioral needs; this narrative focused on the area of reading. If students fail to progress in a skill when compared to their peers, it might be necessary to differentiate the instruction in order to increase the rate of skill acquisition allowing the student to progress toward an established benchmark. RTI is often described as a multitiered intervention system, with each progressive tier offering an increased intensity of instructional support (Stecker et al., 2008).

After the reauthorization of the IDEA in 2004, it was permissible for schools to use a student's response to research-based intervention as a method for the identification of a SLD. The RTI framework required high-quality classroom instruction and universal screening for all

students. The requirement of high quality instruction in conjunction with universal screening allowed systemic difficulties with instruction to be identified and addressed. The National Association of State Directors of Special Education (2005) estimated that 80% of students learning needs were met with quality classroom instruction and universal screening. The RTI model addressed several criticisms of the aptitude-achievement discrepancy model for the identification of a specific leading disability (Mesmer & Mesmer, 2008; Vaughn & Fuchs, 2003).

In secondary intervention, often identified as tier two intervention, a student is identified as at-risk of reading difficulty using universal screening data and in need of additional instruction. In the second tier of intervention changes to the intensity of instruction are required for a sufficient period to allow educators to rule out inadequate instruction as the primary cause for the student's insufficient progress. RTI was a model designed to address more intense learning needs before a student falls significantly behind same aged peers.

RTI focused on providing students with increased opportunities to learn specific skills. The framework prescribed in RTI for the identification of reading disabilities was similar to the intervention model in Reading First. In Reading First participating schools were required to provide literacy screenings in all kindergarten through third grade classes. The process in both Reading First and RTI requires the use of assessments, in most cases CBM-R data, in the form of a tri-annual universal screening to identify weaknesses in overall reading instruction and to identify if a student is at risk of not reaching academic goals. Students not meeting established benchmarks are determined to be in need of additional instruction (Mesmer & Mesmer, 2008). In both models intervention was delivered in smaller groups that allowed additional research-based instruction in areas of identified needs. This universal screening data were also useful to

teachers and administrators in comparing progress between schools, grade levels, schools, and teachers.

Students identified as in need of additional differentiated instruction in RTI required more frequent progress monitoring than tri-annual benchmarking. Progress monitoring is often conducted weekly in order to monitor the effectiveness of the intervention. It is important that the assessment have multiple forms to decrease the possible that the test takers' familiarity with the assessment is not inflating the performance score (Mesmer & Mesmer, 2008). It is also critical that the progress monitoring instrument is sensitive to small changes in students' skills that might not be readily identifiable in traditional comprehensive achievement assessments (Riedel, 2007). If a student shows growth in the targeted skill, the student will continue at the increased intensity of instructional intervention until the student is no longer in need of intervention or the student fails to make sufficient progress at the current level of intervention. School-based teams use a data driven decision-making process to determine if additional academic interventions are warranted. If a student fails to make the expected progress in spite of scientifically based intervention, it may be an indication that the student needs more specialized instruction.

High Stakes Assessment

Wanker and Christie (2005) asserted that NCLB is possibly the most comprehensive education reform legislation in the last 40 years. A cornerstone of NCLB was the legislative requirement that all students were required to show proficiency on state academic standards, as measured by criterion-referenced assessment. This required most states to adopt a system of state assessment that met the federal requirements for subjects evaluated as well as shifting assessment from normative referenced assessments to criterion-referenced assessments. It was necessary for each state to develop a structure to measure the progress of each school and local educational agency made toward the goal of all students meeting proficiency standards by the 2013-2014 school year. A school or district was deemed to have made adequate yearly progress (AYP) only if students in all identified subgroups met state benchmarks. NCLB required that each state publish the status of each school toward meeting AYP (Standerfer, 2006). This has caused schools, districts, and states to adopt research based strategies to improve students' proficiency levels as measured by state mandated assessments. It has also made it advantageous to identify students at risk of academic failure as soon as possible to allow sufficient time to remediate deficit areas before high stakes assessments are administered.

The State of Tennessee overhauled the previous normative referenced state assessment to comply with the NCLB requirement that each state use a criterion-referenced achievement test in the areas of Reading Language Art and Mathematics in grades 3 through 8. The Tennessee Comprehensive Assessment Program (TCAP) serves as Tennessee's high-stake assessment for holding school districts, schools, and teachers accountable for achieving adequate yearly progress (APY) under state and NCLB accountability. All Tennessee students in grades 3 through 8 take this assessment to determine proficiency in mathematics, reading and language

arts, science, and social studies. The TCAP has been determined to be reliable and valid as an assessment instrument by the Educational Testing Service (ETS), which established the content validity of this summative measure of student achievement. This accountability measure serves more than 450,000 students and is administrated each spring to meet the state's accountability standards (Educational Testing Service, 2012).

Construct Validity of Curriculum-Based Measurement of Reading

Fuchs, Fuchs, and Compton (2004) reported finding more than 150 articles on the psychometric properties of CBM. Ardoin and Christ (2008) discussed two approaches to the development of CBM-R passages that may be used as universal tri-annual screening instruments of basic reading fluency. In the first approach multiple reading passages were selected based on a readability formula that determines the difficulty of passages. It is common to administer three separate passages during each benchmark, fall, winter, and spring. The median or middle score of WRC would then be used to record the students' performance at each screening. The median WRC scores (fall, winter, and spring) were then used to derive a rate of growth over time in reading fluency. Ardoin et al., (2004) suggested that a single CBM-R probe might be sufficient to derive a reasonably accurate measure of reading fluency growth with passages properly developed to control for variations in difficulty. Relatively little is known about the standard error as it relates to measurement error inherent in the administration of multiple reading fluency probes (Ardoin & Christ, 2008). Ardoin, Suldo, Witt, Aldrich, and McDonald (2005), suggested that the magnitude of error was minimized if passages difficulty was constant over the course of the school year. The second method used to derive an effective universal screening tool is to use the same passage or group of passages for each administration (fall, winter, and spring). This

method eliminates the inherent standard error of using multiple passages developed to be equivalent based on a readability formula (Ardoin & Christ, 2008). This method was used by Shin and Bamonto (1998) in the development of AIMSweb to decrease the magnitude of measurement error between administrations of universal screening assessments. By administering the same passages across universal screenings, AIMSweb is able to measure growth between evaluations (fall, winter, and spring) with less interference from standard error than that of DIBELS (Dynamic Indicators of Basic Literacy Skills). Ardoin and Christ (2009) evaluated the estimated magnitude of standard error for DIBELS, AIMSweb, (both commercially available measures) and an Experimental Passage Set; named Formative Assessment Instrumentation and Procedures for Reading (FAIP-R). The standard error of estimates were approximately 11, 12, and 15 WRCM for FAIP-R, AIMSweb, and DIBELS respectively. In the AIMSweb constructed CBM-R passages an additional step in developing the passages was taken. Each AIMSweb passage, while under development, was administered to a small sample group of students to observe performance. This resulted in less volatility in students' scores from passage to passage over time.

Predictive Validity

Research substantiates the ability of CBM-R to predict students' outcomes on group administered standardized measures of reading, reading comprehension, and standards based criterion measures of reading (Baker et al., 2008; Crawford, Tindal, & Stieber, 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Roehrig et al., 2008; Sibley, Biwer, & Hesch, 2001; Stage & Jacobsen, 2001; Wood, 2006). Deno, Mirkin, and Chiang (1982), conducted the initial research studies to establish a relationship between CBM-R and standardized tests of

reading performance. The standardized instruments used for the comparison were the Stanford Diagnostic Reading Test, the Woodcock Reading Mastery Test, and the reading comprehension subtest from the Peabody Individual Achievement Test. Results suggested a strong correlation that ranged from (.73) and (.91), while most correlation coefficients remained in the .80s (Deno et al., 1982 p.42).

Roehrig, Petscher, Nettles, Hudson, and Torgesen (2008) evaluated the predictive and concurrent validity of the DIBELS Oral Reading Fluency (ORF) for predicting performance on the Florida Comprehensive Assessment Test and the Stanford Achievement Test (SAT-10) reading comprehensive assessment test. This study included a population of 35,207 third grade students enrolled in schools that benefited from the Reading First Grant in the state of Florida. This cohort reflected similar demographics to other Reading First schools including: 51% Male, 36% White, 36% African American, 23% Hispanic, 3% Multiracial, 1.5% Asian, and less than 1% Native American. Students identified as economically disadvantaged based of free or reduced prepared meal data represented 75% of total students. The population was divided into two groups while ensuring that both groups reflected the demographics of the larger population (2008). In this sample 90% of students in the low risk group, 62% of the some risk group, and 3% of the high risk group met reading goals by the end of third grade. The overall correlations between ORF with FCAT-SSS and SAT-10 were high ($r = .66$ to $r = .71$). In the fall assessment correlations were moderately strong when comparing ORF with the FCAT-SSS and SAT-10 ($r = .66$ to $r = .68$) respectively. The ORF assessment administered during the winter, February-March evaluation window, yielded the strongest relationship ($r = .71$) when compared to both the FCAT-SSS and the SAT-10.

Hintze and Silberglitt (2005) found a strong relationship between CBM-R and the Minnesota Comprehensive Assessment (MCA). This study demonstrated the CBM-R ability to predict proficiency levels of the MCA beginning in the first grade. This study included 2,675 students in one Minnesota school system. To meet the requirement that each participant have CBM-R data available for grades 1, 2, and 3 only 1,815 students met this requirement and were included in the study. CBM-R was given in the winter of first grade and then three times per year (fall, winter, and spring) in second and third grades. The researchers used discriminate analysis and logistic regression to maximize the number of true positives (students likely to pass the MCA) and true negatives (students likely to fail to achieve proficiency). The predictive validity between the MCA and the CBM-R was significant for all analyses. The CBM-R was more highly correlated with the state assessment with less time between administrations of the two assessments. The degree of association of the CBM-R and the MCA ranged from r = .49 in the winter of the students first grade CBM-R to r = .69 in the spring of the third grade. It was also evident that predictive validity between administrations of the CBM-R increased in a linear manner from weaker correlation between the winter first grade and the spring of third grade (r = .63); to a strong correlation (r = .94) between the winter and spring third grade CBM-R. The positive predictive power of the winter third grade CBM-R when compared to the MCA was strong (r = .97) indicating that students who were predicted to fail to reach proficiency on the MCA did generally receive failing scores. The authors of this study concluded that CBM-R appears to be an efficient method for predicting performance on high-stakes tests by demonstrating the ability to predict those students who are likely to pass reading portions of such tests as far back as first grade (Hintze & Silberglitt). The researchers stated that using, "…R-CBM to set the cut sores in a successive manner from one benchmarking period to the next

across grades appeared to be a more accurate and efficient method than using high stakes test constantly as the criterion regardless of the grade level" (p.383).

CHAPTER 3

RESEARCH METHOD

Introduction

The purpose of this study was to investigate how well Reading Curriculum Based Measures (R-CBM) benchmarks predicted performance on the Tennessee Comprehensive Assessment Program (T-CAP) measuring reading and language arts in the third grade. Archival data were used from one East Tennessee school district. The data were collected without individual students' names and were not traceable to any student. All third grade students selected for this study participated in all three R-CBM benchmark assessments during the 2010-2011 school year. The R-CBM benchmarks were conducted in each school during a predetermined 1 week assessment window in September (fall), January (winter), and April (spring). All participants had three R-CBM score and a reading and language arts TCAP, during the 2010-2011 school year. Another purpose of this study was to determine how well fall, winter, spring, and median third grade male, female, economically disadvantaged, and not economically disadvantaged student R-CBM scores predicted scaled scores in the proficient range on the TCAP reading and language arts assessment. An additional purpose of this study was to determine the linear equation that predicts TCAP reading and language arts scores from the fall, winter, and spring R-CBM scores and how accurately the equation predicted TCAP reading and language arts scores. The R-CBM benchmarks identified as F (fall), W (winter), S (spring), and M (median) were compared as matched sets with TCAP reading and language arts scale scores.

Chapter 3 describes methodology used in this study. The chapter includes the introduction, research questions and null hypotheses, instrumentation, population, data collection procedures, types of data analyses, and chapter summary.

Research Questions and Null Hypotheses

Research Question 1

Is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho1: There is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

Research Question 2:

For males, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho2: For males, there is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

Research Question 3

For females, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho3: For females, there is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

Research Question 4:

For economically disadvantaged students, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho4: For economically disadvantaged students, there is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

Research Question 5:

For students not economically disadvantaged, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho5: For students not economically disadvantaged, there is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

Research Question 6

What is the linear equation that predicts TCAP reading scores from fall R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Research Question 7:

What is the linear equation that predicts TCAP reading scores from winter R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Research Question 8:

What is the linear equation that predicts TCAP reading scores from spring R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Research Question 9:

What is the linear equation that predicts TCAP reading scores from median R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

Population

The population selected to participate in this study included all third grade students in one East Tennessee school district. The third graders were students at 13 schools. Each school was organized into a kindergarten through fifth grade level configuration. According to the Tennessee Report Card (2011), 11.4% percent of the school district's students received special education services, 1.4% percent of the students were English Language Learners, and 53.6% percent were economically disadvantaged based on free and reduced lunch eligibility. The participating district had a daily membership of 10,761. All schools participating in this study met the NCLB criteria to be in Good Standing based on meeting State Annual Yearly Progress expectations for all measured populations. The student population among the kindergarten through fifth grade elementary schools ranged from 134 to 732. Nine schools were identified as title one schools with between 57.7% and 77.6% of the students qualifying for free or reduced

priced meals. The four schools not identified as title one schools recorded a free and reduced priced meal rate of 38.4% to 47.7%. The population studied included 911 third grade students who were required to participate the TCAP assessment required under NCLB. Students who did not participate in the standard reading and language arts assessments were eliminated from this study. This included 11 students who participated in alternative assessments and 66 students who participated in the Modified Academic Achievement Standards (MAAS) for students with disabilities. Eight students were excluded because they participated in the English Linguistically Simplified Assessment (ELSA) in which assessment questions were the same as the TCAP achievement but featured simplified language to reduce the linguistic demand on English Language Learners (ELL). One student was absent during the administration of the TCAP reading and language arts assessment. Fifty-five students were also rejected from this study because they did not receive WRC scores in the fall, winter, and spring R-CBM benchmarks. Seven hundred seventy third grade students were included in this study.

Instrumentation

This study used archrival data generated from the standardized administration of three R-CBM measures in September (fall), December (winter), and April (spring) as the predictor, or independent variables. The R-CBM used in this study was a commercially available program AIMSweb. Each AIMSweb passage has between 250-300 words that are printed in the same font style without pictures. Each student included in this study used three standardized 1 minute AIMSweb reading assessment passages. The words read correctly (WRC) and errors are counted and tallied. The median or middle WRC score for each administration of the three R-CBM passages was used in this study. All examiners completed standardized training on the

administration of the AIMSweb benchmark. This training required each examiner to achieve a inter-scorer agreement of 98% (Shinn & Shinn, 2002).

The dependent variable for this study used the TCAP reading and language arts assessment, which is the criterion-referenced achievement test in the areas of Reading Language Arts in grades 3 – 8. The Tennessee Comprehensive Assessment Program (TCAP) serves as Tennessee's high-stake assessment for holding school districts, schools, and teachers accountable for achieving adequate yearly progress (APY) under state and NCLB accountability. This assessment is a group administered reading achievement assessment that is used to determine proficiency in reading and language arts. All examiners in the participating school district received state mandated training on the standardized administration of the achievement test, as well as proper test security. The TCAP has been determined to be reliable and valid as an assessment instrument by the Educational Testing Service (ETS), which established the content validity of this summative measure of student achievement (Educational Testing Service, 2012).

Data Collection

Prior to conducting this study permission was requested and received from the Institutional Review Board (IRB) at East Tennessee State University (Appendix A). Permission to collect data was initiated and received in a letter to the participating school system (Appendix B), Director of Schools. R-CBM scores expressed in words read correctly (WRC) and TCAP reading and language arts scale scores, with connected gender and free and reduced price meal lunch eligibility information, was requested for all third graders from the 2010-2011 school year. Data were requested and received with student identifiable information redacted and a unique numeric number assigned to each data set. Data were also extracted from the public web site of

the Tennessee Department of Education 2011 *Report Card* for the district and 13 elementary schools in the district (Tennessee Department of Education, 2011). Student or schools were not identified at any time in this study.

Data Analysis

Data were collected for all third graders in an East Tennessee School District. Data were received with a unique to this study number assigned to each data set. The historical data were then organized and entered into a data file. A quantitative nonexperimental comparative research design was used in this study. A multiple linear regression analysis for one set of predictors was conducted to evaluate how well R-CBM scores predicted a specific scale score on the TCAP reading and language arts assessment (Research Question 1). Data were analyzed using the IBM-SPSS Statistical Software Package. The statistical results were used to reject or retain the null hypotheses at the .05 level of significance.

The predictors variables were the four R-CBMs (fall, winter, spring, median), while the criterion variables was the TCAP reading and language arts scale score. The significance of the linear combination of R-CBM will be reported as related to TCAP reading and language arts score. Multiple linear regression for one set of predictors was also used to analyze the predictive value of R-CBM for students of different genders and socioeconomic status, as measured though eligibility for free or reduced priced meals (Research Questions 2-5). This study explored the relationship between two variables to develop a linear equation that predicts the scale score on the TCAP reading and language arts assessment for the R-CBM three benchmark scores (Research Question 6-9). It was also determined how accurately this equation predicts the TACP reading and language arts assessment scale score. This study was organized around nine

research questions. Research question 1-5 were analyzed using a multiple regression analysis. Research questions 6-9 were analyzed using a linear regression analysis.

Summery

Historical third grader data were collected from an East Tennessee public school district to determine how well R-CBM scores (fall, winter, spring, and median) could accurately predict students' scores on the T-CAP reading and language arts assessment. The study also explored the ability of R-CBM scores to accurately predict T-CAP English language arts scores for students of different gender and socioeconomic status. An equation was to predict TCAP reading and language arts scores from R-CBM scores in the fall, winter, and spring administrations of the assessment. Chapter 3 presented the research questions and null hypotheses, population, research design, data collection and collection procedures, data analysis, and a summary. In chapter 4 the findings are reported and discussed. In chapter 5 the summary, conclusions, and recommendations for future research and implications for practice are discussed.

CHAPTER 4

ANALYSIS OF DATA

The purpose of the study was to investigate the relationship between a formative reading curriculum based measure (R-CBM) and the Tennessee Comprehensive Assessment Program (TCAP) third grade reading language arts assessment. The population consisted of all third grade students in an East Tennessee school district during the 2010-2011 school year. The data analyzed in this study were archival data of 911 third grade students who participated in the TCAP reading assessment and three R-CBM assessments. Some third grade students were eliminated from the study because they did not complete all assessments.

Research questions and hypotheses were used to guide the study. In five research questions, five corresponding hypotheses were tested. A multiple regression analysis was conducted to evaluate how well R-CBMs predicted TCAP reading scores. Data were analyzed to determine the predictive relationship between fall, winter, spring, and median R-CBM scores and the TCAP reading assessment for the student population based on gender and socioeconomic status. For four research questions, the linear equation that predicts TCAP reading scores from (fall, winter, spring, and median) R-CBM scores was determined as well as how well the equation predicts TCAP reading scale scores. All calculations were performed using the statistical software package IBM-SPSS. All research questions, hypotheses, and data analyses are presented in Chapter 4.

Archrival data were received with student identifiable information redacted and unique number assigned to each data set. Each data set included three R-CBM scores expressed in words read correctly (WRC) and TCAP reading language arts scale scores, with connected

gender and free and reduced price meal lunch eligibility information for all third graders from the 2010-2011 school year. Table 1 displays the distribution of third grade students participating in the study by gender and economic status as determined by students qualifying for free or reduced priced meals.

Table 1

Demographics

Economic Status	Male	Female	Total	%
Economically disadvantaged	169	196	365	47.4
Not economically disadvantaged	203	202	405	52.6
Totals	372	398	770	100.0

Analysis of Research Questions

Research Question 1

Research Question 1: Is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho1: There is not a significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

A multiple regression analysis was conducted to evaluate how well the (fall, winter, spring, and median R-CBM score) predicted TCAP reading proficiency scores. The predictors were the four R-CBM scores, while the criterion variable was the TCAP reading proficiency score. The linear combination of R-CBM scores was significantly related to the TCAP reading score, $F(4, 762) = 212.15$, $p < .001$; therefore, Ho1 was rejected. The sample multiple correlation coefficient was .73, indicating that approximately 53% of the variance of the TCAP reading score in the population can be accounted for by the linear combination of R-CBM scores.

In Table 2 the indices indicate the relative strength of the individual predictors. All the bivariate correlations between the R-CBM scores and the TCAP reading score were positive and two of the four indices were statistically significant. The partial correlations between fall R-CBM and Spring R-CBM reading scores and TCAP readings scores were significant. On the basis of these correlational analyses, fall is about as predictive as spring but winter is the most highly correlated with TCAP reading scores. However, judgments about the relative importance of these predictors are difficult because they are strongly correlated. The correlations among the R-CBM scores ranged from .91 to .99.

Table 2

The Bivariate and Partial Correlations of Predictors with TCAP Reading Language Arts Score

R-CBM	Beta	*t*	*p*	Zero order correlations	Partial correlations
Fall R-CBM	0.23	2.88	.004*	0.70	0.10
Winter R-CBM	0.26	0.88	.380	0.71	0.03
Spring R-CBM	0.25	3.10	.002*	0.71	0.11
Median R-CBM	0.01	0.02	.988	0.72	0.01

*significant

Research Question 2

Research Question 2: For males, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho2: For males, there is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

A multiple regression analysis was conducted to evaluate how well the (fall, winter, spring, and median R-CBM score) predicted TCAP reading proficiency scores for male third grade students. The predictors were the four R-CBM scores, while the criterion variable was the TCAP reading proficiency score. The linear combination of R-CBM scores was significantly related to the TCAP reading score, $F(4, 365) = 108.14$, $p < .001$; therefore, Ho2 was rejected.

The population multiple correlation coefficient was .74, indicating that approximately 54% of the variance of the TCAP reading score in the population of male students can be accounted for by the linear combination of R-CBM scores.

In Table 3 the indices indicate the relative strength of the individual predictors. All the bivariate correlations between the male students' R-CBM scores and the male students' TCAP reading and language art score were positive and two of the four indices were statistically significant. The partial correlations between fall R-CBM scores and spring R-CBM scores and TCAP reading and language art scores were significant. On the basis of these correlational analyses, fall is about as predictive as winter and spring and but median is the most highly correlated with TCAP reading score. However, judgments about the relative importance of these predictors are difficult because they are strongly correlated. The correlations among the R-CBM scores ranged from .89 to .99.

Table 3

The Bivariate and Partial Correlations of Predictors for Males with TCAP Reading Language Arts Score

R-CBM	Beta	*t*	*p*	Zero order correlations	Partial correlations
Fall R-CBM	0.23	2.23	.027*	0.71	0.12
Winter R-CBM	0.11	0.30	.761	0.71	0.02
Spring R-CBM	0.29	2.90	.004*	0.71	0.15
Median R-CBM	0.12	0.27	.786	0.72	0.01

**significant*

Question 3

Research Question 3: For females, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho3: For females, there is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

A multiple regression analysis was conducted to evaluate how well the female students' (fall, winter, spring, and median R-CBM score) predicted TCAP reading proficiency scores for female third grade students. The predictors were the four R-CBM scores, while the criterion variable was the TCAP reading proficiency score. The linear combination of R-CBM scores was significantly related to the TCAP reading score, $F(4, 392) = 99.62$, $p < .001$; therefore, Ho3 was rejected. The population multiple correlation coefficient was .71, indicating that approximately 50% of the variance of the TCAP reading score in the population of female students can be accounted for by the linear combination of R-CBM scores.

In Table 4 the indices indicate the relative strength of the individual predictors. All the bivariate correlations between the female students' R-CBM scores and the female students' TCAP reading score were positive and one of the four indices were statistically significant. The partial correlations between fall R-CBM and spring R-CBM scores and TCAP reading scores were significant. On the basis of these correlational analyses, fall is about as predictive as spring but winter and median are the most highly correlated with TCAP reading score. However, judgments about the relative importance of these predictors are difficult because they are strongly correlated. The correlations among the R-CBM scores ranged from ..92 to .99.

Table 4

The Bivariate and Partial Correlations of Predictors for Females with TCAP Reading Language Arts Score

R-CBM	Beta	*t*	*p*	Zero order correlations	Partial correlations
Fall R-CBM	0.23	2.23	.047*	0.69	0.10
Winter R-CBM	0.46	0.30	.336	0.70	0.05
Spring R-CBM	0.15	2.90	.223	0.69	0.06
Median R-CBM	-0.12	0.27	.830	0.71	-0.01

*significant

Question 4

Research Question 4: For economically disadvantaged students, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho4: For economically disadvantaged students, there no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

A multiple regression analysis was conducted to evaluate how well the (fall, winter, spring, and median R-CBM score) predicted TCAP reading proficiency scores for economically disadvantaged third grade students. The predictors were the four R-CBM scores, while the criterion variable was the TCAP reading proficiency score. The linear combination of R-CBM

scores was significantly related to the TCAP reading score, $F(4, 358) = 98.17$, $p < .001$; therefore, Ho4 was rejected. The population multiple correlation coefficient was .72, indicating that approximately 52% of the variance of the TCAP reading score in the population of economically disadvantaged students can be accounted for by the linear combination of R-CBM scores.

In Table 5 the indices indicate the relative strength of the individual predictors. All the bivariate correlations between the economically disadvantaged students' R-CBM scores and the economically disadvantaged students' TCAP reading score were positive and no indices were statistically significant. On the basis of these correlational analyses, winter is about as predictive as spring, but fall is slightly less predictive, and median is the most highly correlated with TCAP reading score. However, judgments about the relative importance of these predictors are difficult because they are strongly correlated. The correlations among the R-CBM scores ranged from .90 to .99.

Table 5

The Bivariate and Partial Correlations of Predictors for Economically Disadvantaged with TCAP Reading Language Arts Score

R-CBM	Beta	*t*	*p*	Zero order correlations	Partial correlations
Fall R-CBM	0.13	1.35	.179	0.68	0.07
Winter R-CBM	-0.17	-0.45	.650	0.69	-0.02
Spring R-CBM	0.16	1.68	.093	0.69	0.08
Median R-CBM	0.50	1.13	.259	0.70	0.06

Question 5

Research Question 5: For students not economically disadvantaged, is there a significant relationship between a linear combination of the predictor variables (fall, winter, spring, and median R-CBM score) and the criterion variable TCAP reading proficiency scores?

Ho5: For students not economically disadvantaged there is no significant relationship between a linear combination of the predator variables (fall, winter, spring, and median) and the criterion variable TCAP reading proficiency scores.

A multiple regression analysis was conducted to evaluate how well the (fall, winter, spring, and median R-CBM score) predicted TCAP reading proficiency scores for noneconomically disadvantaged third grade students. The predictors were the four R-CBM scores, while the criterion variable was the TCAP reading proficiency score. The linear combination of R-CBM scores was significantly related to the TCAP reading score, $F(4, 358) = 98.17$, $p < .001$; therefore, Ho5 was rejected. The population multiple correlation coefficient was .70, indicating that approximately 50% of the variance of the TCAP reading score in the population of noneconomically disadvantaged students can be accounted for by the linear combination of R-CBM scores.

In Table 6 the indices indicate the relative strength of the individual predictors. All the bivariate correlations between the noneconomically disadvantaged students' R-CBM scores and the noneconomically disadvantaged students' TCAP reading score were positive and no indices were statistically significant. The partial correlations between fall R-CBM and Spring R-CBM reading scores and TCAP readings scores were significant. On the basis of these correlational analyses, fall is about as predictive as spring but winter is the most highly correlated with TCAP reading scores. However, judgments about the relative importance of these predictors are

difficult because they are strongly correlated. The correlations among the R-CBM scores ranged from .91 to .99.

Table 6

The Bivariate and Partial Correlations of Predictors for Non Economically Disadvantaged with TCAP Reading Language Arts Score

R-CBM	Beta	*t*	*p*	Zero order correlations	Partial correlations
Fall R-CBM	0.13	1.35	.179	0.68	0.07
Winter R-CBM	-0.17	-0.45	.650	0.69	-0.02
Spring R-CBM	0.16	1.68	.093	0.69	0.08
Median R-CBM	0.50	1.13	.259	0.70	0.06

Research Question 6

Research Question 6: What is the linear equation that predicts TCAP reading scores from fall R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

A linear regression analysis was conducted to evaluate the prediction of TCAP reading scores from the R-CBM administered in the fall to all third grade students. The scatterplot for the two variables, as shown in Figure 1, indicates that the two variables are linearly related such that as R-CBM scores increase the TCAP reading score increases. The regression equation for predicting the overall TCAP based on fall R-CBM scores is

Predicted TCAP Reading Score = 0.681 *Fall R-CBM Score* + 692.937

The 95% confidence interval for the slope, 0.63 to 0.73, does not contain the value of zero; therefore, fall R-CBM reading scores are significantly related to the TCAP reading score. Accuracy in predicting the TCAP reading score was strong 0.71. Approximately 50% of the variation of fall R-CBM and the TCAP reading score was accounted for by its linear relationship with the TCAP reading score.

Using the above regression equation for predicting the overall TCAP reading score it was determined that a fall R-CBM score of 99 WRC would predict a TCAP reading score of 760.36. A TCAP reading and language arts score of 760 is necessary for a student to score proficient. For example, a score of 153 WRC was obtained on the fall R-CBM would predict a TCAP reading score of 797.13. This score would place a third grade student in the advanced category on the TCAP. A fall R-CBM Score of 24 WRC would predict a nonproficient TCAP reading score of 709.281. This score exceeds the basic TCAP reading and language arts cut score of 709.

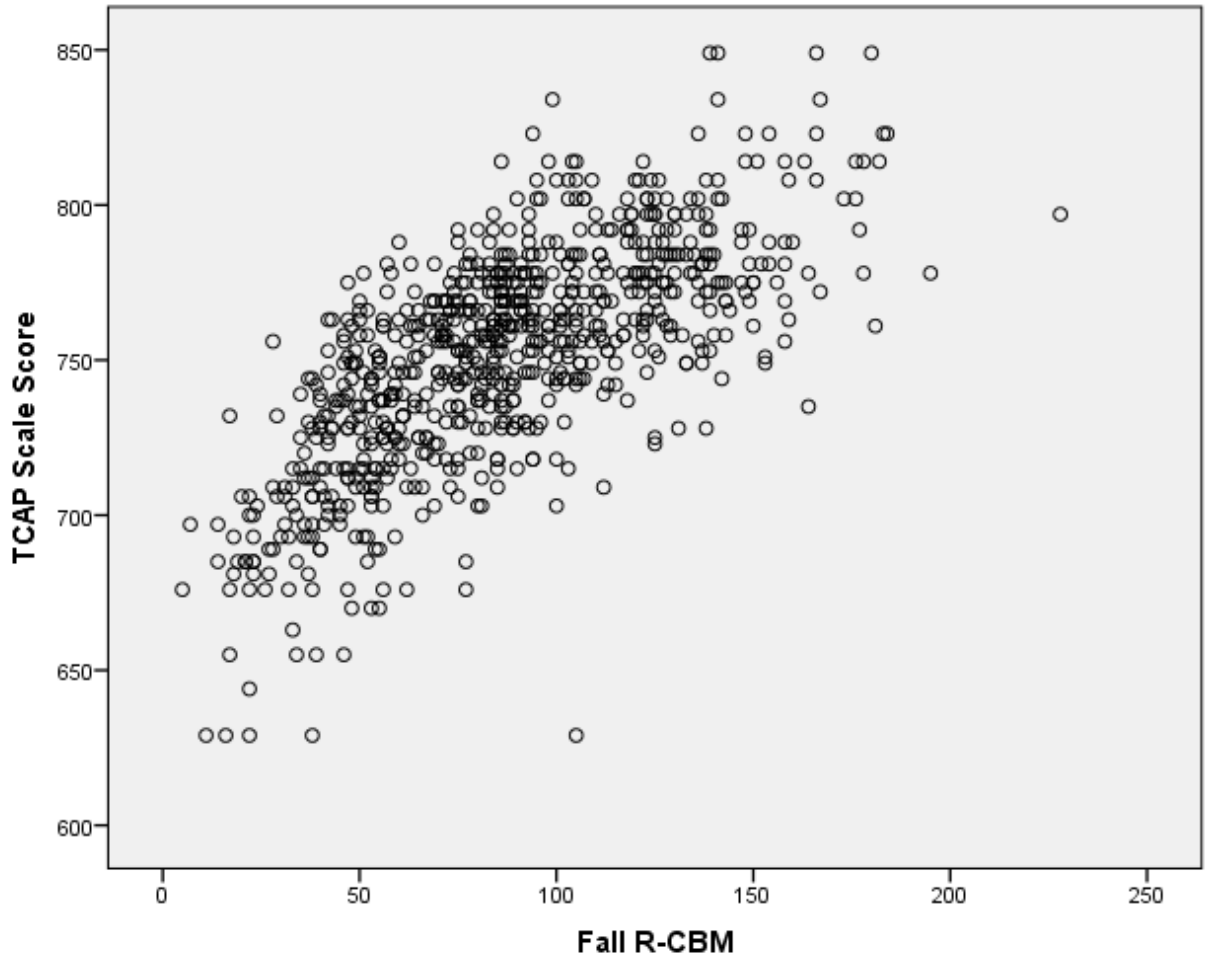

Figure 1. Scatterplot Depicting the Relationship Between Standardized Fall R-CBM Scores and TCAP Reading Scores.

Research Question 7

Research Question 7: What is the linear equation that predicts TCAP reading scores from winter R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

A linear regression analysis was conducted to evaluate the prediction of TCAP reading scores from the R-CBM administered in the winter to all third grade students. The scatterplot for the two variables, as shown in Figure 2, indicates that the two variables are linearly related such

that as R-CBM scores increase the TCAP reading score increases. The regression equation for predicting the overall TCAP based on winter R-CBM scores is

Predicted TCAP Reading Score = .683 *Winter R-CBM Score* + 681.284

The 95% confidence interval for the slope, 0.64 to 0.73 does not contain the value of zero, therefore overall strength is significantly related to the TCAP reading score. Accuracy in predicting the TCAP reading score was strong 0.71. Approximately 51% of the variation of winter R-CBM and the TCAP reading score was accounted for by its linear relationship with the TCAP reading score.

Using the above regression equation for predicting the overall TCAP reading score it was determined that a winter R-CBM score of 116 WRC would predict a TCAP reading score of 760.51. A TCAP reading and language arts score of 760 is necessary for a student to score proficient. If a score of 170 WRC was obtained on the winter R-CBM a TCAP reading score of 797.39 would be predicted. This score would place a third grade student in the advanced category on the TCAP. For example, a winter R-CBM Score of 41 WRC would predict a nonproficient TCAP reading score of 709.29. This score exceeds the basic TCAP reading and language arts cut score of 709.

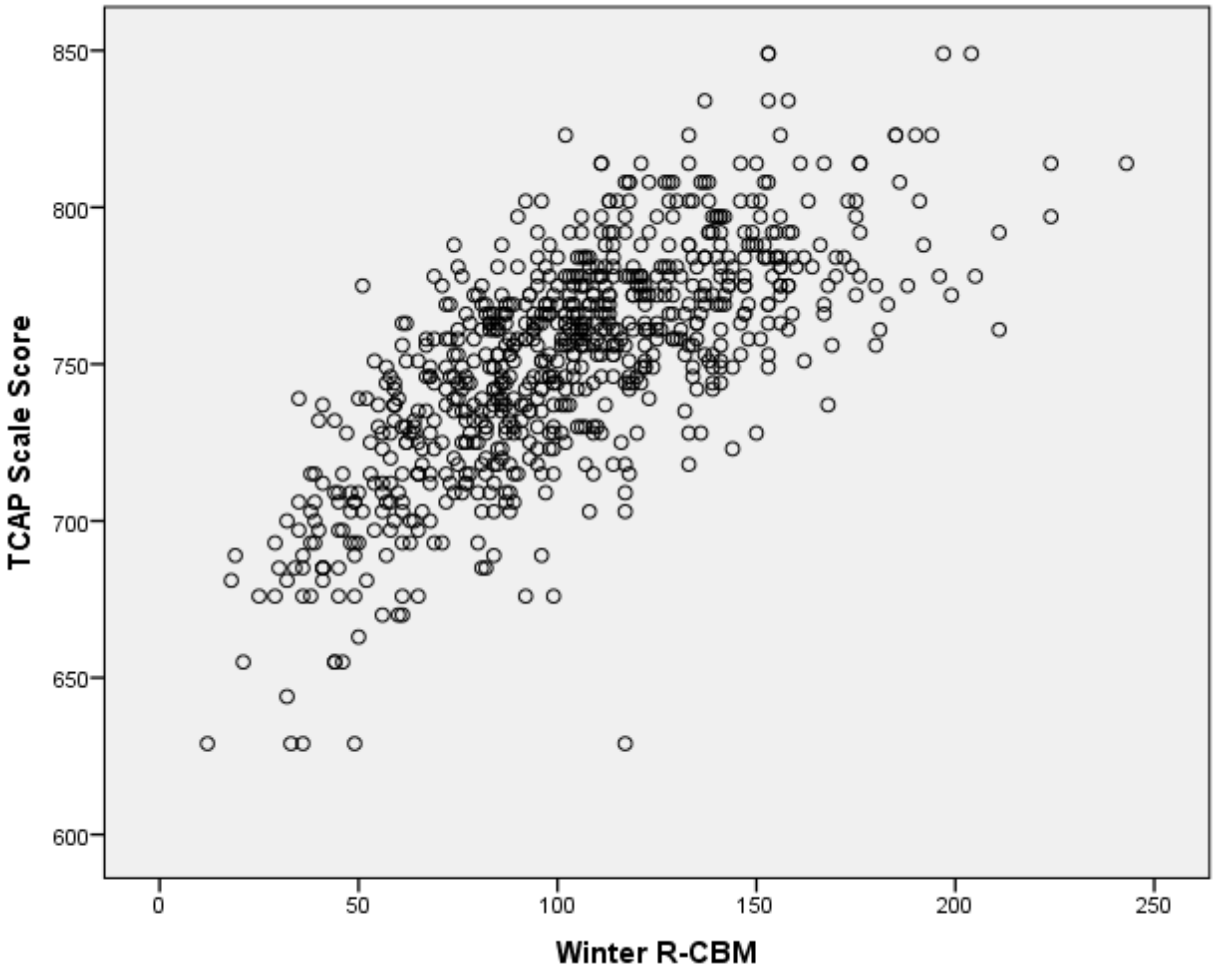

Figure 2. Scatterplot Depicting the Relationship Between Standardized Winter R-CBM Scores and TCAP Reading Scores.

Research Question 8

Research Question 8: What is the linear equation that predicts TCAP reading scores from spring R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

A linear regression analysis was conducted to evaluate the prediction of TCAP reading scores from the R-CBM administered in the spring to all third grade students. The scatterplot for the two variables, as shown in Figure 3, indicates that the two variables are linearly related such

that as R-CBM scores increase the TCAP reading score increases. The regression equation for predicting the overall TCAP based on spring R-CBM scores is

Predicted TCAP Reading Score = .642 Spring *R-CBM Score* + 675.012

The 95% confidence interval for the slope, 0.60 to 0.69 does not contain the value of zero, therefore overall strength is significantly related to the TCAP reading score. Accuracy in predicting the TCAP reading score was strong 0.71. Approximately 50% of the variation of spring R-CBM and the TCAP reading score was accounted for by its linear relationship with the TCAP reading score.

Using the above regression equation for predicting the overall TCAP reading and language arts score it was determined that a spring R-CBM score of 133 WRC would predict a TCAP reading and language arts score of 760.40. A TCAP reading and language arts score of 760 is necessary for a student to score proficient. For example, a score of 191 WRC was obtained on the spring R-CBM would predict a TCAP reading score of 797.63. This score would place a third grade student in the advanced category on the TCAP. A spring R-CBM Score of 53 WRC would predict a nonproficient TCAP reading score of 709.04. This score exceeds the basic TCAP reading and language arts cut score of 709.

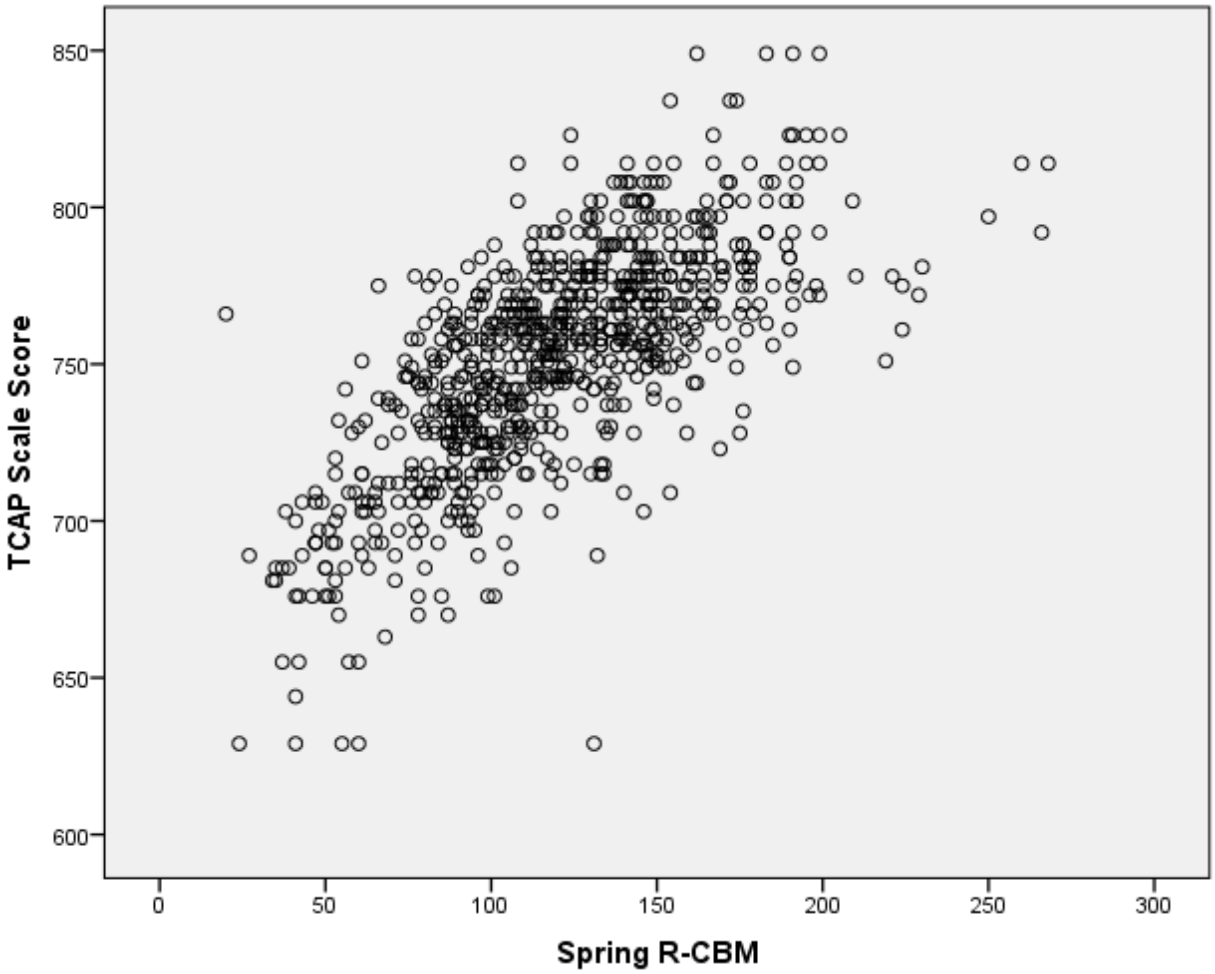

Figure 3. Scatterplot Depicting the Relationship Between Standardized Spring R-CBM Scores and TCAP Reading Scores.

Research Question 9

Research Question 9: What is the linear equation that predicts TCAP reading scores from median R-CBM scores for third graders, and how well does this equation predict TCAP reading scores?

A linear regression analysis was conducted to evaluate the prediction of TCAP reading scores from median R-CBM or middle score of the fall, winter, and spring R-CBM administered to all third grade students. The scatterplot for the two variables, as shown in Figure 4, indicates

that the two variables are linearly related such that as R-CBM scores increase the TCAP reading score increases. The regression equation for predicting the overall TCAP based on median R-CBM scores is

Predicted TCAP Reading Score = .681 *Median R-CBM Score* + 681.553

The 95% confidence interval for the slope, 0.63 to 0.73 does not contain the value of zero, therefore overall strength is significantly related to the TCAP reading score. Accuracy in predicting the TCAP reading score was strong 0.72. Approximately 52% of the variation of median R-CBM and the TCAP reading score was accounted for by its linear relationship with the TCAP reading score.

Using the above regression equation for predicting the overall TCAP reading score it was determined that a median R-CBM score of 116 WRC would predict a TCAP reading and language arts score of 760.55. A TCAP reading and language arts score of 760 is necessary for a student to score proficient. For example, a score of 170 WRC was obtained on the median R-CBM would predict a TCAP reading score of 797.32. This score would place a third grade student in the advanced category on the TCAP reading and language arts assessment. A median R-CBM Score of 41 WRC would predict a nonproficient TCAP reading score of 709.47. This score exceeds the basic TCAP reading and language arts cut score of 709.

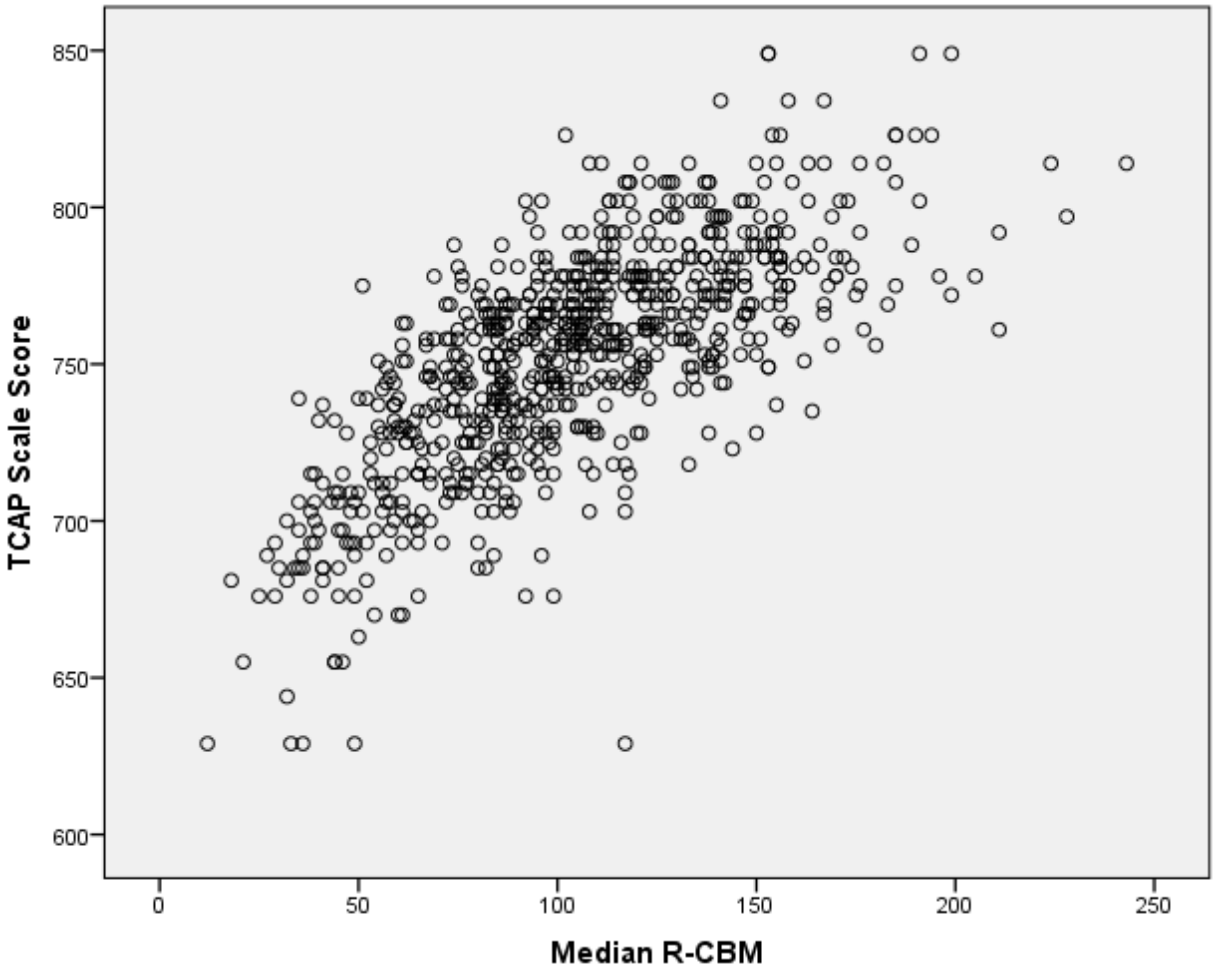

Figure 3. Scatterplot Depicting the Relationship Between Standardized Median R-CBM Scores and TCAP Reading Scores.

Summary

In Chapter 4 each of the first five research questions had hypotheses related to the significant relationship between a linear combination of R-CBM scores and TCAP reading scores. Four research questions created a linear equation that predicted TCAP reading scores from R-CBM scores in the fall, winter, spring, and median. Data were analyzed with SPSS based on R-CBMs administered to all third grade students in the fall, winter, and spring that were compared TCAP reading scores yielding a strong linear relationship.

CHAPTER 5

SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS FOR PRACTICE AND FUTURE RESEARCH

Introduction

The purpose of the study was to determine how well Reading Curriculum Based Measures (R-CBM) benchmarks predicted student performance on the third grade Tennessee Comprehensive Assessment Programs (TCAP) reading language arts subtest. Historical data were analyzed for third grade students from 13 schools in one East Tennessee school district from the 2010-2011 school year. A multiple regression analysis was conducted to determine the predictive relationship between fall, winter, spring, and median or middle score of the fall, winter, and spring R-CBM scores and the TCAP reading and language arts assessment for student population, based on gender and economically disadvantaged based on free and reduced lunch eligibility. Next, a linear equation to predict TCAP reading language arts scores based on fall, winter, spring, and median R-CBM scores was developed. The statistical analysis was conducted using the IBM-SPSS software package.

Summary of the Study

The findings of the study suggest that there is a significant relationship between a linear combination of predictor variables (fall, winter, spring, and median R-CBM scores and the criterion variable TCAP reading and language scores. R-CBM and TCAP reading and language arts scores are likely correlated because they share a common cause such as basic reading fluency skills. Judgments about the relative importance of each predictor are difficult because

they are correlated. The population multiple correlation coefficient ranged from .70 to .74. A multiple regression analysis was conducted with all four R-CBM scores as predictors. The linear combination of the four predictors was significantly related to the TCAP reading and language arts scores $F(4, 762) = 212.15$, $p = <.001$, $R^2 = .53$. Accordingly the R-CBM score is positively correlated across assessments with the TCAP reading and language arts scores and the partial correlation controlling for R-CBM is equal to zero. This correlation coefficient establishes the predictive relationship between the known variable R-CBM (fall, winter, spring, and median) and was useful in predicting an unknown variable, the TCAP reading and language arts score. This finding suggested that fall R-CBM scores have a similarly strong predictive of TCAP reading and language arts scores as winter, spring, and median R-CBM scores.

A multiple regression analysis of four predictors indicated a strong relationship ($R^2 = .53$) between male students' R-CBM score predictors and the TCAP reading and language arts score. The analysis indicated that a student's R-CBM score was strongly related to their actual score on the TCAP reading and language arts assessment. The results establish that R-CBM scores (fall, winter, spring, and median) may be a valid predictor of student scores on the TCAP reading and language arts assessment for this population of third grade students.

When male third grade student R-CBM data were analyzed the bivariate correlational relationship remained positive ranging from .71 to .72. This result was similar to the overall analysis for the total population when compared. Male students' fall R-CBM scores was about as predictive as winter and spring; however, median was slightly higher correlated variable when compared to the TCAP reading and language arts assessment scores. The analysis of female students R-CBM scores also yielded positive bivariate correlations that ranged .69 to .71 between female R-CBM scores and TCAP reading and language arts scores. For female

students, fall was about as predictive as spring but winter and median were slightly most predictive of TCAP reading and language arts scores. R-CBM scores for male students when compared to female scores was a little better predictor of TCAP reading and language arts but not significantly better predictor than female R-CBM scores. A second multiple regression analysis was conducted of all four R-CBM predictors for female students found a strong predictive relationship ($R^2 = .50$) when compared to TCAP reading and language arts scores. While male R-CBM scores were a slightly stronger predictor of TCAP reading and language arts scores, it was not significantly higher than that of female students.

When economically disadvantaged student R-CBM data were analyzed, the bivariate correlational relationship remained positive ranging from .70 to .71 similar to all students analysis. Similar to the total population economically disadvantaged students fall R-CBM scores were as predictive as spring; however, median was slightly higher correlated variable when compared to the TCAP reading and language arts assessment scores. The analysis of noneconomically disadvantaged students R-CBM scores also yielded positive bivariate correlations that ranged .68 to .70 between R-CBM scores and TCAP reading and language arts scores. For noneconomically disadvantaged students, fall was about as predictive as spring but winter and median were slightly more correlated with TCAP reading and language arts scores. R-CBM scores for economically disadvantaged students when compared to noneconomically disadvantaged student scores were a little better predictor of TCAP reading and language arts but not significantly better predictor than noneconomically disadvantaged student R-CBM scores. Two multiple regression analysis were conducted of all four predictors for economically disadvantaged student R-CBM score were strongly predictive ($R^2 = .52$) of TCAP reading and language arts scores. In the second analysis noneconomically disadvantaged students R-CBM

scores were found to be strongly predictive ($R^2 = .50$) of TCAP reading and language arts score. There was not a significant difference between the strength of relationship for the R-CBM predictors and TCAP reading and language arts scores when comparing economically disadvantaged students and noneconomically disadvantaged students.

A linear regression analysis was conducted to evaluate the ability of R-CBM (fall, winter, spring, and median) scores to predict TCAP reading and language arts scores for all third grade students. R-CBM scores for fall, winter, spring, and median were positively linearly related such that as R-CBM scores increase the TCAP reading and language arts score correspondingly increases. The accuracy the equation created to predict TCAP reading and language arts scores from a single R-CBM score was strong 0.71 (fall, winter, and spring) and 0.72 (median). Variation in the R-CBM scores ranged approximately 50% to 52% when compared to TCAP reading and language arts scores.

Conclusions

As a result of the analyses, it was determined that R-CBM scores were strong predictors of TCAP reading and language arts scores for third grade students in this population. Based on results of a multiple regression analysis all four R-CBM predictor variables (fall, winter, spring, and median) demonstrated a strong bivariate correlational relationship. This finding was in agreement with Roehrig et al. (2008) who found a strong to relationship r =.66 to .71 between R-CBM and the Florida Comprehensive Assessment Test (FCAT) reading assessment. The result also supported the findings of Hintze and Silberglitt (2005) that found a strong correlation between R-CBM (fall, winter, and spring) and the Minnesota Comprehensive Assessment (MCA) in reading subtest. This study will contribute to the growing body of research that

substantiates the ability of R-CBM to predict student outcomes on group administered standardized measures of reading, reading comprehension and standards based criterion measures of reading (Baker et al., 2008; Crawford et al., 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Roehrig et al., 2008; Sibley, Biwer, & Hesch, 2001; Stage & Jacobsen, 2001; Wood, 2006).

A significant finding using multiple regression analysis was fall R-CBM scores were about as good as other (winter, spring, and median) R-CBM scores at predicting TCAP reading and language arts scores. Fall R-CBM was similarly predictive .70 as winter, spring and median .71, .71, and .72 respectively. Marston (1992) found that teachers and administrators were difficult to convince that R-CBM could function as a valid indicator of a student's overall reading skills. The results of this study substantiate the use of R-CBM in measuring global reading skills as measured by the TCAP reading and language arts assessment.

The fall R-CBM offers teachers and schools the timeliest data with which to make changes to improve student reading outcomes as measured in the TCAP reading and language arts assessment. The strength of the R-CBM predictive relationship for TCAP reading and language arts scores was similar for all students despite gender or economic status. This allows for grater generalization to all students despite membership in a particular recognized subgroup.

The strength of the positive linear relationship between R-CBM (fall, winter, spring, and median) and the TCAP reading and language scale score made it possible to develop linear equations. This equation predicts with a reasonable level of confidence, approximately 50%, 51%, 50%, and 52%, of the variation for fall, winter, spring, and median R-CBM respectively and TCAP reading and language arts scale score. A fall R-CBM score of 99 words read correctly predicts a TCAP reading and language arts score of 760.36. The cut score for

proficiency on the TCAP reading and language arts was 760. The R-CBM results for predicted proficiency on the TCAP reading and language arts assessment for third grade students in this study was in agreement with Fountas and Pinnell's Recommended Oral Reading Rates (2009) for third grade students which suggested that third grade students should earn a R-CBM score of 100 – 140 words read per minute. A winter R-CBM score of 116 words read correctly predicts a proficient TCAP reading and language arts score of 760.51. A spring R-CBM score of 133 words read correctly predicts a proficient TCAP reading and language arts score of 760.40. The median R-CBM predicted score was not useful as it mirrored the results of the winter R-CBM results and could not be calculated until all three assessments were administered.

The use of this equation may allow a teacher to determine appropriate instructional goals for students in the area of reading and periodically monitor a student's progress using commercially available R-CBM 1-minute assessments. Using the fall R-CBM scores a teacher may be able to identify students at risk of not reaching proficiency while sufficient time exists to change the intensity, duration, and methods of reading instruction in order to prevent the student from achieving a less than proficient score on the high stakes assessment.

Recommendations for Practice

This researcher recommends that the participating school district continue to use a systematic benchmarking framework in reading though the use of fall, winter, and spring universal R-CBM screening. This process when connected to a structured, data driven, response to intervention (RTI) process may be able to identify students in need of additional specialized research based instruction in reading before students fall too far behind. Researchers at the Florida Center for Reading Research (2007) studied characteristics of Reading First schools

whose rate of reading growth exceeded schools with similar demographical populations. The highest performing schools exhibited the following seven common characteristics: strong leadership, positive belief and teacher dedication, data use and analysis, effective scheduling, professional development, scientifically based intervention programs, and parent involvement. It is hoped that the analysis of the findings of this study will convince reluctant administrators and teachers that R-CBM has a predictive value in early identification of students who are at risk of not earning proficient scores on the TCAP reading language arts subtest.

Use of the regression equation for predicting the overall TCAP reading and language arts score for particular levels of reading fluency as measured by a R-CBM will practically connect the use of R-CBM with student success on high stakes assessments. Teachers and principals in Tennessee are evaluated in part on the progress of students demonstrate in high stakes assessments and the use of existing data to improve student outcomes will also improve teacher and principal effectiveness scores. This study provides a tool with which teachers can better use preexisting data determine the effectiveness of a student's instructional program. The ability to predict a student's TCAP reading and language arts score reliably at the beginning of the school year (fall) based on the number of words read correctly in a 3 minutes R-CBM passage should be empowering. If a student is able to achieve a fall R-CBM score of 99, WRC it is likely they will achieve a proficient score of the TCAP reading and language arts assessment. In the winter a student who continues on a trajectory to a proficient score on the TCAP reading language arts assessment would require a winter R-CBM score of 116 WRC. In order remain on a path to a proficient TCAP reading and language arts score a third grade student would have to earn a spring R-CBM score of 133 WRC to predict a proficient score. Using R-CBM data teachers can continually monitor a student's progress though out the school year with existing data. If the

student fails to meet the expected benchmark for proficiency, a plan could be developed to improve the rate of growth before a student is significantly at risk for not achieving a proficient reading score on state assessment. It is also recommended that students who are predicted to achieve a nonproficient score should be more frequently progress monitored using R-CBM to monitor a student's progress toward improved reading skills. By using R-CBM data to target instruction to meet the individual needs of each student all students will continue to develop effective reading skills.

Recommendations for Future Research

This study should be replicated with other populations to further validate its results and strengthen the body of research surrounding the relationship between R-CBM and the TCAP reading and language arts assessment. Replicating in a more culturally diverse population that more closely mirrors the demographic distribution of the State of Tennessee would enhance the ability to generalize the results of this analysis as of 2013. Forty-six states, including Tennessee, have adopted the Common Core State Standards (CCSS) for English Language Arts. The expectations for students must be dramatically increased to meet the CCSS. By school year 2014 – 2015, 22 states have agreed to adopt the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment. A similar study should be replicated when student PARCC English language arts scores become available to determine if R-CBM scores continue to have significant predictive utility. This analysis could also determine the required R-CBM score needed to predict proficiency on the new PARCC assessment (PARCC Fact Sheet, 2013). Future research is needed in the area of R-CBM universal screening and applications at other grade levels in order to determine if R-CBM is predictive of TCAP reading and language arts

scores in grades 4-8. It would also be useful to determine fall, spring, and winter R-CBM scores that would predict proficiency on state reading assessment for each grade level evaluated.

REFERENCES

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum-based measurement slope estimating using data from triannual universal screenings. *School Psychology Review*, 37, 109-125.

Ardoin, S. P., & Christ, T. J. (2009). Curriculum-based measurement of oral reading: standard error associated with progress monitoring outcomes from DIBELS, AIMSweb, and an experimental passage set. *School Psychology Review*, 38, 266-283.

Ardoin, S.P., Suldo, S. M., Witt, J.C., Aldrich, S., & McDonald, E. (2005). Accuracy of readability estimates' predictions of CBM performance. *School Psychology Quarterly*, 20, 1-22.

Ardoin, S. P., Witt, J. C., Suldo, S. M., Connel, J. E., Koenig, J. L., & Resetar, J.L. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probe when conducting universal screening. *School Psychology Review*, 33, 218-233.

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'enui, E. J. & Beck, C. T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high poverty schools. *School Psychology Review*, 37, 18-37.

Berkely, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention. *Journal of Learning Disabilities*, 42, 85-95.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, S. M., & Shaywitz, B. A. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow up. *Educational Psychology*, 96, 444-461.

Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly*, 2, 5-142.

Bradley, R., Danielson, L., & Doolittle, J. (2005). Response to intervention. *Journal of Learning Disabilities*, 38, 8-13.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, 64, 723-733.

Chall, J. S. (1967). *Learning to read: The great debate*. New York, NY: McGraw-Hill.

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

Cummings, K. D., Atkins, T., Allison, R., & Cole, C. (2008). Response to intervention: Investigating the new role of special educators. *Teaching Exceptional Children*. 40, 24-31.

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

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

Deno, S. L., Marston, D., Shinn, M., & Tindal, G. (1983) Oral reading fluency: A simple datum for scaling reading disabilities. *Topics in Learning and Learning Disabilities*, 2, 53-59.

Deno, S. L., & Mirkin, P. K. (1977). *Data-based program modification: A manual*. Reston, VA: Counsel for Exceptional Children.

Deno, S. L., Mirkin, P. K. & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children*, 49, 36-45.

Deno, S. L., Wallace, T., & Espin, C. (2005). History of curriculum-based measurement: An interview with Stan Deno. Proceedings of the CEC conventions and expo, Baltimore, MD. Retrieved from http://progressmonitoring.net/video_stan_deno.html

Denton, C., Foorman, B. R., & Mathes, G. G. (2003). Schools that "beat the odds": Implications for reading instruction. *Remedial and Special Education*, 24, 258-261.

Duncan, A. (2012, February). Approval of Tennessee's request for ESEA flexibility. Retrieved from:http://www.tn.gov/education/doc/ESEA_Flexibility_Approval_Letter.pdf

Educational Testing Services; K-12 Education, Custom Assessments and Programs.(2012) Retrieved from: : http://www.ets.org/k12/programs/custom_assessments

Flesch, R. (1955). *Why Jonny can't read*. New York, NY: Harper & Row.

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology*, 31, 1-29.

Fountas and Pinnell Recommended Oral Reading Rates. (2009). Retrieved June 17, 2013, from http://bernardston.pioneervalley.k12.ma.us/documents/OralReadingRates.pdf

Fuchs, D., Johnson, E., McKnight, M. A., & Mellard, D. F. (2006). *Responsiveness to intervention (RTI):How to do it*. Lawrence, KS: National Research Center on Learning Disabilities

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

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

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

Fuchs, L. S., & Fuchs, D. (2006). A framework for building capacity for responsiveness to intervention. *School Psychology Review*, 35, 621-626.

Fuchs, L. S., Fuchs, D., & Compton, D. L., (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*, 71, 7-21.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R., (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5, 239-256.

Gall, M. D., Borg, W. R., & Gall, J. P., (1996). *Educational research: An introduction*, White Plains, NY: Longman.

Gerber, M., (2005). Teachers are still the test: Limitations of response to instruction strategies for identifying children with learning disabilities. *Journal of Learning Disabilities*, 38, 516–524

Gersten, R., & Keating, T., (1987). Long term benefits from direct instruction. *Educational Leadership*, 45, 28-31.

Graney, S. B., & Shinn, M. R. (2005). Effects of reading curriculum-based measurement (R-CBM) teacher feedback in general education classrooms. *School Psychology Review*, 34, 184-201.

Heller, K. A., Holtzman, W., & Messick, S. (1982). *Placing children in special education: A strategy for equity*, Washington, DC: National Academy Press.

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

H.R. 1350--108th Congress: Individuals with Disabilities Education Improvement Act of 2004. (2003). In www.GovTrack.us. Retrieved from http://www.govtrack.us/congress/bills/108/hr1350

Justice, L. M. (2006). Evidenced-based practice, response to intervention, and the prevention of reading difficulties. *Language, Speech, and Hearing Services in Schools*, 37, 284 – 297.

Individuals with Disablities Education Act of 2004, P.L.108-446 Retrieved from: http://idea.ed.gov/download/statute.html

Johnson, E., Mellard, D., Fuchs, D., & McKnight, M. (2006). *Response to intervention (RtI): How to do it*. Lawrence, KS: National Research Center on Learning Disabilities.

Kame'enui, E. J., Fuchs, L. S., Francis, D. J., Good, R., O'Connor, R. E., Simmons, D. C., Tindal, G., & Torgesen, J. K. (2006). The adequacy of tools for assessing reading competence: A framework and review. *Educational Researcher*, 35, 3-11.

Mathes, P. G., Denton, C. A., Fleche, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). An evaluation of two reading interventions delivered from diverse models. *Reading Research Quarterly*, 40, 148-182.

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review*, 33, 193-203.

Mesmer, E. M., & Mesmer, H. A. E. (2008). Response to intervention (RTI): What teachers of reading need to know. *The Reading Teacher*, 62, 280-290.

Moats, L. (2009). Knowledge foundations for teaching reading and spelling. *Reading and Writing*, 22, 379-399.

National Association of State Directors of Special Education.(2005). Response to intervention: Policy considerations and implementation. Alexandria, VA: Author. Retrieved from http: http://nasdse.org/portals/0/school.pdf

National Commission on Excellence in Education, (1983). *A Nation at risk: The imperative for educational reform*. Washington, DC: Department of Education.

National Center on Accessing the General Curriculum. (2002). Curriculum-based evaluations: Effective classroom practices report (OSEP Publication No. H324H990004). Retrieved from http://www.cast.org/system/galleries/download/ncac/CurBasEval.pdf

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards*. Washington DC: Author.

National Institute of Child Health and Human Development. (2000). *Report of the National reading panel teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: Author.

No Child Left Behind Act (NCLB) of 2001, P.L. 107-110. § 115, Stat. 1425 (2002).

Oller, J. W., & Jonz, J. (1994). *Cloze and coherence*. Cranbury, NJ: Associated University Presses.

Partnership for Assessment of Readiness for College and Careers. (2013). PARCC fact sheet. Retrieved June 17, 2013, from: http://www.parcconline.org/sites/parcc/files/PARCCFactSheetandFAQsbackgrounder_FINAL.pdf

U.S. Department of Education, Office of Special Education Programs. (2007, February). Alignment of IDEA and NCLB. Retrieved from U.S. Department of Education website: http://idea.ed.gov/object/fileDownload/model/TopicalBrief/field/PdfFile/primary_key/3

U.S. Department of Education. (2011, September). What ESEA flexibility means for students, teachers, and parents: Answering the public's questions. Retrieved from: www.ed.gov/sites/default/files/answering-public-questions.doc

U.S. Department of Education. (2007). Questions and answers on response to intervention (RTI) and early intervening services (EIS). Retrieved from U.S. Department of Education website: http://idea.ed.gov/object/fileDownload/model/QaCorner/field/PdfFile/primary_key/8

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, 47, 427-469.

Reschly, D. J., & Hosp, J. L. (2004). State SLD identification policies and practices. *Learning Disability Quarterly*, 27, 197–213.

Riedel, B. (2007). The relationship between DIBELS, reading comprehension, and vocabulary in urban first grade students. *Reading Research Quarterly*, 42, 546-567.

Roehrig, A. D., Duggar, S. W., Moats, L. C., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction: Identifying potential supports and challenges. *Remedial and Special Education*, 29, 364-382.

Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measures for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, 46, 343-366.

Rozin, P., & Gleitman, L. R., (1977). The structure and acquisition of reading II: The reading process and the acquisition of the alphabetic principle. *Toward a Psychology of Reading*, 55-141.

Rupley, W. H., Blair, T. R., & Nichols, W.D. (2009). Effective reading instruction for struggling readers: The role of direct/explicit teaching. *Reading and Writing Quarterly*, 25, 125-138.

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

Shinn, M. R, & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of reading curriculum-based measurement (R-CBM) for use in the general outcome measurement*. Eden Prairie, MN: Edformation.

Shores, C., & Chester, K. (2009). Using *RTI for school improvement*. Thousand Oaks, CA: Corwin Press.

Sibley, D., Biwer, D., & Hesch, A. (2001). *Establishing curriculum-based measurement oral reading fluency performance standards to predict success on local and state tests of reading achievement*. Arlington Heights, IL: Arlington Heights Public School District 25.

Sindelar, P., Shearer, D., Yendol-Hoppey, D., & Liebert, T. (2006). The sustainability of inclusive school reform. *Exceptional Children*, 72, 317-331.

Spach, G. D., & Spache, E. B. (1973). *Reading in the elementary school* (3rd ed.) Boston, MA: Allyn & Bacon.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review*, 30, 407-419.

Stahl, S. A., Osborn, J., & Lehr. F. (1990). *Beginning to read: Thinking and learning about print, A summary*. Champaign: University of Illinois Center for the Study of Reading.

Standerfer, L. (2006). Before NCLB the history of ESEA. *Principal Leadership*, 6, 26-27.

Stecker, P. M., Fuchs, D. S., & Fuchs, L. S. (2008). Progress monitoring as essential practice within response to intervention. *Rural Special Education Quarterly*, 27, 10-17

Tennessee Department of Education; Division of Assessment (2012). Tennessee comprehensive assessment program. Nashville, TN: Retrieved from http://tennessee.gov/education/assessment/achievement.shtml

Tennessee Department of Education; Division of Special Education (2010). Disability eligibility standards. Nashville, TN: Retrieved from http://www.state.tn.us/education/speced/doc/41610apprvdelig.pdf

Tennessee Department of Education (2011, November). *ESEA flexibility request*. Retrieved from http://www.tn.gov/education/doc/ESEA_Flexibility_Request.pdf

Tennessee Department of Education. (2012, March). Grand division meetings: ESEA flexibility wavier regarding accountability and title 1 provisions (Presentation Slides Presented: March 29, 2012.

Tennessee Department of Education, Tennessee Reading First. (2006). Tennessee reading first intervention guide. Retrieved from The University of Tennessee – Knoxville website: http://www.cls.utk.edu/pdf/Intervention_Guide_final.pdf

Tennessee Department of Education (2011). Report card on Tennessee schools. Nashville, TN: Retrieved from: http://www.tn.gov/education/reportcard/

Torgesen, J., Houston, D., Rissman, L., & Kosanovich, K. (2007). Teaching all students to read: A guide for principals. Portsmouth, NH: RMC Research Corporation, Center of Instruction in Elementary School Florida Center for Reading Research.

Vaughn, S., & Fuchs, L. S. (2003). Reading learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice*. 18, 137-146.

Vaughn, S., & Roberts, G. (2007). Secondary intervention in reading: Providing additional instruction for students at risk. *Teaching Exceptional Children, 39*, 40–46.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early interventions as a vehicle for

distinguishing between cognitive and experimental deficits as basic causes of specific reading disabilities. *Journal of Educational Psychology*, 88, 601-638.

Villaume, *S.K., &* Brabham, *E. G.* (*2003*). Phonics instruction: Beyond the debate. *The Reading Teacher*,. 56, 478-482

Wanker, W., & Christie, K. (2005). State implementation of the No Child Left Behind Act. *Peabody Journal of Education*, 80, 57-72.

Wood, D. E. (2006). Modeling the relationship between oral reading fluency and performance on a statewide reading test. *Educational Assessment* 11, 85-104.

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

Zimmerman, B. J., & Dibenedetto, M. K. (2008). Mastery learning and assessment: Implications for student and teachers in an era of high stakes testing. *Psychology in the Schools*, 45, 206-216.

APPENDICES

APPENDIX A

Letter to Director of Schools

May 13, 2013

Mr. [redacted], Director of Schools
[redacted]
[redacted]
[redacted] Tennessee [redacted]

RE: Permission to Conduct Research Study

Dear Mr. [redacted],

I am writing to request permission to conduct a research study utilizing historical data. All data is requested without individual student identifiable information. I am currently enrolled in the Department of Educational Leadership and Policy Analysis, School Leadership Program, at East Tennessee State University (ETSU) in Johnson City, TN, and am in the process of writing my doctorial dissertation. The study is entitled, "Predicting Performance on the Tennessee Comprehensive Assessment for Reading for Third Graders using Reading Curriculum Based Measures". This project will be conducted under the supervision of Dr. James Lampley, Ed.D.; Research Specialist and Assistant Professor of Education.

With your permission I request archrival data from the Tennessee Comprehensive Assessment for Reading and Reading Curriculum Based Measures for third grade students during the 2011 school year. This data is requested in a format that is not traceable to any individual student or school. If approved, study results will be utilized to determine: how well fall, winter, spring, and median third grade R-CBM scores predict reading proficiency scores on the Tennessee TCAP assessment for students and the linear equation that predicts TCAP reading scores from winter R-CBM scores for third graders and how well does this equation predict TCAP reading scores. It was determined by the East Tennessee State University, Office for the Protection of Human Subjects that this research did not meet the Food and Drug Administration (FDA) nor the Department of Health and Human Services (DHHS) definition of research involving human subjects.

Your approval to conduct this study will be greatly appreciated. If you have any questions or concerns regarding this request please contact me at my email address: zrsk8@goldmail.etsu.edu.

Sincerely,

Robert Scott Kirkham
Graduate Student ETSU
Cc: James Lampley, Ed.D. Chairperson, ETSU

APPENDIX B

Letter from Director of Schools

May 13, 2013

Mr. Robert Scott Kirkham

506 Belle Meade Drive

Maryville, TN 37803

(865) 379-1055

Dear Mr. Kirkham,

The purpose of this letter is to inform you that I give permission to conduct the research titled, "Predicting performance on the Tennessee Comprehensive Assessment for Reading for Third Graders using Reading Curriculum Based Measures" utilizing historical data from Blount County Schools. This also serves as assurance that this District complies with requirements of the Family Educational Rights and Privacy Act (FERPA) and the Protection of Pupil Rights Amendment (PPRA) and will ensure that these requirements are followed in the conduct of this research.

Sincerely,

[redacted]

Director of Schools

VITA

ROBERT SCOTT KIRKHAM

Personal Data:	Date of Birth: September 16, 1974
	Place of Birth: Trenton, New Jersey
	Marital Status: Married
Education:	A.A. Humanities and Social Science, Mercer County Community College, West Winsor, New Jersey 1994
	M.A. Educational Psychology and Counselor Education, Tennessee Technological University, Cookeville, Tennessee 1999
	Ed.S. Educational Psychology and Counselor Education, Tennessee Technological University, Cookeville, Tennessee 2000
	Ed.S. Instructional Leadership, Tennessee Technological University, Cookeville, Tennessee 2000
	Ed.D. Educational Leadership, East Tennessee State University, Johnson City, Tennessee 2013
Professional Experience:	Graduate Assistant, Tennessee Technological University, College of Education, Office of Laboratory Experiences, Cookeville, Tennessee, 1998 – 2000

Teaching Assistant, Tennessee Technological University, College of Arts and Science, Department of Counseling and Psychology, Cookeville, Tennessee, 2000

School Counselor, Blount County Schools, Community Schools Grant, Maryville, Tennessee, 2000 -2003

School Counselor, Lanier Elementary School, Maryville, Tennessee, 2000-2003

School Psychologist, Blount County Schools, Maryville, Tennessee, 2003 – 2007

Student Services and Special Education Supervisor / Director, Blount County Schools, Maryville, Tennessee, 2007 – 2013

Honors and Awards:

Craig Steven Johnson Scholarship Award 1995
Outstanding Young American Award 1996
National Collegiate Education Award – 1998
All-American Scholar – 1999
Pi Lambda Theta – International Honor Society and Professional Association in Education – 2000
Phi Kappa Phi – Multidiscipline Honor Society – 2002
Phi Delta Kappa – Graduate Educational Honor Society – 1999
Tennessee Association for Administrators in Special Education (TAASE) Cup East Tennessee Special Education Administrator of the Year – 2010
Special Education Connection “Star Special Education Director” Recognition – 2011