Curriculum-Based Fluency and Comprehension Measurements as Predictors of Elementary Student Performance on State-Wide MAP Communication Arts Assessment Scores in an RTI Setting

Sandra K. Rice
B.A. UCM State University, 1999
M.S. William Woods University, 2004

Submitted to the Graduate Department and Faculty of the School of Education of Baker University in partial fulfillment of the requirements for the degree of Doctor of Education in Educational Leadership

_____
Harold Frye, Ed. D., Major Advisor

_____
Charlsie Prosser, Ed.D.

_____
Ryan Rostine, Ed. D.

Date Defended: November 29, 2016

Abstract

This quantitative study included fourth-grade students from District XYZ, a large suburban school district located in the mid-west. The purpose of this study was to determine if CBM oral reading fluency and CBM MAZE comprehension interim benchmark assessment scores obtained in an RTI framework predicted the summative Missouri MAP Communication Arts scores. The multiple regression analyses models revealed that CBM oral reading fluency and MAZE comprehension interim assessments for fall and spring were moderately strong predictors of the Missouri summative MAP assessment. The winter CBM oral reading fluency and MAZE comprehension interim assessments did not increase the predictive value of the equation; therefore, these were not included in the results. Based on the results of this study, the recommendation is made to utilize CBM oral reading fluency and MAZE comprehension interim assessments in an RTI framework to provide data to district leaders and educators of students who may need academic remediation.

Dedication

First, this dissertation is dedicated to my husband Ron, who without you, I could not have endured the difficult trials during this doctoral study. You have always been my rock, cheerleader, and best friend who encouraged and believed in me always. Thank you with all my heart for your love, understanding, and patience. Next, I want to dedicate this study to my wonderful daughters, Jennifer and Emilee who have supported their mother's determination to make the world a better place. I hope you find that hard work and perservence are qualities that will help you reach your goals. Finally, I want to dedicate this study to my parents and brother, Bob, who passed on to our Lord during this journey. Although you are not here in flesh to see me graduate, I know you are here in spirit. I miss you dearly.

Acknowledgements

I want to express my gratitude to my family, friends and colleagues for the professional and moral support during my doctoral study and the writing of this dissertation. To my parents, who as a young child, instilled in me a hard work ethic and the tenacity to persevere even in tough times. To my brother, Rick and sister-in-law Kathy, for the many encouraging words and belief in my success. To my stepdad, Bob for his encouragement and love. I would like to give a heartfelt thank you to Dr. Harold Frye, my major advisor, for his immense knowledge that he has shared with me over the past decade. Without his support and encouragement, this project would not have been possible. I would like to thank Dr. Phillip Messner, my research analyst for his relentless work with me to understand the analytical process, and the rest of my committee, Dr. Charlsie Prosser, and Dr. Ryan Rostine for their valuable feedback and countless hours on this study.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction

Never before have literacy skills been as critical to our nation's success as they are today. In 2008, the National Governors Association (NGA), the Council of Chief State School Officers (CCSSO), and Achieve, Inc. advised in *Benchmarking for Success: Ensuring U.S. Students Receive a World-Class Education*, that due to technological advancement and the expansion of world trade and political trends, quality jobs demanding higher skills are more difficult to obtain due to increased competition for these positions. Further, lower skill positions are being automated and outsourced to areas of the world where production and service pay is lower than in the United States (p. 5). The National Governors Association further contends, that for American workers to compete in this world's knowledge-and-innovative economy, they must have postsecondary education or training, strong fundamental skills in math and reading, and the ability to solve unfamiliar problems and communicate effectively (p. 5). "More jobs are going to the best educated no matter where they live, which means that Americans will face more competition than ever for work" (2008, p. 5).

According to labor economists Frank Levy and Richard Murnane, "Over the long run, better education is the best tool we have to prepare the population for a rapidly changing job market" (2012, p 155). Hanushek and colleagues (2008) analyzed economic data and found that to truly maximize growth, it is not enough to produce a

high-achieving elite; a nation's economic success depends upon closing achievement gaps to ensure that all students attain a solid foundation of knowledge and skills (p 68). Another recent study of 14 developed countries by Coulombe and Tremblay (2006) concluded, "Increasing the average level of literacy will have a greater effect on growth than increasing the percentage of individuals who achieve high levels of literacy skills" (p 23).

Background of the Study

Over the last two decades, American students have made little progress in reading. According to the latest National Center for Education Statistics (NCES) report (2014), "The number of children and youth ages 3–21 receiving special education services was 6.4 million in 2011–12, or about 13 percent of all public school students" (p. 8). The 2008 National Governors Association report, Benchmarking for Success: Ensuring U.S. Students Receive a World-Class Education, stated that on international assessments in 2003 and 2005, "American students performed only about average among industrialized countries, and fourth graders' reading scores have stagnated while other countries have made sizable gains" (p.12). The (NCES) reports only 35 percent of all fourth-grade students scored proficient or above in reading on the 2013 National Assessment of Educational Progress; up from 34 percent in 2011 (May, 2014).

Over the past five decades, the federal government has continued to pass and improve upon general and special education laws with the intention to improve the academic achievement of all children. The most ambitious are the No Child Left Behind

Act (NCLB) of 2001, and its recent plan to improve the Elementary and Secondary Education Act (ESEA) by President Barack Obama in *A Blueprint for Reform* (United States Department of Education, 2010). In 2004, United States Congress reauthorized the Individuals with Disabilities Improvement Act (IDEA), which provided an avenue to offer immediate assistance to at-risk students by initiating researched-based interventions when students first experience difficulties (IDEA, 2004). In 2009, 48 state departments, District of Columbia and two territories, along with governors and state commissioners of education began work to develop the Common Core Standards (CCS). The Common Core Standards depict what students should be able to know and do to be prepared upon graduation for college, career, and life. Upon completion of the CCS, 43 states voluntarily adopted them. Districts in those states began to align the CCS with research-based curriculum and high-quality instructional strategies (National Governor's Association, 2015). Together these laws and other recommendations serve a common purpose; to hold schools, districts and states to higher standards, with rigorous assessments, and increased accountability for the academic achievement of all students, and upon graduation, to ensure readiness to compete in a global economy.

The *Blueprint for Education, American Recovery, and Reinvestment Act of 2009* and more specifically the Individuals of Disabilities Education Act (IDEA) of 2004 laid the groundwork for the implementation of Response to Intervention (RTI) as a general education initiative. RTI is a multi-tiered system that ensures students at-risk of failure make adequate academic progress, using high-quality instruction and interventions,

guided by regular assessment data (National Association of State Directors of Special Education [NASDSE] and Council of Administrators of Special Education [CASE], 2006). RTI serves two purposes. First, it is a general education model to ensure that students receive high-quality instruction based on best practice. Further, classroom teachers use assessment data and early intervention if a student falls behind their peers academically or behaviorally. Classroom teachers provide scientifically researched-based interventions when students demonstrate a need and track their progress or lack thereof through progress monitoring. Second, RTI assists in the eligibility determination for special education services under specific learning disabilities (SLD) for students who do not respond to adequate instruction and scientific research-based interventions (U.S. Department of Special Education, 2007a; IDEA, 2004).

The implementation of the RTI framework necessitates a shift in the focus of regular education. Along with new laws, and revised curriculum, many researchers and educators alike, have resolved to transformative change in assessment practices in order to evaluate student knowledge, teacher quality, and school and district effectiveness. Now more than ever, general education teachers are responsible for meeting the needs of individual students. Teachers collect classroom data, analyze that data, change the intervention or make a referral to the special education system (U.S. Department of Special Education, 2007a).

There is no doubt that RTI has changed the landscape of assessment by broadening the use of and for evaluations. Stiggins states, “Todays schools are less

focused on merely sorting students and more focused on helping *all* students succeed in meeting standards" (2007 p.22). He further states, "Thus in order to fully inform instruction, we need both assessment *of* student learning and assessment *for* learning (Stiggins, p. 327). As early as 1998, Black and William found that consistently applying formative assessment principles in the classroom increased student achievement (p. 56).

Researchers cite that assessment is critical to effective instruction because data can be used to identify instructional goals and student needs, provide relevant instruction, and assess intended learning outcomes or mastery of skills, diagnosis gaps in learning, or to evaluate a particular program or pedagogy, or to predict student performance on end-of-the-year assessment (Linn & Gronlund, 2000; Perie, Marion, Gong & Wurtzel, 2007; Burns, 2010b; Stiggins, 2007). A rich learning environment has a delicate balance of formative, interim, and summative assessment.

In the school year of 2013-2014, the state of Missouri had 562 school districts that educated over 886,000 students who were enrolled in pre-kindergarten through grade 12 (Missouri Department of Secondary and Elementary Education [MO DESE], 2014, pg. 1). These school districts must comply with Missouri Senate Bill 380 (the Outstanding Schools Act), and NCLB and ESEA mandates. The Missouri School Improvement Program (MSIP) is the school district accountability system for review and accreditation based on compliance with state and national board policies (MO DESE, 2016a). The Missouri Improvement Program: Support and Intervention Plan utilizes a differentiated plan for each district based on district performance, and student needs based on five core

components. These components are; 1) A focus on children and families, 2) High expectations for all students, 3) Access to high-quality schools for every child, 4) Solutions to meet the needs of each district and community, 5) Early intervention and prevention (MO DESE, 2016a).

Missouri Assessment Program (MAP) standardized assessments were created to evaluate district's progress toward the Missouri Outstanding Schools Act of 1993 (MO DESE, 2016a). In alignment with this act, Missouri adopted the Missouri Show-Me Standards, a demanding set of content and process standards that identify knowledge, skills and competencies all students should acquire before graduation (MO DESE, 2016d). Grade Level Expectancies (GLE) are grade and course level expectations for specific subjects. Each district is mandated to participate in a summative statewide standardized assessment for designated grade level students in certain curricular areas each year to determine whether they are meeting the requirements of innovation, higher standards, and educational excellence (MO DESE, 2016a).

This study was conducted in a growing, mid-western suburban community with a district comprised of 117 square miles, serving surrounding six communities comprised of approximately 17,600 students. This researcher will refer to this district as District XYZ. One thousand three hundred ninety-nine certified staff work in eighteen elementary schools, three middle schools, three high schools, an alternative secondary school, a secondary technology academy, an early education, and a special education day treatment center. More specifically, approximately 700 certified elementary (K-6) staff

work with over 9,000 students (see Table 1) (Missouri Department of Secondary and Elementary Education, 2014).

In the last two years, a team of administrators and teachers redesigned the district's curriculum and instruction using components of the Common Core State Standards (CCSS) with the goal of increasing student achievement and ensuring the academic success of every student (District XYZ, Five Year Plan, 2013). This process will take five years to complete and will include Essential Standards, Learning Targets, District Summative Assessments, and Alignment to National and State Standards. During the 2013-14 school year, teachers began phase one of the five-year plan by teaching the newly updated curriculum. On the website it stated,

> The elementary curriculum emphasizes mastery of foundational skills in reading, math, science, social studies, communication arts, health, art, music and physical education. The curriculum is individualized to meet special needs, and lessons are provided to develop social skills, positive character traits and drug-abuse resistance skills (District XYZ, Curriculum and Instruction page, July 28, 2014).

District XYZ initiated Response to Intervention (RTI) in all eighteen elementary schools during the 2013-14 school year, to meet the needs of the lowest functioning, academic students in math and reading (District XYZ, Five Year Plan, 2013). Response to Intervention would provide effective practice to promote the achievement of all students while meeting the district mandated goals outlined in Missouri's Department of

Secondary and Elementary (DESE) No Child Left Behind Flexibility Waiver, Top 10 by 20 Initiative, and MSIP 5 goals.

In a Brief, the National Center on Response to Intervention (NCRTI, 2010) provided the following definition of RTI:

> Response to Intervention integrates assessment and intervention within a multi-level prevention system to maximize student achievement and to reduce behavior problems. With RTI, schools identify students at risk for poor learning outcomes, monitor student progress, provide evidence-based interventions and adjust the intensity and nature of those interventions depending on a student's responsiveness, and identify students with learning disabilities. (p. 2)

As part of their Assessment Plan, District XYZ began the process of utilizing AIMSweb, a data management system, to periodically provide formative or interim evidence of student academic progress or lack of on grade level skills and standards. An RTI consultant, employed by the district, provided an implementation manual and training to teams of teachers and administrators from each school building (District XYZ, Assistant Superintendent, 2013).

District XYZ also participated in the MAP summative assessment for grades 3-8 in Communication Arts. The Communication Arts MAP summative assessment is important because it provides information on what students learned and were able to do by the end of their school year. However, teachers needed periodic assessments or

checks in the way of formative assessments, to ensure students were reaching appropriate academic goals and standards throughout the course of the year.

In order to provide the needed on-going evidence, District XYZ, utilized Curriculum-based Measurements (CBM) several times during the course of a school year District XYZ, RTI manual, 2013). Shinn defines CBM's as a "set of standardized and validated short-duration tests in reading, math computation, math applications, spelling, written expression, early literacy and early numeracy" (2007, p. 608). CBM's have been identified as an accurate and efficient measure of general readability, an indicator of instructional effectiveness, student academic growth, and success on high-stakes tests (Deno, Fuchs, Marston, & Shinn 2001; Deno, & Mirkin, 1977; Deno, 2003). R-CBM is widely used for progress monitoring, screening for academic difficulties, and setting and monitoring individualized goals of students (Deno et al., 2001).

Statement of the Problem

The desire of the district is that teachers will be able to determine if a student has an academic deficit and intervene immediately to remediate any risk of failure in reading. However, the district and others do not know if there is a relationship between the on-going AIMSweb R-CBM and MAZE CBM formative assessments and Missouri's spring Communication Arts MAP summative assessment. Between 2009 and 2013, students in the state of Missouri averaged 51% in the proficiency and advanced range on the fourth grade Communication Arts MAP assessment (MO DESE, 2016b). During that same time, District XYZ fourth grade students averaged 62 % on the same assessment.

Therefore, within the state and the district itself, 37 to 49 percent of students were reading below the rate of their age-appropriate peers.

Beginning in the 2013-2014 school year, all students in grades K-6 were given benchmark reading curriculum-based measurements (R-CBM) and MAZE CBM reading comprehension measurements three times each year; fall, winter, and spring. From these criterion assessments, students who score below 25 percent are provided classroom remediation using the district RTI model (District XYZ, 2016). It is important to know if there is a relationship between the interim and summative assessments to justify the usage of the R-CBM and MAZE CBM as a source of predictive assessment data.

Purpose of the Study

The purpose of this study is to determine the extent of the relationship between reading scores on AIMSweb Communication Arts interim benchmarks assessments (R-CBM and MAZE CBM) and that of the Communication Arts MAP summative assessment for students enrolled in district XYZ. By using formative assessments that are aligned with the summative Communication Arts MAP test, teachers can easily attend to students needs immediately, when problems first arise. Also, the educators and patrons will be able to determine that AIMSweb R-CBM and MAZE CBM criterion assessments align with the District, State and National Curricular Standards in reading.

Significance of the Study

The intended result of this study is to determine if R-CBM and MAZE CBM formative assessments are predictive of the summative Communication Arts MAP

assessment. When a student demonstrates a score below 25% on benchmark formative tests, early remediation with research-based interventions, can be created to lower the achievement gap in reading. This study will help to inform educators and the public of the impact of interim assessments that are aligned with academic goals and standards have on improving reading for elementary students in District XYZ in all academic levels. Also, this study may provide evidence the district needs to meet district and state annual progress as outlined in Missouri's Top 10 by 20 and MSIP 5 goals.

Delimitations

"Delimitations are self-imposed boundaries set by the researcher on the purpose and scope of the study" (Lunenburg & Irby, 2008, p. 134). The study was limited to one specific school district chosen for this study.

1. The sample was delimited to 4th graders in District XYZ.
2. Outcomes were limited to student interaction in reading on the CBM and MAZE CBM and MAP Communication Arts compared to student RTI Tier level I, II or III.
3. The sample only included students who participated in each of the fall, winter, and spring CBM assessments and the spring MAP assessment in 2014.

Assumptions

The following assumptions regarding the effectiveness of the Response to Instruction process and student achievement were made as part of the organization of this study.

This study included the following assumptions:

1. Teachers administered the CBM assessments in a standardized manner.
2. Teachers scored the CBM assessments for their students in a standardized manner.
3. Individual school personnel entered their student data into the NCS Pearson, Inc. PsychCorp AIMSweb program accurately.
4. Teachers administered the MAP assessments in a standardized manner.
5. State personnel scored the MAP assessments in a standardized manner.
6. Handling of the materials for the MAP assessments was performed in an ethical and legal manner, following state guidelines.
7. Students put forth their best effort on all given assessments.

Research Questions

RQ1. What combination of variables (fall Oral Reading Fluency CBM score, winter Oral Reading Fluency CBM score, spring Oral Reading Fluency CBM score) best predict student Communication Arts scale scores?

RQ2. What combination of variables (fall Comprehension Fluency MAZE CBM score, winter Comprehension Fluency MAZE CBM score, spring Comprehension Fluency MAZE CBM score) best predict student Communication Arts scale scores?

Definitions of Terms

Lunenburg and Irby (2008) reported the definition of terms included "all key terms central to the study and used throughout the dissertation." (p. 118)

AIMSweb. Brown-Chidsey and Steege defines AIMSweb as an internet-based data management service that provides an RTI-specific data management tool that encompasses curriculum-based measurements (CBM) for universal screening and progress monitoring. AIMSweb manages students' scores as they move between different stages of RTI (Brown-Chidsey & Steege, 2010).

Fidelity. Fidelity refers to the accuracy, loyalty, and attentiveness with which an intended research design for instruction and intervention is implemented. To ensure standardization, intervention specialists must generally follow a prescribed protocol in order to attend to a program's or strategy's fidelity (Center on Innovation and Improvement, n.d.).

Specific Learning Disability. A specific learning disability (SLD) is a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations (U.S. Department of Special Education, 2007b).

Student Progress Monitoring. Student progress monitoring is a scientifically based practice that is used to frequently assess students academic performance and evaluate the effectiveness of instruction. Progress monitoring procedures can be used with individual students or the entire class (Klotz & Canter, 2007).

Universal Screening. Klotz and Canter describe universal screening as a step taken by school personnel during the school year to determine which students are "at risk" for not meeting grade level standards or those who have behavior or emotional problems that may interfere with their learning (2007).

Organization of the Study

This dissertation is divided into five chapters. Chapter one presented the introduction, background and conceptual framework, and statement of the problem. Additionally, the significance of the study was described along with a purpose statement and the delimitations and assumptions underlying the study. The chapter concluded with a listing of the research questions that guided the study, a brief overview of the methodology used to conduct the research and the definitions of terms. Chapter two presents a review of the literature including an exploration of the national movement to common standards and assessments and meeting the needs of all learners in a timely manner. Chapter three examines the methodology used, research design, population, and sampling procedures. Also, instrumentation, measurement, data collection, reliability, and validity are addressed. Chapter four presents the results of the study through analysis, statistical and hypothesis testing. Finally, chapter five reveals a summary of the

study, discussion of the findings, implications for practice and recommendations for future research.

Chapter Two

Review of Literature

Introduction

This chapter reviews several strands of literature surrounding RTI. The first strand provides a historical evolution of the learning disability construct and controversy surrounding the identification of learning disabilities past and present. The second strand provides several definitions and the essential components of RTI, including the use of formative and summative assessments. The third strand discusses the RTI process with a focus on prevention and intervention in reading. The rationale for including the current and historical literature of RTI will support and demonstrate that for the past four decades, researchers and educators have worked to establish a system of checks and balances to ensure all students learn to the best of their ability. This chapter also supports this study, as this district strives to close the reading gap of its students, while adding to the RTI conceptual framework.

Historical Evolution of Response to Intervention

Samuel Kirk is believed to have coined the expression Learning Disability for the first time in 1962 in his text, Educating Exceptional Children (Kirk & Kirk, 1983; Hallahan & Mercer, 2001) to describe children who suffered from a discrepancy between achievement and capacity to learn. Kirk and Kirk provided the definition of a learning disability as:

> A learning disability refers to retardation, disorder, or delayed development in one or more of the processes of speech, language, reading, spelling, writing or arithmetic resulting from a possible cerebral dysfunction and/or emotional or behavioral disturbance and not from mental retardation, sensory deprivation, or cultural or instructional factors. (p 20)

In 1965, Barbara Bateman offered another definition of learning disabilities. This definition played a significant role in the adoption of the IQ-achievement discrepancy model as a way to identify students with learning disorders to the forefront of research (as cited in Hallahan & Mercer, 2001). It read:

> Children who have learning disorders are those who manifest an educationally significant discrepancy between their estimated potential and actual level of performance related to basic disorders in the learning process, which may or may not be accompanied by demonstrable central nervous system dysfunction, and which are not secondary to generalized mental retardation, educational or cultural deprivation, severe emotional disturbance, or sensory loss. (p. 14)

After half a century of research on students who struggle academically, Federal Legislation PL 94-142, The Education of All Handicapped Children Act was enacted, to provide federal funding for students diagnosed as having a Learning Disability (LD) (Hallahan & Mercer, 2001). The law guided educators to determine if a student had a learning disability using an ability-achievement discrepancy formula (Kirk & Kirk, 1983; Hallahan & Mercer, 2001; Lyon, 1996). Students who had a significant gap between

scores on aptitude or cognitive tests and lower performance on academic achievement testing were diagnosed as Learning Disabled (Kirk & Kirk, 1983; Fuchs, Mock, Morgan, & Young, 2003). The law PL 94-142 defines a learning disability similar to Kirk's earlier definition. It read:

> ….children who have a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. Such disorders include such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Such term does not include children who have learning problems, which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, of emotional disturbance, or environmental, cultural, or economic disadvantage. (1975, p. 89 STAT. 794)

The purpose of this legislation provided an avenue for all children with disabilities to have a free and appropriate education. It had four goals:

- to assure all children with disabilities have available to them…a free appropriate public education which emphasizes special education and related services designed to meet their unique needs.
- to assure the rights of children with disabilities and their parents…are protected.
- to assist States and localities to provide for the education of all children with disabilities.

- to assess and assure the effectiveness of efforts to educate all children with disabilities. (U.S. Department of Education, 2010a)

From the work of researchers beginning in the 1970s, new insights allowed the learning community to work with all students at-risk of learning that eventually led to the development of RTI. Five specific research institutions, funded by the United States Office of Education (USOE) housed at the University of Kansas, University of Minnesota, Columbia University, University of Illinois and University of Virginia focused on different aspects of learning disabilities. Most of these research institutions focused on empirical research and developing educational methods for working with students with LD as a strategic, information-processing problem. As a result, some educational interventions, curriculum-based assessments, and instructional strategies were developed and tested. Deno and colleagues from the University of Minnesota created curriculum-based assessments to provide data to determine whether students were making academic progress. Researchers from the Kansas Institute worked on interventions for adolescents while the Virginia Institute focused on children with learning disabilities who also had attention problems (Hallahan, 2001). One other important study that resulted in a program called Direct Instruction by Sigfried Engelmann, Wesley Becker, and their colleagues "emphasized the systemic teaching of language subskills and the integration of these subskills into broader language competence." (as cited in Hallahan & Mercer, 2001, p. 23)

In a 1982 report, *Placing Children in Special Education: A Strategy for Equity,* was edited by Heller, Holtzman, and Messick (1982) for the National Research Council provided two key themes regarding the improvement of education for children. First, they assert that the validity of assessment must be relevant and useful to instruction. To fully understand a child's learning problems, the attributes of the child and the educational process need to be assessed, including the assessment of a student's learning environment for well-established instructional strategies as well as cognitive testing. Getting to know each individual is as important as understanding the learning environment (p. xi).

Second, the report contended that there should be evidence of quality instruction and classroom management within the learning environment. This data would be accrued through observation or evidence that other students are progressing well in the classroom before any determination of special need was established. Also, they recommended frequent assessments, analysis of the data, and interventions to remediate the lack of skill for at-risk students (Keller, Holtzman, & Messick, 1982).

By the end of the twentieth century, the number of students identified as LD skyrocketed from approximately 2% to 6% (Lyon et al., 2001). Eighty percent of those students were reading disabled (U.S. Dept. of Education, 2002). The education of a child with disabilities costs on average two to three times that of a regular student. School districts began having financial difficulties (Fuchs and Fuchs, 2006; Lyon, 1996). According to the National Center for Education Statistics (2015):

> The number of children and youth ages 3–21 receiving special education services was 6.4 million, or about 13 percent of all public school students, in 2012–13. Some 35 percent of students receiving special education services had specific learning disabilities. (Children and Youth with Disabilities, para. 2)

Although most school districts continued to use the discrepancy model to identify students with learning disabilities, there was growing evidence against it. Problems with the intelligence tests and the discrepancy formula included (a) researchers had not been able to show strong student outcomes using the ability-achievement discrepancy model (Gresham, Watson, & Skinner, 2001; Fuchs & Fuchs, 1998; (b) the original studies using IQ discrepancy methods could not be replicated, leading researchers to believe the original studies were flawed; (c) the intelligence protocols did not provide instructional implications (Case, Speece & Malloy, 2003, Hallahan & Mercer, 2001); (d) researchers were often unable to identify differences between low-achievers and those labeled LD (Fuchs, Fuchs & Speece, 2002; Lyon, Fletcher, Shaywitz, Shaywitz, Torgensen, Wood, Schulte, & Olsen, 2001; Case, Speece & Malloy, 2003); (e) children were not identified early enough to provide remedial services when problems first occurred. This inability to diagnosis caused a lack of treatment often until the third or fourth grade or what is called a "wait to fail model" (Hallahan & Mercer, 2001, p. 28; Lyon, et al. 2001, p. 266; Case, Speece & Malloy, 2003). Also, the discrepancy formulas differed from state to state, often causing a child to lose services after moving, and an overall reliable, measurable

discrepancy between IQ and achievement was impossible to achieve (Lyon et al., 2001). Lyon et al. (2001) stated the ability-achievement discrepancy model this way:

> The use of IQ discrepancy to identify children with LD appears to move many students further away from the education they need. Because the discrepancy hinges on the IQ level of students, rather than on their specific academic needs, the emphasis is on eligibility rather than instruction. (p. 267)

Federal special education mandates of Public Law 94-142, The Education of All Handicapped Children Act defined special education as "Individualized instruction, at no cost to the parents or guardians, to meet the unique needs of a child with a disability" (P.L. 94-142 ,1975). Embedded in special education law has been a precursor that before consideration for special education, every child was provided appropriate instruction in a regular classroom setting. (United States Department of Education, IDEA Federal Register, 2006, Section 300.309b, p. 5) Therefore, assessing a child's needs and providing an instructional program that fits his / her need is at the core of special education (Jimerson, Burns & VanDerHeyden, 2007).

According to Lyon, the prevalence of special education referrals was due to several social and political factors. First, due to inexperienced or inadequate teachers who did not have the knowledge or ability to differentiate instruction, some students continually fell behind academically and eventually were referred to special education services (1996). Second, although researchers identified best practice for learning to read, there were few preventive programs in place (Lyon, 1996). His team found that in

1976-77, the first year of full implementation of Public Law, 92-142, 1.8 percent of the entire school population was serviced in programs for learning disabilities, and in the 1992-93 school year that number had risen to 5.4 percent (Lyon, 1996).

Furthermore, students who did not meet eligibility requirements to obtain special education services were provided no support programs or services. Classroom teachers were expected to provide adequate instruction to all students in their classroom, even those who had academic and behavior struggles. As a result, problem solving teams evolved in schools in the 1970's with the goal to provide assistance to teachers who were experiencing academic, instructional and behavior difficulties with some of their students. The goal of the problem-solving team was to identify the student's problem, provide teachers with immediate strategies and resources in which to support them in the general education environment, while further evaluating the outcome (Gresham, 2007; Brown-Chidsey & Steege, 2010).

In the 1910 book, *How We Think*, John Dewey first conceptualized a five-step problem solving method that founded on logical and reflective thinking (Hermanowicz, 1961). Dewey believed that logical thought, or reflective thought, began when a problem or difficulty arose. The individual (or group) naturally develop a hypothesis in response to the problem and tests his or her hypothesis through experimentation (Dewey, 1910). His methods involved: observation of a difficulty or a problem, analysis of the problem, determine the goals to solve the problem, analyze the solutions, and finally, determine the best option to try (Hermanowicz, 1961). He concluded that problem solving was the

process of reflective thinking through inductive and deductive reasoning. Dewey's problem solving method affected many realms of educational research, the scientific method and working with difficult to teach students (Dewey, 1910).

Development of Problem Solving Teams

The key components of RTI have evolved in our nation's schools; however, not until recently have these components and procedures been organized into a comprehensive multi-tiered system to support struggling students (Kovaleski, 2007a). Following is a discussion of the earlier problem solving team efforts to help struggling students in general education which served as the keystone of current RTI models.

Deno and Mirkin (1977) developed a problem-solving assessment called Data-Based Program Modification (DBPM) that closely resembles Dewey's model of reflective thinking. DBPM was a systematic and decision-making method that was initially designed to determine the effectiveness of special education interventions for individual students. Components include monitoring the student's progress through data collection and altering the program to fit the needs of the individual. DBPM was the first program to evaluate the success of the interventions through frequent measurement of student performance and growth. According to the Deno, this program modification system evolved into the foundation of RTI as a progress monitoring tool for all students (Deno & Mirkin, 1977, Deno, 2003).

Problem solving method first described as a behavioral consultation model by Bergan (1977), and later by revised and updated by Bergan and Kratochwill (1990).

Bergan and Kratochwill's behavioral consultation model involved a consultant who partnered with the teacher to help in the four-step process of; problem identification, problem analysis, plan implementation and problem evaluation (1990). Research has documented the effectiveness of behavioral consultation as a vehicle for delivering interventions to students with a wide variety of learning and behavioral problems (Fuchs, Mock, Morgan & Young, 2003).

In addition to the behavioral consultation model, another type of pre-referral intervention became popular in the mid-1980s to help reduce the number of special education referrals. This new intervention that focused more attention on interpersonal relationships became known as a collaborative consultation model (Fuchs, Mock, Morgan & Young, 2003, p. 160-161). According to Fuchs et al., a team of colleagues consulted with the teacher, then collaboratively modified instruction, or some other aspect of the learning environment, to better accommodate a difficult-to-teach student before a formal referral of the student for testing and special education placement (2003, p. 160).

One of the most popular collaborative consultation problem-solving models that many schools embraced was the Teacher Assistant Team (TAT) developed by Chalfant and colleagues (Fuchs et al., 2003; Chalfant, Pysh, & Moultrie, 1979). The TAT provided valuable administrative and social support to isolated teachers with students who were difficult to teach, helped to keep students mainstreamed in the classroom, and provided classroom-based strategies to help at-risk students (Kruger & Struzziero 1995; Burns, Vanderwood & Ruby, 2005; Telzrow, McNamara, & Hollinger, 2000; Gresham,

2007; Fuchs, Mock, Morgan & Young, 2003). Even though most of the teams were not trained in classroom-based strategies or the problem solving or group processes, the TAT became a pre-referral process for special education (Kruger & Struzziero, 1995; Bahr & Kovaleski, 2006, Fuchs et al., 2003).

Some districts combined both behavioral and collaborative consultation models, in what became known as "The Collaborative -Solving Team" (Fuchs et al., 2003, p. 161). They were valuable for several reasons. First, they were composed of consultants (specialists) as well as teachers. Second, all members were trained in both the four-stage problem solving process and interpersonal relations. Third, the process was an efficient way of delivering pre-referral interventions to teachers. Fourth, the popularity of these teams represented needed change in education (i.e. collaboration, bottom-up decision making, and egalitarianism) (Fuchs & Fuchs, 1996; Kovaleski, 2002). In a phone survey conducted in the years 1999 and 2000 of 200 randomly selected State of Education Departments and Washington, D.C., (four per state), 85 percent of all schools had Pre-referral Intervention Teams (PIT) or similar teams. Each team frequently comprised of an administrator, counselor, school psychologist, and other school personnel, who recommended additional services, testing, and easy classroom interventions (Truscott, Cohen, Sams, Sanborn, & Frank, 2005; Johnson, Mellard, Fuchs, & McKnight, 2006). While these teams had many labels such as Teacher Assistance Team, School-based Intervention Team, Multidisciplinary Team, or Building Assistance Team, they employed similar problem solving processes and functions, supporting one student at a time.

Advocates of these collaborative consultation models believe it has enhanced academic and behavioral outcomes for students (Telzrow, McNamara, & Hollinger, 2000; Fuchs et al., 2003).

Many schools had a type of problem-solving team, which developed general education plans for targeted students at risk of failure, whether due to behavior or learning difficulties. The problem-solving approach included the team working through a complex four stage process of problem identification, problem analysis, plan implementation and plan evaluation (Fuchs, Mock, Morgan & Young, 2003). Interventions were carefully selected based on student data and individual need. The student was frequently monitored to ensure; the intervention was moving the student toward the desired outcome, if a revision in the intervention was warranted, or if the student needed further evaluation. According to Fuchs et al., some limitations to the problem-solving approach included a lack of a strong core curriculum that resulted in positive student outcomes (2003, p. 139).

Response to Intervention

While the RTI model has similar core characteristics, it takes on a scientific approach to problem solving. To improve student outcomes, a RTI model evaluates the core (Tier 1) academic and behavior program through universal screening (Kovaleski, 2007; Fuchs et. al, 2003). Interventions provided to at-risk students are based on scientifically validated principles of effective curriculum and instruction. There is a system of frequent monitoring of student progress with curriculum-based measurement

(CBM) to assist in problem identification and effectiveness of the intervention. Classroom teachers embed evidence-based interventions and instruction in daily routines and collaboratively work with others in the educational system to impact the learning of students (Kovaleski, 2003; Mellard & Johnson, 2008; Kovaleski & Glew, 2006; Brown-Chidsey & Steege, 2010; Fuchs, et. al, 2003).

According to Vaughn and Fuchs (2003), this problem-solving approach tracked the progress of every student, identified those at-risk, and provided appropriate pedagogy and curricular interventions to remediate them. Each at-risk student's success was monitored using curriculum-based measurements (CBMs) and those who did not make gains were recommended to special education for further testing (p. 138). Vaughn and Fuchs provide evidence of success using CBM assessments which helped to identify each student's deficit and ideas for the teacher on remediation in the regular education classroom (Vaughn & Fuchs, 2003, p. 139). In conclusion, they state:

> A response-to-instruction model could yield several promising benefits: (1) identification of students using a risk rather than a deficit model, (2) early identification and instruction of students with LD, (3) reduction of identification bias, and (4) a strong focus on student outcomes. (p. 140)

In their research, Lyon et al. argued that children are entering kindergarten with a deficit in phonological processing skills and failure to learn to apply phonemic and phonic skills often begin a process of reading difficulties, which can lead to more delays in vocabulary, fluency, and comprehension. In other words, children who begin school

lagging behind often never catch up without remediation, resulting in an LD diagnosis later in their educational career. Lyon et al., (2001) further contend that teachers lack the skill set to address and respond to each child's individual differences (p. 269), yet believe the only way to solve this issue is to provide immediate interventions when a deficit is first recognized.

Further, authors Fuchs & Fuchs (2006) state in *Introduction to Response to Intervention: What, Why, and How Valid is it?*, "that the IQ-achievement discrepancy approach fails to distinguish a qualitatively different and more deserving group of students from a much larger group of low achievers." (p. 96). Studies also suggest that young, poor readers with and without an IQ-discrepancy perform similarly on many reading-related cognitive tasks (as cited by e.g., Fletcher et al., 1994; Foorman, Francis, & Fletcher, 1995; Stanovich & Siegel, 1994), and demonstrate phonological processing deficits that are correctable with appropriate instruction" (as cited e.g., Fletcher, 1995; Morris et al., 1998; Stanovich, 1999; Torgesen, Morgan, & Davis, 1992; Vellutino et al.,1996, p. 96).

There is a growing body of research to support the Response to Intervention process for student reading achievement. Lyon et al. (2001) report "that there is substantial evidence that early identification and intervention in kindergarten and Grade 1 may substantially reduce the number of children that might otherwise be eligible for special services" (p. 276). Researchers found that after the first year of RTI implementation, there was a decrease in the number of special education referrals and

placements (VanderHeyden et al., 2007). Likewise, the National Joint Committee on Learning Disabilities (2005) provided six potential benefits for the RTI process:

1. Students with learning disabilities are identified sooner than the IQ-discrepancy model.
2. There has been a reduction in the number of students referred for special education.
3. There has been a decrease in the overidentification of minority students.
4. The data is maximally relevant to student instruction.
5. There is a focus on student outcomes with increased accountability.
6. RTI promotes shared responsibility and collaboration. (p. 14)

As part of the IDEA National Assessment Implementation Study, Bradley et al., (2011) reports, by the year 2008-2009, all states but two had RTI task forces, commissions, or internal working groups and 70 percent of elementary schools reported using RTI for reading/language arts (p. 32-33).

Although RTI is supported in the nation's special education law to assist in the determination for special education services under the specific learning disabilities (SLD) categories, researchers and school psychologists have raised some concerns. Some researchers have argued that RTI is only a pre-referral system because it lacks the descriptive language it needs to identify students who may be SLD (Kavale, Kauffmann, Bachmeier, and LeFever, 2008).

Wanzek and Vaughn (2011) found that more students qualified for special education in grade 3-5 than K-2 (p.106). O'Connor et al., (2013) noted the process of RTI required increased flexibility of additional time, scheduling and allocating personnel and other resources for program success. Also, students who eventually qualify for special education services may face increased reading difficulties that are more challenging to remediate than those who were eligible before RTI. The authors recommend that special education teachers may need advanced training and strategies to meet the needs of those students (p. 106).

Brown-Chidsey (2010) report that some models use the word *intervention* in Response to Intervention to reflect the activities used to help students, while others use the word *instruction*. It is important to note that researchers use both terms; Response to Instruction and Response to Intervention synonymously. Others put the two together as RTII or RTI². However, the I in Response to Intervention refer to "instruction" of a more specific or intensive nature, so both terms are often interrelated (p. 3).

RTI Support from Key Groups

In a 1982 National Research Council study on RTI and special education, Heller, Holtzman & Messick, documented three criteria that when met, would identify students with a learning disability. This was further researched and implemented by Sharon Vaughn and Lynn Fuchs in 1995, as a way to "conceptualize and identify" students with learning disabilities (Vaughn & Fuchs, 2003, p. 138). The three conditions ensure the use of (1) a quality general education program that used scientifically-based materials and

methods in which students made adequate academic gains; (2) a special education program that would improve student outcomes; (3) an assessment process that was accurate and meaningful. (p. 138) These researchers agreed that this approach would be a valid method of identifying LD when all three criteria were met (Vaughn & Fuchs, 2003, p. 138).

In 2001, Congress authorized No Child Left Behind (NCLB) legislation, a historic effort by the federal government to overhaul education and to close the gap for disadvantaged students (U. S. Department of Education, 2015). While this legislation is essential to the development and adoption of RTI in schools, providing a comprehensive review or critical analysis of NCLB is outside the perimeters of this study. Following is this researchers attempt to address the specific components of NCLB legislation that are relevant to RTI. The areas of importance include; prevention and intervention, scientific-based research and evidenced-based practice and accountability.

Elementary and Secondary Education Act (ESEA, 1965) was President Johnson's "war on poverty," which sought to improve education for economically disadvantaged students. NCLB 2001was the seventh reauthorization of ESEA. This standards-based education reform was a strict movement in which schools focused on (1) accountability for every student's academic progress; (2) ensured high-quality teachers; (3) certified curricular programs founded on scientifically-based research; (4) evidenced-based practice that created an educational system that aligned with state learning regulations

(U.S. Department of Education, n.d.; Elementary and Secondary Educaton Act (ESEA), 1965; Mellard & Johnson, 2008).

NCLB stated that by 2014, all students must be proficient in reading and math, a goal that many critics believed was unattainable (Hursh, 2007). NCLB required states to adopt educational standards and conduct annual assessments in reading and mathematics from grades 3-8 and 10-12 which to evaluate and monitor the achievement of all students (Connor, Compton & O'Connor, 2014). Each state set the standard for school districts to improve on their rate of students performing in the proficient or above average areas academically in reading and mathematics each year (U. S. Department of Education, 2015).

Accountability is a strong component of NCLB. State education agencies provided detailed reports to the U. S. Department of Education on whether schools and/or districts failed to meet Adequate Yearly Progress (AYP). Districts or schools failing to meet the strict goals had to take corrective action. Failure the first and second year resulted in a reallocation of Title I revenue to teacher professional development, and the adoption of an improvement plan. Failure to make improvement gains in the three-year time frame placed the district in correction action, along with parental rights to move their child(ren) to a non-failing school. Sanctions and more intensive measures lead to additional failures to reach required goals in subsequent years (U.S. Department of Education, 2015).

On the heels of No Child Left Behind, President Bush established the President's Commission on Excellence in Special Education (PCESE). In 2002, the PCESE conferred with researchers and experts on the subject and presented a report called *A New Era: Revitalizing Special Education for Children and Their Families.* The Commission's report became significant because it was one of the first federally sponsored examinations of special education since the inception of IDEA in 1975 (Berdine, 2003; PCESE, 2002). The report outlined the findings and recommendations for improving the educational performance of the joint roles of general and special education in meeting the educational needs of all children (Berdine, 2003; PCESE, 2002; Brown-Chidsey & Steege, 2010 p. 7). The Commission endorsed Response to Intervention as a direct link between instruction and student outcomes (PCESE, 2002). They state:

> Children should not be identified for special education without documenting what methods have been used to facilitate the child's learning and adaptation to the general education classroom. The child's response to scientifically based interventions attempted in the context of general education should be evaluated with performance measures, such as pre- and post-administration of norm-referenced tests and progress monitoring. (p. 26)

The original special education law, The Education of All Handicapped Children Act (PL 94 142), had undergone several revisions since it was enacted in 1975 and became known as the Individual with Disabilities Education Act (IDEA) (IDEA, 2004). The PCESE agreed and made three broad recommendations:

1. Focus on results, not on process. IDEA must return to its educational mission: serving the needs of every child. While the law must retain the legal and procedural safeguards necessary to guarantee a "free appropriate public education" for children with disabilities, IDEA will only fulfill its intended purpose if it raises its expectations for students and becomes results-oriented—not driven by process, litigation, regulation and confrontation. (PCESE, 2002, p. 1)
2. A Model of Prevention and Intervention. Place more emphasis on early and accurate identification and intervention when students first show signs of trouble. (PCESE, 2002, p. 1)
3. Children placed in special education are general education children first. General education and special education share responsibilities for children with disabilities. They are not separable at any level—cost, instruction or even identification. (PCESE, 2002, p. 2)

Despite the many concerns, the IQ-discrepancy model was used by most school districts to identify learning disabled students. In 2004, Congress transformed the process of special education with the reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA, 2004). Effective July 1, 2005, school districts had the option to stop using the IQ-Achievement Discrepancy method. Instead, IDEA suggests that special education eligibility, primarily those of learning-disabled, be

determined by the child's response to scientific, research-based interventions (IDEA, 2004, p. 46706; Brown-Chidsey, 2010).

Rachel Brown-Chidsey (2010) defined scientifically based practice as "those instructional methods and pedagogical approaches that have been verified by numerous research studies" (p 16). Scientific research-based practice is referred to as evidenced-based practice in diagnosis, assessment or intervention (Kratochwill & Shernoff, 2004). In 2006, federal regulations 34 C.F.R. 300 & 301 also required that school districts provide evidence that ensured a child suspected of having a learning disability has received appropriate instruction by a highly qualified teacher; and that timely, data-based documentation of repeated formal assessments of student progress during instruction is collected (as cited in Wright, 2008, p. 12).

The field of special education has undergone and continues to experience reforms and initiatives (Gersten & Dimino, 2006). Looking back at the Nation at Risk report, it is evident that reforms and initiatives as the No Child Left Behind of 2001, Least Restrictive Environment, Free Appropriate Public Education Act, Reauthorization of Individuals With Disabilities Education Act of 2004, and American Recovery and Reinvestment Act of 2009 (ARRA) have all impacted special education. Many of these reforms and initiatives improved the educational experience of special education students by ensuring an education as much like their peers as possible.

Not only did these changes have a significant impact on the special education field, but they also had a dramatic effect on reading instruction. Recently, Response to

Intervention (RTI) has gained momentum as the "new" means for determining specific learning disabilities in school-age students. Another key element of the RTI model is its provision of early intervention when students first experience academic difficulties, including those who have a specific learning disability. In addition to the preventative and remedial services that RTI provides to at-risk students, RTI also provides the data required to screen students who may be in need of special education or related services (Wright, 2007).

In 2004, IDEA was revised to include many of the recommendations from the President's Commission and researchers in the field. In short, IDEA required schools to use research-based interventions in the process of assisting students with learning difficulties or determining eligibility for special education. As a result, many schools have implemented Response to Intervention (RTI) as a method of meeting the new requirements set by IDEA 2004 (U.S. Department of Special Education, 2007; IDEA, 2004). As cited in Gleason, 2013:

> The majority of literature on RTI, as well as recent proposals from the U.S. Department of Special Education and President Obama's A Blue Print for Educational Reform: The Reauthorization of the Elementary and Secondary Education Act (2010), support the view that RTI is a framework that will increase the effectiveness of instruction provided to all students. (p. 17)

Response to Intervention Defined

The goal of RTI is to identify specific and effective intervention strategies that produce high learning rates for most students by matching instruction directly to student need (Batsche et al., 2005). Jimerson, Burns, and VanDerHeyden define RTI as "a systemic use of assessment data to most efficiently allocate resources in order to enhance student learning for all students and to effectively identify those who are eligible for special education services" (2007, p 4). In a Brief, the National Center on Response to Intervention (NCRTI, 2010) provided the following definition of RTI:

> Response to Intervention integrates assessment and intervention within a multi-level prevention system to maximize student achievement and to reduce behavior problems. With RTI, schools identify students at risk for poor learning outcomes, monitor student progress, provide evidence-based interventions and adjust the intensity and nature of those interventions depending on a student's responsiveness, and identify students with learning disabilities. (p. 4)

Brown-Chidsey and Steege (2010) define RTI as "a systemic and data-based method for identifying, defining, and resolving students' academic and/or behavioral difficulties" (p.3). Also, "They contend that RTI methods are problem-solving activities that provide a context to engage in ongoing assessment and evaluation" (p. 8). According to International Reading Association (IRA), "RTI is a comprehensive, systemic approach to teaching and learning designed to address language and literacy problems for all students

through increasingly differentiated and intensified language and literacy assessment and instruction" (2010, p. 2).

All of these definitions refer to RTI as a fundamental process of school-wide reform to improve the achievement of at-risk students. The process empowered school stakeholders to share their problem-solving resources and to work together collaboratively to provide a high quality research-based general education. Assessment is embedded in all areas of RTI and plays a crucial role to determine what students know and are able to do. Further, RTI established an avenue for the early identification of at-risk students and to provide evidenced-based interventions for such students (Brown-Chidsey & Steege, 2010; Wright, 2007; Marston, Muyskens, Lau, & Canter, 2003; Torgesen, 2007; Fuchs, 2007). For this paper, RTI is "conceptualized as a comprehensive, systematic, problem-solving framework for improving instruction and student achievement and reforming the LD process" (Barnhardt, 2009). The emphasis will be on RTI as it relates to reading instruction and students who encounter difficulty in learning to read.

A successful RTI process is dependent on interim and formative assessments embedded in the three tiers. The assessment system at each tier provides educators and others, data in which to make informed decisions about which students need intervention, which tier of support is needed, and when students should move tiers (Fuchs & Fuchs, 2006; Johnson, Mellard, Fuchs, & McKnight, 2006; Mellard & Johnson, 2008). According to Linn and Gronlund, assessment data can be used to determine effective

instruction and to identify instructional goals and student needs, relevant instruction, and assess intended learning outcomes (2000). Many educators, schools and districts analyze the assessment data as a predictor of student performance on state-wide summative accountability assessments (Kovaleski, 2007a).

Types of assessment. Interim assessments, often referred to as benchmarks or universal screenings, are given at systemic intervals during an RTI process. The Council of Chief State School Officers (CCSSO), commissioned a study of interim assessments by Technical Issues in Large Scale Assessment State Collaborative on Assessment and Student Standards (TILSA SCASS) in 2008 with the goal "to better understand the assessments being used, their purposes and potential value" (p. 1). From the previous work of Perie, Marion and Gong (2007), TILSA SCASS (2008) revised the definition of interim assessments to be

> Assessments administered multiple times during a school year, usually outside of instruction, to evaluate students' knowledge and skills relative to a specific set of academic goals in order to inform policymaker or educator decisions at the student, classroom, school, or district level. The specific interim assessment designs are driven by the purposes and intended uses, but the results of any interim assessment must be reported in a manner allowing aggregation across students, occasions, or concepts. (p. 2-3)

In other words, "interim assessments are those that fall between large scale assessments and classroom formative assessments, for the formative purpose of influencing

curriculum and instruction, while at the same time measuring and documenting each student's growth as a summative measure" (TILSA SCASS, 2008, p. 3). When administrated correctly, interim or universal screening assessments provide a school or district relevant information in which to diagnosis student placement for remediation, predict outcomes on state summative accountability tests, and evaluate student understanding of completed studies (TILSA SCASS, 2008, p. 6).

Summative assessments are given one time at the end of a unit of study, the semester, or school year to evaluate students' performance against a defined set of content standards. These assessments are widely known as statewide tests (but can be national or district) and are often used as part of an accountability program or to inform policy. They are the least flexible of the assessments.

According to Popham, (2008), formative assessment is defined as "a planned process in which assessment-elicited evidence of students' status is used by teachers to adjust their ongoing instructional procedures or by students to adjust their current learning tactics" (p. 6). Burns (2010b) define formative assessments as "instructionally relevant protocols that provide immediate information that can assist in obtaining a diagnosis, course of action or feedback to improve the quality of an educational experience" (p. 22). Moss and Brookhart (2009) believe that formative assessment is a means in which to change instruction not to merely audit it (p. 6). For this study, formative assessment will be defined as, a process used by teachers and students in the classroom and during instruction that provides feedback to adjust ongoing teaching and

learning to improve students' achievement of intended instructional outcomes (Formative Assessment for Students and Teachers State Collaborative in Assessment and Student Standards [FAST SCASS], 2006, as cited in Popham, 2008, p. 5).

Most often formative assessment is embedded within the learning activity and linked directly to the current unit of instruction therefore delivered in a classroom setting. Providing corrective feedback, modifying instruction to improve the student's understanding, or indicating areas of further instruction are essential aspects of formative assessment (Moss & Brookhart, 2009; Popham, 2008). According to Black and William (1998), formative assessment does improve learning (p. 56). In the RTI process, formative assessments are systemically given to students during Tier 2 and Tier 3 to evaluate whether students are benefiting from interventions. This process, called progress monitoring, provides the on-going data on which to base decisions, whether to change an instructional program, or to move a student into a different tier (Mellard & Johnson, 2008). While this is a brief overview of assessments found in the RTI process, further discussion and details are provided in later sections of this paper.

Table 1. Three Types of Assessment

Formative	Interim	Summative	
Typical Use	feedback to adjust teaching and learning	ongoing monitoring student progress	student placement; school and district accountability
Frequency of Administration	continual; multiple times a day	generally two to six times per school year	usually once a school year
Scope of Administration	student and classroom	usually school or district	usually state

Note. Sourced from Interim assessment practices and avenues for state involvement by the TILSA SCASS Interim Assessment Subcommittee, 2008, p. 4.

RTI Models

Fuchs, Mock, Morgan, and Young (2003) characterized existing RTI models as either standard protocol or problem-solving based (p. 159). According to Fuchs and Fuchs (2006) the two RTI methods, problem-solving and standard treatment protocols, differ operationally causing different prevalent rates of reading disability and different subsets of nonresponsive children identified for special education. In other words, if an RTI model provided common interventions among all children who were not reading proficiently they were implementing a standard protocol approach. If a problem-solving team provided specific individualized interventions for students, they were implementing a problem-solving approach (Burns & Coolong-Chaffin, 2006). However, researcher Daniel Reschly (2003) presented both of these approaches within one model, which seemed to make conceptual sense in that both sought to improve student learning and could probably work best within a unified model (Christ, Burns, & Ysseldyke, 2005, p. 1-

2). Christ et al. (2005) go on to suggest that the "problem-solving" model be renamed "problem analysis" because both subtypes have common solving elements. Burns, Deno and Jimerson (2007) recommend a unified system called IDEAL. The steps of this system include (a) Identify the problem, (b) Define the problem, (c) Explore alternative solutions to the problem, (e) Apply a solution, (f) Look at the effects of the application (IDEAL; Bransford and Stein, 1984 as cited in Burns, Deno and Jimerson, 2007, p. 429). A brief summary of Problem-Solving, Standard Protocol, and IDEAL approaches follow.

Problem-solving approach. The problem-solving approach is a systematic analysis of individuals to determine skill deficits and provide specific interventions to target those deficits (Barnett, Daly, Jones, & Lentz, 2004). Therefore, no student characteristic (e.g., disability label) determines what the intervention will be. Instead, solutions are determined by analyzing preliminary data, and evaluating the student's responsiveness to interventions regarding a student's specific performance (Carney & Stiefel, 2008).

In a school, RTI as a problem solving method requires the team to analyze the problem correctly, select an empirically-based intervention, progress monitor effectively, and make team decisions regarding student progress and data (VanDerHeyden, Witt & Gilbertson, 2007). There may be several solutions to individual problems. Therefore, problem solving trial and error approaches rely on the careful collection of data on students' performance in response to treatment (Fuchs, Mock, Morgan & Young, 2003). Due to the individualization of the problem-solving model, researchers found desired

levels of fidelity were difficult to maintain in a school setting, and conclusions were not always reliable and consistent (Telzrow, et al., 2000).

Carney and Steifel (2008) conducted a study of the long-term outcomes of the RTI problem-solving method using the Instructional Support Team (IST) model. In a field based setting involving students needing behavioral supports, over the 3.5 year scope of this study, eight percent of the school population was referred for Tier 2 interventions. The IST problem solving method advocate that students should positively respond within 50 days of beginning an intervention. At the conclusion of this study, twenty percent of the students referred for behavioral supports were moved to an independent level while just fewer than forty percent were functioning independently of extra help. Many of the other participants were considered low risk but still in need of some additional support (p. 70-71). These false positives, or children who appear nonresponsive and physically challenged, prove that with intensive instruction they are not learning or behaviorally handicapped, are common with problem solving methods (Fuchs & Fuchs, 2006).

Standard protocol approach. The standard protocol approach to RTI requires the use of the same empirically validated treatment implemented to all students with similar academic problems (Fuchs, Mock, Morgan and Young, 2003). Lyons and others first advocated standard protocol as an early reading intervention approach (Lyon, Fletcher, Shaywitz, Shaywitz, Torgensen, et al., 2001), while other researchers applied this method in scientifically-based research (Vellutino, Scanlon, Sipay, Small, Pratt, et

al.,1996). Mellard & Johnson define standard protocols as "interventions that researchers have validated as effective in experimental studies" (2008, p. 85). Examples of standard protocol method in reading include the use of explicit instruction to support students with decoding skills (McMaster, Fuchs, Fuchs, & Compton, 2003; O'Connor, Fulmer, Harty, & Bell, 2005; Vaughn, Linan-Thompson, & Hickman, 2003), the use of strategy instruction to support students with comprehension challenges (Kamps & Greenwood, 2003) or the use of fluency development to encourage more fluent reading of text (Hasbrouk, 2006). Fuchs et al. (2003) cite five advantages of the standard protocol method:

1. A large number of students can participate in the treatment protocol.
2. The intervention's accuracy is easy to assess.
3. The accuracy of fidelity and ease of training staff to provide the intervention is cost effective.
4. There is no decision-making process on what intervention to implement.
5. It lends itself to group analysis where outcomes for students can be assessed against the "goal-line" criteria. (p. 166)

Carney and Stiefel (2008) cite several scientific research-based standard protocol interventions for reading remediation from the research literature. These interventions include the Auditory Discrimination in Depth (ADD) program (Lindamood & Lindamood, 1998), the Embedded Phonics (EP) program (Torgesen et al., 2001), and Reading Recovery (Ruhe & Moore, 2005; Schmitt & Gregory, 2005, p. 62).

When comparing the two models, the standard protocol approach seems more likely to facilitate greater quality control than the problem-solving model. According to Fuchs and Fuchs (2006), the standard treatment protocol seems to identify "true positives" or children truly in need of special education services. At the same time, the standard protocol method identifies "false negatives, those children who are responsive in a multi-level model but when returned to regular classroom instruction they once again demonstrate the same learning problems as before (p. 97). The problem solving model is less intensive and systematic, yet appears to be more sensitive to individual student need. Both the problem solving model and the standard protocol model identify students who are "true positives" or have a disability and in need of special services. At the same time both methods identify students who are unsuccessful in the mainstream classroom but with interventions prove they are not physically challenged (Fuchs & Fuchs, 2006).

IDEAL: a unified RTI approach. Bransford and Stein (1993) developed an IDEAL five basic steps to problem solving models include:

1. **I**dentify the problem by providing universal screening for all students on a continuous basis.
2. **D**efine the problem by examining criterion-referenced and norm-referenced screening and progress monitoring data, as well as classroom data to determine the magnitude of the discrepancy, and the source of the condition (Burns et al., 2007). The assessment data collected will provide analysis of educational

environment, instructional placement, and instructional modifications to discrepancies. (Deno, 2003)

3. **E**xplore alternative solutions to generate and explore potential remediations. The data may suggest small group interventions often found in the standard protocol approach or a more individualized problem-solving approach. Research is unclear as to the role of the PST or who would serve on this team (Burns et al., 2007). However, the specific intervention would be based on the individual child's need, which would be assessed and grouped according to the severity of the problem. (Burns et al. 2007)
4. **A**pply the chosen solution for the recommended period. Some models may include special education services as part of the delivery model (Tilly, 2002). Other models may use remediation in the general education classroom as individual, small-group, or class-wide interventions (Kovaleski, Tucker, & Stevens, 1996). However, analyze the data to determine whether individual children would benefit from small groups standardized interventions or require more individualized interventions as found in the problem-solving model. (Burns et al., 2007)
5. **L**ook at the effects. Outcome assessment of Tier 2 occur at least monthly and should address the same general outcome measure used in Tier 1. Consistency between Tiers is important because of the level of student skill and slope of growth (dual discrepancy) should both be compared with the general population

(Burns et al., 2007). Provide supplemental interventions to the student(s) in small groups of two to five students for thirty to sixty minutes a day. The required duration of each intervention is three or four days a week for ten to fifteen weeks. Spplemental remediation is repeated as needed. (Johnson et al., 2006; Burns et al., 2007; Fuchs & Fuchs, 2001)

Structure of RTI

While there is more than one model of Response to Intervention, there is a consensus regarding core features and activities that comprise each tier of the structure. The National Center on Response to Intervention (NCRTI) identified the "four essential components of RTI framework in each tier as screening, progress monitoring, multi-level or multi-tier prevention system, and data-based decision making" (2010, p. 1).

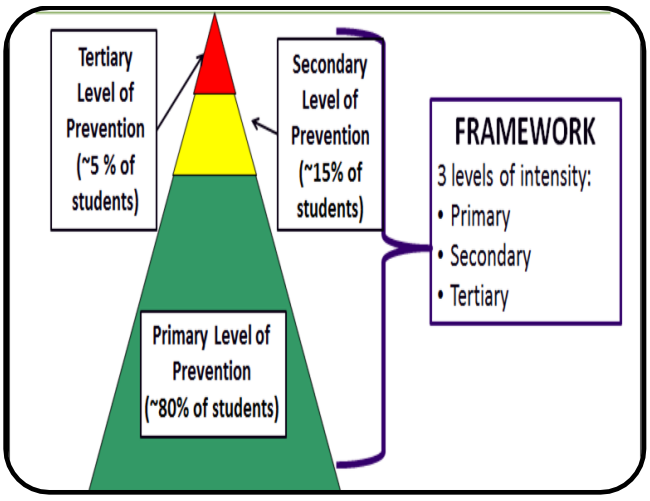

Figure 1. NCRTI's RTI Intervention Levels. This figure illustrates the NCRTI's three-tier example of Response to Intervention and the recommended percentage of students that on average benefit from that level of instruction (NCRTI, 2012).

The green primary level or first tier depicts the high-quality general education core curriculum and instruction received by all students. The majority of students are predicted to be successful at this level. The yellow secondary-level or Tier 2 depicts the students who need more research-based intensive small group instruction to benefit from the core or first tier core-curriculum. Researchers predict that approximately ten to fifteen percent of students in a school may fit into this category. As a supplement, general education teachers provide differentiated, focused, small group instruction for 20-40 minutes, several times each week. Teams monitor student progress by analyzing interim, formative, and summative assessments. Finally, students who do not make adequate gains academically or behaviorally after the allotted time, move to the red tertiary level or Tier 3. They require the most intensive small group or 1:1 instruction by specialized members of the staff. Approximately five percent of students may fit into this category (NCRTI, 2012).

Tier 1

Tier 1 is the first level of prevention (Jones, Yssel, & Grant, 2012; National Association of State Directors of Special Education [NASDSE, 2006]; Berkeley, Bender, Peaster, & Saunders, 2009). The focus of this level is for all children to receive a high-quality, developmentally appropriate foundation of curriculum and instruction. All within a tightly structured general education classroom (Mellard & Johnson, 2008; NASDSE, 2006; Fuchs & Fuchs, 2006; IRA, 2010) that ensures accurate and consistent instructional delivery through measures of fidelity (Johnson et al., 2006). In how-to

manuals, research articles, and national and state documents on RTI, scientifically-validated practice is referred to as a critical component of the RTI process and endorsed by numerous research studies embedded in Tier 1 (Brown-Chidsey, & Steege, 2010).

High-quality research-based core curriculum and instruction. According to many researchers, scientifically validated or evidenced-based curriculum and instruction meet the needs of seventy-five to eighty percent of regular education students (Brown-Chidsey & Steege, 2010; IRA, 2010; Jones, Yssel & Grant, 2012; NASDSE, 2006; Burns, 2010a). In 2001, NCLB mandated the use of "scientifically-validated instruction" and defined scientifically based research as "research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to educational activities and programs." (P.L. 107-110) The language of the Individuals with Disabilities Education Improvement Act of 2004 (P.L. 108-446) (IDEA, 2004) indicate that students must receive appropriate learning experiences before a disability can be considered as a basis for achievement or behavioral difficulties. High-quality curriculum and instruction ensure that student difficulties are not attributed to an ineffective teacher (Johnson, Mellard, Fuchs & McKnight, 2006). In a RTI model, a Tier 1 system of scientifically-based strategies and interventions, must be available to all educators, along with a strategic plan of interventions for students who do not respond (Daly, Martens, Barnett, Witt, & Olsen, 2007).

Differentiated instruction. Classroom teachers need to be empowered to provide an environment for learning that includes a variance of flexible grouping, differentiation,

and instruction based on data and observations (Kosanovich, Ladinsky, Nelson, & Torgensen, 2007; Jones, Yssel, & Grant, 2012). Differentiation is responsive instruction that incorporates a variety of strategies designed to meet the unique needs of each student (Watts-Taffe et al., 2012; Tomlinson, 2003). According to Tomlinson (2003), classroom teachers that differentiate instruction often modify their content, process, product and/or learning environment, based on individual student readiness, to ensure maximum learning occurs (p. 6). According to researchers, in a RTI framework, a differentiated and responsive instruction is an essential prerequisite to referring a child for special educational services (Watts-Taffe et al., 2012, Tomlinson, 2003). A benefit of Tier 1 is the safety net it provides to catch students who are at risk academically and need the services of additional support.

RTI works within the context of general education to provide needed support to students by causing a shift in many educational systems where the classroom teacher is responsible for screening, identification, and the first tier of intervention (Johnson et al., 2006). Such a system requires an integrated approach to service delivery that includes "leadership, collaborative planning, and implementation by professionals *across* the education system" (NASDSE, 2006, p. 3). Kratochwill, Volpiansky, Clements & Ball, (2007) cite a sample study of 104 school districts across 12 states, all of which adopted research-based programs but showed low implementation rates due to lack of teacher training, required lessons in teaching strategies and failure to develop age-appropriate instructional procedures (p. 621).

High-quality reading instruction. Providing high-quality reading instruction to all students is considered a primary intervention and should produce proficient readers. Scientific research has provided an abundance of knowledge that can be used to inform reading instruction (National Reading Panel, 2003; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004; Kim, Apel, & Al Otaiba, 2013; Torgeson, 2002). Studies of reading instruction have consistently shown that nearly all students can be taught to read (Mathes, Denton, Fletcher, Anthony, Francis, & Schatschneider, 2005; Jenkins, & O'Connor, 2002; Torgeson, 2002). After analyzing research on the attributes of effective instruction for students at-risk of reading problems, Denton, Vaughn and Fletcher (2003), found five areas must be addressed to enhance the reading development for all students. They are (1) an efficient and knowledgeable teacher; (2) integration of the main instructional components; (3) differentiated instruction for students with reading difficulties; (4) explicitness of instruction; and (5) bridging the gap between research and practice (p. 202). Although a complete discussion of this empirical research is beyond the scope of this review, a brief summary of recent conclusions about effective reading instruction follows.

A document titled, *Put Reading First: The Research Building Blocks for Teaching Children to Read* was published by the Partnership for Reading (National Institute for Literacy, 2003, 2nd ed.). The purpose of this collaborative research effort of the National Institute for Literacy, the National Institute of Child Health and Human Development, the U.S. Department of Education, and the U.S. Department of Health and Human Services

was to make scientifically based research available to school districts, administrators, educators, parents, and others with an interest in helping all people learn to read (p. i). Five components successful reading instruction came to the fore front from this research. They are:

1. Phonemic Awareness Instruction. Phonemic awareness is the ability to notice, think about, and work with the individual sounds in spoken words. (p. 2)
2. Phonics Instruction. Phonics instruction teaches children the relationship between the letters of written language and the individual sounds of spoken language. (p. 12)
3. Fluency Instruction. Fluency instruction is the ability to read a text accurately and quickly with expression; recognize words automatically and provide support to sound out unknown words, and comprehend what is read. (p. 22)
4. Vocabulary Instruction. Vocabulary instruction refers to the words we know and recognize to communicate effectively in oral and written works. (p. 34)
5. Text Comprehension Instruction. Comprehension is when a reader can recognize and read words with understanding with purpose and actively think about the meaning of the text while they read. (National Institute for Literacy, 2003, p. 48)

In 2014, a report called "Improving Reading Outcomes for Students with or at Risk for Reading Disabilities: A Synthesis of the Contributions from the Institute of Education Sciences Research Centers" was published (Connor, Alberto, Compton &

O'Connor). This research was funded by the Institute of Education Sciences (IES) National Center for Education Research and National Center for Special Education Research and published in peer-reviewed outlets through December, 2011 (p. 3). The focus of this report described what has been learned regarding reading instruction and the improvement of reading outcomes for children with or at risk for reading disabilities (p. 28). This synthesis of information provided scientific evidence about effective reading instruction from peer-reviewed articles and chapters, projects and grants initially awarded from 2002 through 2008 through the National Center for Education Research and the National Center for Special Education Research (p. 11). Based on their findings, teachers of preschool and kindergarten students should base instruction on oral language, and vocabulary development to aid comprehension. Kindergarten and first-grade educators teach the reading skills that lead to decoding and word recognition (phonemic awareness, letter-sound knowledge, and quick recognition of high-frequency words). Beginning in second grade and through the remaining elementary school years, teacher's instructional focus is on refining students reading rate and accuracy, vocabulary, and reading comprehension. During the elementary years, students gradually increase in reading ability related to fluency and comprehension, and by middle and high school years, students are expected to use reading as a tool for learning, finding, and using information (Connor et al., 2014).

In an article titled, "The Prevention of Reading Difficulties" (2002), Torgesen recommends early reading instruction of kindergarten through grade 3 be explicitly and

systematically balanced between word-level and reading comprehension skills. Next, procedures must be in place to identify students at-risk of reading difficulties early in their school career. Finally, at-risk students must have a high-quality reading instruction that is explicit, intensive and supportive that includes (1) critical components of phonemic awareness and decoding skills; (2) fluency in word recognition and processing; (3) reading comprehension strategies; (4) oral language vocabulary; (5) spelling and writing skills (p. 14). The author concluded, "There are numerous examples that when working effectively in all these areas will dramatically increase our success in teaching all children to read well during elementary school" (p. 17).

These studies, research, and peer, reviewed articles provide evidence that when reading instruction is systematic, explicit, evidenced-based, and taught by highly-qualified teachers, positive student outcomes result. Recently the U.S. Department of Education's Institute of Education Sciences supports three websites that assist school districts, schools, educators, and others that are interested in identifying research-based educational supports and studies related to the effect of curriculum and instruction. They are:

1. "What Work's Clearinghouse" (WWC) has been a central and trusted source of scientific evidence for what works to improve student outcomes in education since 2002. (retrieved from http://ies.ed.gov/ncee/wwc/aboutus .aspx)

2. "Florida Center for Reading Research" (FCRR) has been available since 2002. The center provides a) Support for identifying key language, cognitive, and regulatory components necessary for strong comprehension of spoken and printed language. b) Developing, refining, and evaluating instructional activities that teachers can effectively use to increase the language and reading comprehension skills of children in pre-K through 5th grade. c) The development and maintenance of an on-line professional development system that will support teachers' use of instructional activities. (retrieved from http://www.fcrr.org/about-fcrr/projects.asp)
3. "The University of Oregon Center on Teaching and Learning" (CTL) was established in 2000 to conducting rigorous research on the design, delivery and efficacy of curriculum, instruction, and assessment as individual elements used in schools, especially in the primary, elementary, and middle school grades. CTL's mission is to conduct, translate, and disseminate research that focuses on the solutions to serious but practical problems in school systems. (retrieved from https://ctl.uoregon.edu/about/ history)

Universal screening. The RTI framework begins with a critical component of establishing a baseline, often obtained using general outcome measures. This interim assessment, often referred to as universal screening (Mellard & Johnson, 2008; Brown-Chidsey & Steege, 2010; NASDSE, 2006; Fuchs & Fuchs, 2006; NCRTI, 2010) focuses on essential academic areas to determine both group and individual's current level of

proficiency on specific skills (Fuchs & Fuchs, 2001; NASDSE, 2006; Mellard & Johnson, 2008; Brown-Chidsey & Steege; NCRTI, 2010). In 2010, the National Center on Response to Intervention [NCRTI] defined universal screening as:

> ...brief assessments that are valid, reliable, and demonstrate diagnostic accuracy for predicting which students will develop learning or behavioral problems. Screenings are conducted to identify those who are at risk of academic failure and, therefore, need more intensive intervention to supplement primary prevention. (i.e., the core curriculum p. 8)

Researchers believe the analysis of system-wide interim data at Tier 1 provides three functions. First, it provides evidence of the high-quality curriculum and instruction over a larger population such as at a school, district or state level. If the majority of students in general education is making adequate academic progress, it can be inferred that the instruction is effective. Second, universal screening leads to data-based decision making in the curricular and instructional program for students who are not making adequate academic progress and who need more support from Tier 2 to help them be successful if repeated several times a year (Hughes & Dexter, 2008; Brown-Chidsey & Steege, 2010; NASDSE, 2006; Mellard & Johnson, 2008; Buffum et al., 2009; Jimerson et al., 2007; Vaughn and Fuchs, 2012; NCRTI, 2010). Third, if implemented on a regular basis across grade levels, interim data will help to prevent two errors that emerge on preliterate predictive screenings. One type is false negatives or students who score above a cut score on a screening tool and later on exhibit reading difficulty. The second type of

error is called false positives or students who score poorly on a predictive instrument but later become good readers. Children identified as false positives increase the number of Tier 2 intervention students and stress school resources (NASDSE, 2005; Hughes & Dexter, 2008; Compton, Fuchs, Fuchs & Bryant, 2006).

Predicting which preliterate children will be at risk of proficiency in reading is difficult (Jenkins, Hudson & Johnson, 2007; Torgesen, 1998; Connor et al., 2014). Studies show the proportion rate for false positives with preliterate children is between twenty to sixty percent, with an average around forty-five percent, (Torgesen, 1998) while false negative errors range from ten to fifty percent and average around twenty-two percent for this same group (Torgesen, 1998; Torgesen, 2002). Torgesen (1998) states, "prediction accuracy increases significantly the longer a child is in school" (p. 4). To implement a successful RTI model, D. Fuchs and Deschler, (2007) state that a process must be in place to identify and manage true positives and negatives (p. 2). Although a complete discussion of this empirical research is beyond the scope of this review, a brief summary of factors that may challenge the accuracy of student outcomes on interim universal screening and progress monitoring measures follows.

Elements of effective universal interim screening measures. The main purpose of an interim screening instrument is to identify students who have not required critical skills and are in need of further investigation. There are four criteria to a good screening tool; sensitivity, specificity, practicality, and consequential validity (Hughes, & Decker, 2008; Jenkins and Johnson, 2007).

A screening measure's accuracy is determined by two statistics, sensitivity and specificity. According to Jenkins and Johnson, (2007), sensitivity is the screen's accuracy to identify individuals who later fail the outcome measure. It is calculated by dividing the number of true positives by the total number of students who fail the later reading assessment. Specificity, on the other hand, determines the screen's ability to identify the individuals who will pass the criterion measure. Specificity is calculated by dividing the number of true negatives by the total number of individuals who perform successfully on the outcome measure (p. 3). The accuracy of a screening tool is judged primarily by its correctness of these two measures (Johnson et al., 2006). Although no screening tool will provide 100 percent accuracy, "the goal is to have very few false negatives by using instruments that yield true-positive rates approaching 100 percent" (Johnson et al., 2006).

Researchers determine the sensitivity and specificity of a measure by adjusting the cut score of a tool. According to Johnson et al. (2006), "Cut scores, also called cut points, represents the dividing line of students who are not at risk and those who are potentially at risk" (p.14). Schools and districts must select its screening cut point based on the criteria with the specific criterion measure it will use to determine whether a student is reading accurately (Jenkins & Johnson, 2008). Mellard et al. (2008) remind us that adjusting the cut scores may alter the sensitivity and specificity of a particular assessment (p. 27). This is information is valuable to a RTI model because false

positives are costly and lead to students getting instruction in which they do not need (Hughes, & Dexter, 2008).

Next, interim assessments must be practical, easy to use, inexpensive, and quick to use while maintaining reliability. Practicality refers to the short and simplicity of implementation of the screening tool that is given to all students (Mellard et al., 2008, p. 27). The screening tool must be simple enough to be given on a wide scale by an average person (Hughes, & Dexter, 2008). In other words, a classroom teacher or tester can adequately give the instrument to her class of students in the regular environment. Third, it must have consequential validity, meaning those students who are determined to be at risk receive timely interventions that help to remediate any weakness or disability and it does no harm to students (p. 3). For example, the tool should avoid inequitable treatment, and the assessments should be linked to effective interventions (Hughes, & Decker, 2008; Jenkins and Johnson, 2007).

Finally, the screening tool has classification accuracy. In other words, the screening tool identifies students who are not at risk for reading failure and do not develop reading problems. Further, it identifies students who are in danger of failing and later develop problems. Researchers call this the criterion measure (Hughes, & Decker, 2008; Jenkins and Johnson, 2007).

Screening measures can use either a criterion referenced or normative comparison standard of performance (National Reading Center of Learning Disabilities [NRCLD], 2006). When first identifying a formative universal screening tool it is important to

understand if the screening tool is criterion or norm-referenced, and what cut score distinguishes appropriate readability (Mellard & Johnson p. 25). A criterion measured assessment is designed to determine a student's proficiency against a fixed set of criteria or learning standards while a norm-referenced assessment ranks each student according to the general knowledge of like peers (NRCLD, 2006).

The purpose of criterion-referenced assessments is to find out how much students know before instruction begins and after it ends (Huitt, 1996). An example of a criterion reading measure would include a student reading a grade level reading passage with the goal of reading a specified number of words per minute (Mellard and Johnson, 2008). When using a criterion referenced measure, Hosp and Madyun (as cited in Jimerson et al., (2007), remind us that the screening tool must accurately measure what is being assessed (p. 175). In other words, the assessment must reflect the standards and skills taught to students. Also, Mellard et al., (2008) advised selecting a criterion measure with a strong predictive validity (p. 26), or evidence that the skill being assessed (in this case, fluency) will predict student's readability.

The purpose of a norm-referenced measure is to discriminate between high and low achievers (Huitt, 1996). In other words, if the student body assessed is in excess a minority group and the norm-referenced test are based on non-minority students, the student outcome may not be dependable. If the screening tool is norm-referenced, it is important to ensure adequate representation of students similar to the one being assessed (Hosp & Madyun, as cited in Jimerson et al. 2007, p. 176).

Criterion-referenced measures are the usually preferred screening measures because they provide more accurate data about performance on particular skills and standards (Mellard et al., 2008; Burns & Coolong-Chaffin, 2006; Good III et al., 2001; NRCLD, 2006). Schools should link screening measures to existing performance measures, including existing standards in the school curriculum.

Common universal screening measures. For over 3 decades, researchers and educators alike have used general outcome measures (GOM) such as Curricular-Based Measurement (CBM) to frequently measure foundational reading skills, predict future academic success, and to provide ongoing curricular goals to ensure the success (Deno, 2003; Good III, Simmons, & Kame'enui, 2001). In 2005, Stecker and Lembke defined CBM as "a scientifically validated form of student progress monitoring that incorporates standard methods for test development, administration, scoring and data utilization" (p. 1). It is no surprise that such a tool would be used in the RTI process.

Curriculum-Based Measurement (CBM) was first created in the 1970s at the University of Minnesota. CBM was developed as a special education initiative to provide special education teachers a way to evaluate instruction and improve their effectiveness (Deno, & Mirkin, 1977).

As the most well-known and most researched GOM, CBMs have cited many benefits. CBMs:

(1) Employ standardized administration and scoring methods that yields accurate and meaningful information about student performance and growth over time. (Fuchs, Fuchs, & Speece, 2002; Deno, 2003, p. 185)

(2) Are reliable, valid as well as easy to aggregate for accountability with widely used standardized assessments and state tests. (Deno, 2003, p. 185; Fuchs, Fuchs, & Speece, 2002)

(3) Are efficient, only taking 1-3 minutes in duration depending on the skill being measured and some can be administered in a group setting. (i.e. MAZE CBM) (Deno, 2003, p. 185; Marston, 1989; Fuchs & Vaughn, 2005)

(4) Are simplistic enough that schools can develop their own set of norms for CBM or schools can choose to use a variety of commercial products, such as the Dynamic Indicators of Basic Early Literacy (DIBELS), and AIMSweb. (Good & Kaminski, 2002; Shinn & Shinn, 2002a)

(5) Are norm referenced and criterion referenced that allows for the aggregation of data at the systems level to provide information that may be used to examine the effectiveness of the instructional supports within a classroom, school, or district to help determine when changes should be made. (Kaminski & Cummings, 2007; Deno, 2003, p. 185)

Despite having a rich scientifically-based curriculum and instruction, there is usually a percentage of pupils who need further supports to be successful. If less than 80% of students in the general education classroom are meeting benchmarks, school

personnel need to review the core curriculum (Mellard et al., 2008). Students identified at-risk need to be provided interventions in an appropriate timeframe thus linking assessments to intervention to remediate those who need it when they need it. In the RTI framework, those students who need more support would move to the level of Tier 2.

Tier 2

According to Burns and Ysseldyke (2005) no more than fifteen to twenty percent of students in a particular school should need additional support beyond that of Tier 1. A critical component of Tier 2 is identifying those who need remedial support and small group interventions (Fuchs & Fuchs, 2005; Brown-Chidsey & Steege, 2010; Compton et al, 2006; Fuchs & Vaughn, 2012; O'Connor et al., 2013; Mellard et al., 2008; Burns & Coolong-Chaffin, 2006). There are two reasons for Tier 2. The first is to prevent reading difficulty by providing intensive services to the students, due to their lack of growth within the specific time frame (McMaster & Wagner, 2007; Brown-Chidsey & Steege, 2010; Mellard & Johnson, 2008; Compton et al., 2006). The second reason for Tier 2 is to assess students' responsiveness to instructional intensity to determine if they will return to regular education or move into Tier 3 or special education due to persistent reading difficulties (Compton et al., 2006; O'Connor et al., 2013).

Students are provided intense Tier 2 interventions by their classroom teacher in small groups of two to eight students in similar areas of need and skill level (Shapiro, 2009; Mellard & Johnson, 2008). During intervention, teachers provide explicit instruction involving more teacher-student interaction, frequent opportunity for practice,

and comprehensible specific feedback. Instruction should be systematic, building skills gradually, in isolation and then integrating them with other skills usually for three times each week for twenty to forty minutes a day (Gersten, Compton, Dimino, Santoro, et al., 2008). Student progress is monitored weekly, bi-weekly, or monthly for nine to twelve-week duration (Mellard et al., 2008).

Progress monitoring in tier 2 and beyond. Johnson, Mellard, Fuchs, and McKnight (2006) define progress monitoring as a "set of assessment procedures for determining the extent to which students are benefiting from classroom instruction, and for evaluating the effectiveness of the curriculum" (p. 2.1). In the context of the RTI model, students not responding adequately to Tier 1 instruction are provided alternate interventions. There are two important indicators of performance standards considered in Tier 2 and Tier 3: performance level and growth (Mellard et al., 2008). At-risk student's progress is monitored by evaluating growth, determining the rate of achievement, and assessing specific academic skills or behavior goals (Johnson et al., 2006; Connor et al., 2014). Since Tier 2 and 3 specify the targeted area of deficit, frequent performance levels must be considered when analyzing individual student data (Mellard et al., 2008, p. 81). Therefore, explicit decision rules need to be created by the RTI team for when a student (a) may not need intensive interventions and can return to Tier 1; (b) needs a different intervention within Tier 2 or; (c) may need to be identified for special education as Tier 3 (Johnson et al., 2006). Timely decisions within the RTI model are critical to

ensuring a student's long-term success. Research provides recommendations for timely decision making:

1. Assess student progress using CBM in Tier 2 and beyond twice per week.
2. Chart results and analyze student's progress regularly.
3. Use preset rules to determine when a student is not adequately responding. (commonly suggested rules are that four consecutive data points below the goal line warrant changes to the intervention; four above the goal line warrant raising the goal. (Stecker & Lembke, 2005, p. 4; Johnson et al., 2006, p. 2.4)

Researchers recommendations vary in response to the frequency of progress monitoring in Tiers 2 and 3. Mellard et al., (2008) cite that progress monitoring takes place once to three times per week (p. 81). Kovaleski (2007b) cites that direct assessment in the intensive phase can range from twice a month to twice a week (p. 84). Individual students or groups of students can be progress monitored on a daily, weekly, bi-weekly or monthly basis, depending upon students' need. Frequent progress monitoring (weekly or bi-weekly) provides data for decision making (Lembke, McMasters & Stecker, 2010, p. 2).

The National Association of Special Directors of Special Education (NASDSE, 2005) has identified nine essential characteristics for progress monitoring to be useful in an RTI context. Progress monitoring should do the following:

1. Assess the specific skills embodied in state and local academic standards.

2. Assess marker variables that have been demonstrated to lead to the ultimate instructional target.

3. Be sensitive to small increments of growth over time.

4. Be administered efficiently over short periods.

5. Be administered repeatedly using multiple forms.

6. The result in data can be summarized in teacher-friendly data displays.

7. Be comparable across students.

8. Be applicable for monitoring an individual student's progress over time.

9. Be relevant to development of instructional strategies and use of appropriate curriculum that addresses the area of need. (p. 224-225)

DIEBELS and AIMSweb are two commercial products commonly purchased by schools and districts to provide preventive researched-based, norm-referenced measures of reading and pre-reading skills (Good, Gruba, & Kaminski, 2002; Good, Simmons, & Kame'enui, 2001). However, other criterion measures such as the Comprehensive Inventory of Basic Skills-Revised (CIBS-R) (Brigance, 1999), or even some norm-referenced measures such as the Test of Oral Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999) may also be appropriate. According to researchers, the assessment system that is chosen (1) should directly assess the skill of interest (e.g., reading fluency rather than expressive or receptive language); (2) should address both fluency and accuracy; (3) should include multiple measures; and (4) should be easy and

efficient enough to use with all students (Mellard, 2008; Burns & Coolong-Chaffin, 2006).

Johnson, Pool, and Carter, in their article titled, "Screening for Reading Problems in an RTI Framework," provide other factors that challenge the accuracy of identifying reading problems from screening (2009). These include the amount of time between screening and outcome measures, the complexity of the reading construct, and the reliance on brief measures to predict these complex outcomes. Each of these is discussed in turn.

Time between screening and outcomes. The goal of screening students is to identify whether said student will be a successful reader in the future. Studies show that the time between screening and outcomes measures are more accurate when performed close together (Coyne, Kame'enui, & Simmons, 1998; Johnson et al., 2009; Hintze & Silberglitt, 2005).

The complex nature of reading. Learning to read is a complex skill that comprises of five components, called the "Big 5" by the National Reading Panel. They are phonemic awareness, decoding, vocabulary, comprehension and fluency (National Institute of Child Health and Human Development, 2000). Research indicates that using screening tools that address all the areas of learning to read will increase the accuracy of student outcomes (Compton et al., 2006; Johnson et al., 2009; O'Connor & Jenkins, 1999).

Reliance on brief measures to predict reading outcomes. Progress monitoring provides the data for efficient and effective decision-making. For example, the information can be used to group students with similar needs, and to provide evidence of student success with a particular intervention or service delivery over a specific period of time (Johnson et al. 2006; Mellard & Johnson, 2008, Burns et al., 2006). An important purpose of progress monitoring at Tier 2 is to determine if the intervention was successful in helping the student learn at an appropriate rate (Johnson et al., 2006; Connor et al., 2014). Providing timely services in Tier 2 for at-risk students can be the difference between them returning to regular education as disability-free or entering special education (Compton et al., 2006; Mellard & Johnson, 2008; Johnson et al., 2006). Once Tier 2 interventions have been delivered, students are evaluated as responsive or nonresponsive. Torgesen (2000) estimated that two to six percent of the general school population will be nonresponsive to Tier 2 instruction (p. 61). In the United States, data documenting a student's response to evidence-based intervention can be used as part of the process of identification of a learning disability (IDEA, 2004). Approximately 5 percent of the student population fail to demonstrate minimal progress in the less intensive Tier 2 level, are provided Tier 3 or tertiary interventions (Mellard & Johnson, 2008, p. 3; Brown-Chidsey & Steege, 2010).

Tier 3

Waznek and Vaughn state, "students who fail to make minimal gains after taught with high quality, validated classroom instruction in Tier 1, and additional intervention in

Tier 2, are considered to be inadequately responding to intervention or nonresponders" (2010, p. 306). In some school districts, Tier 3 is often synonymous with special education. NASDE defines Tier 3 as "the third level of intervention for special education students or students who did not respond to 20 weeks of Tier 2 intervention. This is the most intense level of instruction. Instruction is changed promptly based on progress monitoring data to ensure effectiveness of instruction" (p. 17). Mellard and Johnson (2008) define special education "as specially designed instruction to meet the unique needs of students with disabilities" (p. 101). Johnson et al., in the Response to Intervention manual define special education as "generally specialized instruction to meet the unique needs of students with disabilities" (p. 80). To achieve academic success, students with Learning Disabilities (LD) require intensive, iterative (recursive), explicit, and scientifically based specialized instruction that is monitored on a continuing basis (National Joint Committee on Learning Disabilities [NJCLD], 2002).

When students have been unresponsive at Tier 1 and 2, they are often provided an individualized, comprehensive evaluation to determine if they have a disability (Mellard et al., 2008). Although IDEA (2004) provides flexibility to LEA's in determining a student with learning disability (SLD), the National Joint Committee on Learning Disabilities provide four recommendations to help guide the development of these procedures (NJCLD, 2005):

1. Decisions regarding eligibility for special education services must draw from information collected from an individual comprehension evaluation using

multiple methods including clinical judgment and other sources of relevant information.

2. Students must be evaluated on an individual basis and assessed for intra-individual differences in the seven domains that comprise the definition of SLD in the law-listening, thinking, speaking, reading, writing, spelling, and mathematical calculation.
3. Eligibility decisions must be made through an interdisciplinary team, must be student-centered and informed by appropriate data, and must be based on student needs and strengths.
4. As schools begin to execute a process of decision-making that is more clinical than statistical in nature, ensuring through regulations that this team of qualified professionals represents all competencies necessary for accurate review of comprehensive assessment data will be critical. (p. 11)

The intensity of Tier 3 is accomplished by increasing the length of time in intervention, decreasing the group size, the frequency of the sessions, and the duration of the intervention (Denton et al., 2013; Vaughn, Denton & Fletcher, 2010; Reschly, 2005; Wanzek & Vaughn, 2010). Targeting student deficits may also increase intensity by increasing its efficiency (Denton et al., 2013; Wanzek & Vaughn, 2010). According to National Research Center on Learning Disabilities, interventions in special education must be designed to meet the specific learning and behavioral needs of the student, implemented on a timely basis, provided by a highly qualified teacher or specialist, and

monitored to determine progress and achievement of desired outcomes (p. 3.15). Swanson, (1999) provided specific forms of special education instruction that have found most effective in teaching students with learning disabilities combine direct instruction with strategy instruction. The main features of this model include:

1. Control of task difficulty
2. Small-group instruction
3. Directed questioning and response
4. Sequencing – breaking down the task
5. Drill-repetition-practice
6. Segmentation
7. Use of technology
8. Teacher-modeled problem solving
9. Strategy cues. (p. 504)

Of these features, the first three had the most influence on student achievement. Additionally, general educators will be informed of the types of supports required in the general education classroom as students with SLD receive accommodations, modifications, and remediation specifically designed for their individual needs (Swanson, p. 504).

Implementation of Fidelity

Implementation of instruction/intervention as intended has a number of synonyms, such as fidelity of implementation, treatment integrity, efficacy of implementation, procedural reliability, and intervention integrity (Sanetti, & Kratochwill,

2009). For this paper, fidelity is defined as "the extent to which essential intervention components are delivered in a comprehensive and consistent manner by an interventionist trained to deliver the intervention" (Sanetti and Kratochwill, 2009, p. 448).

According to researchers, all the components of RTI need to be monitored for fidelity (Brown-Chidsey & Steege, 2010; Jimerson, Burns & VanDerHeyden, 2007) as it applies to the implementation of instruction and intervention in schools (Sanetti and Kratochwill, 2009). According to researchers, fidelity is an important component of the problem solving model and must be demonstrated before benefits can be determined. Without fidelity, promising reforms and practices can lead to failure (Telzrow, 2000; Mellard & Johnson, 2008).

Mellard and Johnson provide three factors of intervention fidelity that attribute to student success research. First, the various components are implemented consistently across classrooms and grade levels at the school level. Second, the program or intervention has a strong evidence-based outcome. Finally, the person implementing the intervention does so in the way in which it was designed (2008, p. 118-119).
Fidelity checks of the RTI system that are delivered and implemented as intended are considered to be best practice (Mellard & Johnson, 2008). Future research of problem solving implementation, empirical intervention and intervention adherence and treatment integrity will provide an increased understanding of the conditions in which students in the learning environment are most conducive to improvement (Telzrow, 2000).

Research Support for RTI

The purpose of this research support for RTI section is to provide an overview of the nature and extent of the published research in which to provide the reader with knowledge to make informed judgments about evidence base for RTI. In a report by What Works Clearinghouse (2008), studies that provide the strongest evidence of success are randomized controlled trials or a quasi-experiment with one of the following three designs; a) quasi-experiment with equating, b) regression discontinuity designs, c) or single-case designs (p. 5). However, school districts often adopt RTI as a district-wide initiative, alleviating randomized studies that would provide strong evidence to support its use (Burns, 2010a). This researcher provided studies examining the efficacy of RTI programs as well as studies reporting on the five typical core components of RTI.

Although RTI had its roots in special education, it gained momentum as a prevention and intervention framework to support all students (Fuchs et al. 2003; Gersten et al., 2009; VanDerHeyden et al., 2007). The following section reviews the key groups that have conceptualized, researched and promoted RTI as a means to prevent reading deficits and to identify students who may be learning disabled.

Many studies of RTI provide comprehensive data on success in all educational settings (e.g., Marston, Muyskens, Lau, & Canter, 2003; Compton, Fuchs, Fuchs, & Bryant, 2006). Burns, Appleton and Stehouwer (2005), completed a meta-analysis of four existing RTI models (1) the Heartland Agency (Iowa) Model, (2) Ohio's Intervention Based Assessment, (3) Pennsylvania's Instructional Support Team, (4)

Minneapolis Public School's Problem-Solving Model, and 10 other studies developed and implemented by researchers.

> The results found that large-scale RTI models were effective although the four field-based models were stronger than the ten research initiated models. Experts agreed that the field-based models had been in existence longer and had gone through a refinement process. Also, student and systemic outcomes both showed improvement. Finally, less than two percent of students had been identified as learning disabled in this meta-analysis, a much lower percentage than predicted from previous analyses. (p. 389)

Several studies provide alternate insights into students who may be at-risk. According to Vellutino, Scanlon, Small, and Fanuele (2006), most reading difficulties are a result of environmental or instructional deficits rather than cognitive disability (p. 167). In another study, Vellutino, Scanlon and Lyon (2006) concluded that intelligence testing did not differentiate reading impairment students from normal readers, or between reading impaired students who were difficult to remediate and those who were easy to remediate. Finally, IQ scores did not provide evidence of reading achievement levels in normally developing readers. What they found was that reading impairment is predictable by testing phonological skills (p. 236).

RTI has been recommended as a model for early identification and prevention of student's at-risk for reading difficulty and as a method of determining LD and special education eligibility. Researchers and contemporary RTI models are beginning to

document that RTI practices provide accurate, valid, and reliable identification of students at-risk of reading difficulty and identification of a Learning Disability (Fuchs, Fuchs, & Speece, 2002; McMaster, Fuchs, Fuchs, & Compton, 2005; Speece, Case, & Malloy, 2003; Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino, Scanlon, Small, & Fanuele, 2006).

State of Missouri and MAP

The Missouri Assessment Program evolved as a result of Missouri's Outstanding Act of 1993 in order to assess student achievement based on the Missouri Show-Me Standards. This grade span test was originally created to assess Communication Arts and in grades 3, 7 and 11; Mathematics in grades 4, 8, and 10; and Science in grades 4, 8, and 11 (MO DESE, 2014, p. 4). In response to NCLB of 2001, MO DESE increased testing to assess Communication Arts in grades 3-8, and once in grades 10-12 by 2003; and Science in grades 5 and 8 by the 2007-2008 school year. During this time, MO DESE contracted with CTB/McGraw-Hill to expand their construct grade span assessments (MO DESE, 2014, p. 4). In 2008, MO DESE revised the high school assessments by eliminating the MAP test and expanding to End of Course (EOC) assessments (MO DESE, 2014, p. 4).

State of Missouri and RTI

This current study is interested in investigating the whether the use of R-CBM formative measurements embedded in the RTI process in one suburban Midwestern school district XYZ, has a relationship with the summative Missouri Assessment

Program (MAP) Communication Arts summative assessment. Nationally, RTI was promoted as a promising avenue to support at-risk and identify students with a Specific Learning Disability (SLD) as part of IDEA in 2004. However, it was not until 2008 that the state of Missouri provided guidelines for districts who desired "to use a scientific research-based intervention process to determine non-responsiveness to instruction" and eligibility for special education services. Any district wishing to implement an RTI process to determine students with SLD is required to: 1) use evidence or scientific-based interventions that are proved effective in the remediation of specific deficits for which the student is lacking, 2) use at least two interventions, 3) document a minimum of 24 intervention sessions over the two interventions, 4) ensure that documentation includes fidelity of implementation, 5) collect and document student progress at least once a week, 6) show student's criteria as it relates to the rate of progress and academic skill level by the end of instruction and student targets established along the way (MO DESE, 2008a, p. 1-2). DESE's Special Education Advisory Council explains that Missouri's three tiered process is called the Missouri Integrated Model (MIM). MIM "supports academic achievement and successful behavior through tiered levels of support that acknowledge and address diversity in student learning" (MO DESE, 2008b, p. 2).

RTI in a Missouri Suburban School District

The Missouri Suburban School District, labeled XYZ in this study, implemented Response to Instruction district-wide in 2013-2014 as a comprehensive school improvement initiative with the intended goal of preventing student academic difficulties

before students fail. The three tier initiative, modeled after MIM, begins in Tier 1, with all students engaged in a high quality, evidenced-based curriculum that is aligned to Missouri standards. These standards provide a clear understanding of what students are expected to learn in reading, math, writing, speaking and listening. This district's model includes the use of a continuum of student performance data to inform, monitor and enhance the education of all students. Three times during a school year, all students are universally screened for academic progress. Student scores are nationally normed. District XYZ employs AIMSweb assessment, data management and reporting system for universal screening and progress monitoring K-12 grade.

Students enter Tier 2 when they are not making grade level progress and have a universal screening score of 25% or below. Tier 2 provides timely, intensive and targeted instruction by a certified teacher in the area of need while continuing the core curriculum. After setting specific individualized goals for each child, instruction is offered in small, differentiated groups, for the duration of twenty to forty minutes, three to four times each week. Students are progress monitored weekly. Teams of professionals called the Progress Monitor Data Team, monitors student data, the fidelity of implementation, and flexible grouping as students move through the tiers. These teams provide support to establish and analyze student progress, or lack thereof, of year-long goals. Long range goals are reliable predictive validity to trend-lines allowing the team to predict sufficient or insufficient progress long term. Sufficient/adequate progress is determined when the student's trend line and goal line will intersect on or before the last day of intervention

and progress monitoring. Students who have four consecutive data points below their goal line or whose trendline will not intersect with their goal line on or before the last day of intervention and progress monitoring (only after a sufficient number of intervention sessions and a final number of 6 data points) is scheduled immediately for Problem Solving and possible change of intervention (Intervention Manual, 2013).

Students enter Tier 3 when they score at 10 percent or below on universal screening, or they are not making progress in Tier 2. Students who were ranked at the 10^{th} percentile or below were provided Survey Level Assessment or tested to find their instructional level for instructional materials and progress monitoring. When they meet established criteria, they should be moved up one level or back to grade level as soon as possible for continued progress monitoring. Since these students are at greater risk of not meeting benchmark targets, intensive instruction is often delivered by a specialist, in small groups or 1:1, more often and often longer in scope. Progress is monitored weekly. Progress Monitor Data Consult Team monitors the progress of the student. Embedded within the context of the continuum is a team approach to data-based decision making that evolves from members of the building that includes a process coordinator within Tier 3.

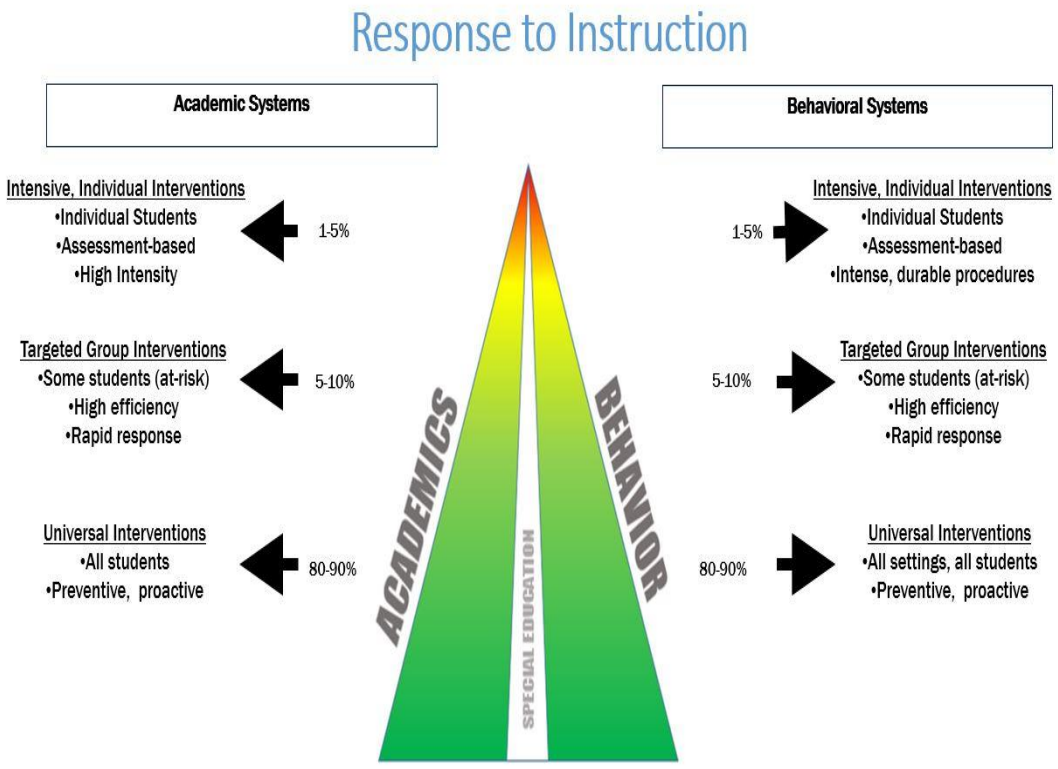

Figure 2. Response to Intervention for Midwestern Suburban School District XYZ

Summary of Literature Review

The literature review outlined the key historical events that played a role in the development of RTI. This included the progression of educational laws, developments of teacher assessment teams, the downfall of the discrepancy model and the renewal of the RTI process to intervene before students fail. This review sought to give an extensive view of the RTI models, its core features, relationship to student achievement and role it plays in education today, particularly on Midwestern Suburban School District XYZ. This review indicates that a substantial body of research supports the implementation of

RTI and its core features to improving student outcomes, predominantly in reading. Furthermore, large field-based studies from researchers of RTI programs have demonstrated that when these practices are implemented with fidelity, RTI holds great promise for improving student achievement. Although a foundation has been laid as to the success of RTI in controlled studies, research is far from complete. If the purpose of RTI is to improve academic outcomes for all students, then it must achieve the goal of bringing those at risk students to proficiency. The focus of this study is whether in a large suburban district these components of RTI can achieve the goal of decreasing the number of students at risk of failing and increase the number of students scoring within the national norms on benchmark tests and state mandated tests.

Chapter Three

Methods

Introduction

The purpose of this study was to determine the success of RTI implementation by determining what relationship exists between R-CBM and MAZE CBM benchmark assessments and Communication Arts MAP assessment for fourth grade students in District XYZ. This district employed a scientific evidenced-based core curriculum, early detection of at-risk elementary students, specific interventions for each student, frequent progress monitoring and continued benchmarking to determine student success. To date, early intervention of all at-risk students within the district has taken place at the elementary level. It is the hope of District XYZ that implementing RTI will close the achievement gap for all students at the district and elementary building levels. This chapter serves to identify the methods and procedures utilized in this study, including a description of the design, population, and sample of students studied, data collection and analysis with respect to student achievement, and limitations used in this study.

Research Design

This study followed a one year non-experimental research design that used archival data to examine prediction of formative oral reading fluency and reading comprehension fluency scores to the MO MAP summative state test for Communication Arts. This study employed a multiple regression model to address the research questions. "Multiple regression is a prediction equation that determines the correlation between a

combination of two or more predictor variables and a criterion variable" (Lunenburg & Irby, 2008, p. 80). The goal of a stepwise multiple regression model is to enable a researcher to determine the combination of independent variables (i.e., predictor variables) that would best predict the dependent variable (i.e., criterion variable). The independent variables in this study were the fall R-CBM, winter R-CBM, the spring R-CBM, the fall MAZE CBM, the winter MAZE CBM and the spring MAZE CBM given to all fourth grade students. The dependent variable was the 2014 Communication Arts MAP summative assessment scores for all fourth grade students in District XYZ.

Population and Sample

The student population used for this quantitative study included all fourth grade students in District XYZ that participated in all 2013-2014 fall, winter, and spring oral reading fluency (R-CBM) and reading comprehension (MAZE) interim assessments and the spring MAP summative assessment. District XYZ had 1,382 students enrolled in 18 elementary school buildings (MO DESE, 2015). The student population was due to proximity of the students to the researcher and the availability of the researcher to collect data on these students.

Sampling Procedures

Participants were selected from fourth grade in XYZ school district. All students were eligible to participate in the study if they were enrolled in the school district during the assessment period of the 2013-2014 school year and participated in AIMSweb screenings and the MAP assessment. If the student did not participate in all of the

required assessments, (fall R-CBM, winter R-CBM, spring R-CBM, fall MAZE CBM, winter MAZE CBM, and spring MAZE CBM and MAP assessment) then they were excluded from the sample.

Instrumentation

AIMSweb Assessment, Reporting, and Data Management System by NCS Pearson, Inc. were utilized by this study to measure oral reading fluency and comprehension. AIMSweb is a benchmark and progress monitoring system based on direct, frequent and continuous student assessment in reading, mathematics, spelling, written expression and behavior (NCS Pearson, Inc., 2012b). NCS Pearson, Inc. developed high quality sets of Standard Reading Assessment Passages (RAPs) to provide scientific evidence of technical properties of the CBMs for use of benchmarking, intervention, and monitoring students' development of reading (p. 1).

For over 30 years, researchers and educators alike have used general outcome measures (GOM) such as Curricular Based Measurement (CBM) to; frequently measure foundational reading skills, predict future academic success, and to provide ongoing curricular goals to ensure the success (Deno, 2003; Good III, Kaminski, Simmons, & Kame'enui, 2001). According to Burns (2010b), GOM are characterized as standardized measures that assess proficiency of global outcomes associated with an entire curriculum (p. 27). The goal of GOM is to assess instructional effectiveness and quickly make changes as needed. In 2005, Stecker and Lembke defined CBM as "a scientifically validated form of student progress monitoring that incorporates standard

methods for test development, administration, and data utilization (p. 1). According to Reschly et al., (2009) decades of research have proven CBM assessments are valid and reliable measures of moderate to high correlations in reading achievement (as cited in Patton et al., 2014).

R-CBM. AIMSweb provides two CBM measures to determine a student's reading ability. First, the R-CBM measures a student's oral reading fluency (OAF). The narrative, fictional passages have 250 words for first and second grade, and 350 for third through six grade. Within specific grade levels, all passages were written to match each other (alternate-form reliability), in difficulty (average WRC), while representing the general curriculum (Shinn & Shinn, 2002a; NCS Pearson, Inc., 2012a). For interim assessments, a tester listens as students orally read three one-minute passages and determines the median score of WRC and number of errors (NCS Pearson, Inc., 2012a, p. 8). The passages used for interim benchmarks are housed in the AIMSweb Assessment System. The first three passages are used for benchmarking and the remaining 30 passages are used for progress monitoring at-risk.

Reliability. Reliability is the degree to which an instrument consistently measures whatever it is measuring. It requires demonstrating the operations of a study can be repeated with the same results (Lunenburg & Irby, 2008). Reliability is the correlation between two or more tests. Coefficients values of reliability range from 0.0 to 1.0 with 1.0 indicating that all test score variances are true. According to Del Siegle, the minimum reliability required for research purposes is normally .70 (2002).

NCS Pearson's 23 passages for first grade and 33 passages for second- eighth grades reliability were rated according to alternate-form reliability and Lexile level. Alternate- form reliability is the compatibility of scores between the different versions of RAPs. To ensure readability at each grade level, MetaMetrics aligned reading Lexile levels to the RAPs (NCS, Pearson, 2012a). Also, split-half reliability or, a correlation in which scores on the first 30 seconds of the 1-minute R–CBM administration were correlated with the second 30 seconds was performed and provided consistent data over time. According to NCS, Pearson, 2012a, a sample study was performed in 2011 with 63 fourth grade students from 5 public schools. Results indicated a reliability of .95 on individual probes. Table 3 demonstrates the reliability of three assessment RAPs from their article, *Standard Reading Assessment Passages (RAPs) for General Outcome Measurement* (Howe & Shinn, 2002). These correlations prove that AIMSweb is a reliable measure of readability.

Table 2.

Reliability of the Standard Reading Benchmark Passages

Passage	Mean WRC	Standard Deviation	Alternate Form Reliability[a]	Lexile
1	121.5	20.1	0.87	770
2	121.8	27.2	0.82	650
3	122.8	24.5	0.86	670
Mean	122.0	23.9	0.85	

Note. Adapted from "Standard Reading Assessment Passages (RAPs) for General Outcome Measurement," by K. B. Howe & M. M. Shinn, 2002, p. 8.

[a]Mean correlation for each alternate form RAP.

Validity. While reliability of an assessment measure is important, validity is essential for showing the accuracy of the assessment. "Validity is the degree to which an instrument measures what it purports to measure" (Lunenburg & Irby, 2008, p. 181). Criterion validity determines the relationship between the test score and another measure such as a test or a program (NCS Pearson, 2012). Table 3 depicts criterion validity from the North Carolina end of the fourth grade test as compared to AIMSweb fall, winter, and spring benchmarks.

Table 3

Validity of R–CBM Screening Scores

Criterion	Grade/Season	n	Correlation
North Carolina End of Grade Test	4/ Fall	1,174	.70
	4/Winter	1,174	.71
	4/Spring	1,174	.72

Note: Adapted from "Criterion Validit of R-CBM Screening Scores, From the AIMSweb database" by NCS Pearson, Inc, 2012a, p. 11

The National Center on Intensive Intervention (n.d.) provides evidence to support that R-CBM ORF is a valid for benchmark testing, interim screening and progress monitoring students.

MAZE CBM. The second measure of readability is the MAZE CBM or reading comprehension tool used to measures a student's comprehension. Although the same CBM passages are used as in the R-CBM assessment, MAZE CBM is a multiple-choice cloze task that students complete while reading silently. The first sentence of a 150-400 word passage is left intact. Thereafter, every 7th word is replaced with three words inside parenthesis. Students choose one of the three words to complete the sentence. One of the words is the exact one from the original passage, while the other two are distractors. One distractor is from the same part of speech as the correct word but does not make sense, and the second distractor is from a different part of speech as the correct word and does not make sense in context (NCS Pearson, Inc., 2012a; Shinn & Shinn, 2002b).

The MAZE CBM is a standardized assessment therefore; it must be administered, scored and interpreted in a standard way. Educators can assess students individually, in a small group, or the entire class at the same time. There are three standardized benchmark MAZE reading passages for each grade level. Students are prompted with testing instructions, and given three minutes to complete.

Reliability of MAZE CBM. In a research study by Shinn, Deno and Espin (2000), technical characteristics of the MAZE CBM were examined in terms of reliability, sensitivity and validity (p. 165). In order to determine student overall performance over time, CBM measures use alternate forms of the assessment to obtain multiple data points. Shinn, et. al., (2000), determined that MAZE CBM's had strong alternate-form reliability with a mean coefficient of .81 and 1-3 month intervals between testing" (p 164). The project further confirmed that "MAZE CBM performance scores were sensitive to group and individual growth rates while student reading performance can be positively correlated with earlier growth rate" (2000, p. 164). Another study by Brown-Chidsey, Johnson Jr., & Fernstrom (2005), confirm that "MAZE CBM can reliably monitor students silent reading skills over time as they progress throughout school" (p. 392).

Validity. There is also evidence of validity with the MAZE CBM. In a study by National Center on Intensive Intervention, (n.d.) found the criterion validity on the Comprehensive Reading Assessment Battery: Words Read Correctly in grades 1-7. The scores ranged from .71- .91 with a mean of 83.5 (n.p.). National Center on Intensive

Intervention supports that MAZE CBM is a reliable and valid tool of student reading comprehension.

In Table 4, NCS Pearson, Inc. provides evidence on validity (2012a) as the median correlation of .59. This modest relationship suggests that MAZE CBM is moderately reliable and valid. As cited in Scholin & Burns (2012, p. 386), "a moderately high correlation between CBM-R and achievement tests provides support for the use of CBM-R data in identifying students at risk for not passing state achievement tests" (Reschly, Busch, Betts, Deno, & Long, 2009; Stage & Jacobsen, 2001).

Table 4

Median Correlations of Maze Scores with State-Test Scores for Grade 4

Grade	Total N	Number of States	Median Correlations
4	8.41	10	.59

Note: Adopted from "AIMSweb Technical Manual", by NCS Pearson, Inc. Inc., 2012a, p. 16.

Measurement

District XYZ administers reading interim benchmarks three times a year during predetermined time frames (called windows) using standard universal screening probes from AIMSweb assessment system. Table 5 outlines the assessments given to all fourth students annually (District XYZ RTI Manual, 2013, n.p.).

Table 5

Annual Reading Benchmark Assessments Given to all Fourth Grade Students

Fall	Winter	Spring
R-CBM[a] *Reading*	R-CBM[a] *Reading*	R-CBM[a] *Reading*
MAZE CBM[a] *Reading*	MAZE CBM[a] *Reading*	MAZE CBM[a] *Reading*

Note. Adapted from "Annual AIMSweb Benchmark Assessments" by Shinn & Shinn, 2002a; Shinn & Shinn, 2002b

[a]R-CBM Benchmark and MAZE CBM assessments are administered each Benchmark period. The median score of corrects and the median score of errors are then used as the Benchmark score. All other measures are administered once per Benchmark period.

These performance interim assessments provide raw scores that when compared to a norm group or comparison population; provide evidence of student academic success. According to NCS (2012b), "national and state norm groups are the most stable because they have the largest populations." (p. 2) Reports from the oral reading fluency, and the reading comprehension interim benchmarks, identify students readability using a set of national norms that detail student's performance-level within a percentile category. NCS defines a percentile as a useful type of score for understanding whether a student is performing higher, lower, or at the same level as his peers. (p.2). These categories are based on percentile ranks; ninety-ninth, seventy-fifth, twenty-fifth and tenth. Students who perform in the twenty-fifth percentile or higher are considered to be within the average percentile (NCS, Pearson, 2012a). In District XYZ, students who score below

twenty-fifth percentile on an interim assessment are provided remedial instruction as indicated in the district's RTI process.

Three variables were analyzed based on the research questions and hypotheses: the R-CBM oral reading fluency, MAZE comprehension scores and the Communication Arts portion of the Missouri State MAP assessment scores. The R-CBM and MAZE scores served as the independent variables and the Communication Arts MAP assessment scores served as the dependent variable for the study.

R-CBM measurement. During ORF screening, errors made by the student are marked by the assessor on their copy of the RAP. Upon assessment completion, two scores are provided; a primary score of Words Read Correctly (WRC), and a secondary score of accuracy. Accuracy accounts for the number of errors and the percentage of words read correctly. For universal benchmark screening, the tester calculates each of the three probes and the final score is the median (or middle) of the three values. Educators are instructed on how to score the R-CBM and how to manually enter the student's scores into the database. There is a computer version that allows the tester to electronically mark errors during the assessment, then automatically scores it, and enters the results into the AIMSweb database. To score the probe, the tester marks errors on his copy of the test. Scores are averaged and students are provided a score of words read correctly (WRC) in one minute. A WRC score that reflects the number of errors and the percentage of words read correctly is then compared to national norm R-CBM cut scores for fourth grade. Table 6 provides those cut scores.

Table 6

AIMSweb National Norms R-CBM Cut Scores for Grade 4

	Fall	Winter	Spring
%	WRC	WRC	WRC
90	160	178	196
75	134	152	168
50	107	125	139
25	84	101	112
10	61	78	90

Note: WRC= words read correctly. Adapted from Pearson Aimsweb (2013-2014). Retrieved from aimsweb.pearson.com/Report.cfm?inc=rReport18.cfm&ClearSchool =1&Aggregate=1

MAZE CBM measurement. After MAZE CBM testing, scorers use a master key with the correct word bolded. Answers are incorrect if the student circles the incorrect word or skips a word selection other than those they were unable to complete in the three minute period. Scorers record the number of correct over incorrect answers although the number of correct responses is the most valuable (Shinn & Shinn, 2002). The score is then compared to national norm MAZE cut scores for fourth grade. Table 7 provides those cut scores.

Table 7

Aimsweb National Norms MAZE Comprehension Cut Scores for Grade 4

	Fall	Winter	Spring
%	WRC	WRC	WRC
90	22	32	34
75	18	27	28
50	14	20	20
25	10	15	15
10	6	11	11

Note: WRC= words read correctly. Adapted from Pearson Aimsweb (2013-2014). Retrieved from aimsweb.pearson.com/Report.cfm?inc=rReport18.cfm&ClearSchool =1&Aggregate=1

MAP summative assessment. In 2014, Missouri's Assessment Program (MAP) standardized assessments were given to public school pupils to evaluate student learning toward mastery of the Show-Me Standards and Grade Level Expectancies (GLE). The GLE were specific academic indicators of what each child in the state of Missouri should be able to know and do in each content area (CTB/McGraw-Hill, 2014b, p. 5).

According to the MAP Grade-Level Assessment Guide to Interrupting Results, (CTB/McGraw-Hill, 2014a) each Communication Arts assessment takes 2-4 hours of time, and may include selected response (multiple-choice), constructed-response, and performance events (writing prompts) (p.1). In 2014, the MAP assessment was

comprised of questions authored from Missouri educators and edited by MO DESE and CTB staff, and from the Survey edition of the *Terra Nova,* a nationally normed test developed by CTB (CTB/McGraw-Hill, 2014b, p. 10). The *Terra Nova Survey* is an abbreviated version of the Complete Battery and provides a general measure of achievement, with a minimum amount of testing time. The Survey generates norm-referenced achievement scores, criterion-referenced objective mastery scores, and performance-level information (CTB/McGraw-Hill, 2014b). CTB/McGraw-Hill state in MAP Grade-Level Assessment Technical Report, (2014b), "Portions of the MAP from CTB's item pool were also aligned to Missouri Content Standards, Process Standards, and GLEs to further solidify the Show-Me Standards as the foundation of the MAP." (p. 10)

MAP measurement. From the MAP assessment, schools and districts are provided individual student, classroom, and school reports in each content area. This information can be used by educators and districts to determine academic strengths and weaknesses in relation to instruction and to help in instructional planning (CTB/McGraw-Hill, 2014b). There are two types of test level scores that are reported to indicate students achievement on the MAP; 1) the scale score, 2) its associated level of achievement (CTB/McGraw-Hill, 2014b, p. 10). A scale score communicates students' achievement in a content area on the MAP assessment. The higher the scale score equates to higher levels of achievement and lower scores equates with lower achievement. Next, students' performance is categorized into an achievement level that is associated to the scale score

of; *Below Basic, Basic, Proficient, or Advanced.* Achievement level scores reflect what students should know and be able to do according to the Missouri Show-Me Standards/GLE (CTB/McGraw-Hill, 2014b, p. 11). It is the goal of the State of Missouri's Department of Education for students to score in the proficient or advanced levels (CTB/McGraw-Hill, 2014b). Table 8 compares the MAP Communication Arts achievement descriptors and score range.

Table 8

Grade 4 Communication Arts Achievement-Level Descriptors/Scores

Level	MAP Score Range
Below Basic	460-611
Basic	612-661
Proficient	662-690
Advanced	691-820

Note. Adapted from "Missouri Achievement Program: Guide to Interpreting Results," by CTB/McGraw-Hill, 2014b, p. 5.

MO DESE offered to district, content standard sub scores as a means to provide specific strengths and weaknesses in tested content areas (CTB/McGraw-Hill, 2014, p. 11). While classroom teachers find this helpful to determine what their previous students knew and didn't know, districts often use this information to write curriculum, and provide teacher in-service for curricular areas of need. For each MAP test content area, a pool of content and process standards sub scores are created from the same pool of items (CTB/McGraw-Hill, 2014b, p. 11). The purpose of reporting the process standard sub

scores on the MAP assessment is to show the achievement in relation to the aligned Missouri Process Standards.

Missouri Department of Education employed CTB to analyze test items and the overall test function. “The data analyses undertaken by CTB address multiple best practices of the testing industry but, in particular, are related to the following *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999)” (CTB/McGraw-Hill, 2014b, p. 74). Selected-Response items on the MAP Communication Arts Test were scored by CTB using an electronic scanning machine. Highly trained examiners and scorers were trained by Department of Education and Kelly Services to complete the hand scoring or constructed and extended response items (writing essays and performance events) (CTB/McGraw-Hill, 2014b, p. 57). Several monitoring practices were put into place for constructed response or extended response items.

1. Daily Accuracy Checks. Calibration sets of pre-scored papers were administered daily to each scorer to monitor scoring accuracy and maintain a consistent focus on the established rubrics and guidelines.
2. Read-behinds. Approximately 5% of hand scored items were dually graded by a team leader who rereads a scored item. This technique is a valuable rater-reliability technique where a scorer’s proficiency is checked. Scorers are provided counseling as needed to improve scoring ability.

3. Recalibration of Raters. This process realigns or recalibrates scorers or raters who begin to drift away from scoring accuracy. This consists of reviewing the rubrics, anchors, and training papers and being administered a recalibration round. (CTB/McGraw-Hill, 2014b, p. 65-66)

MAP reliability and validity. Reliability is the precision of a test when given repeatedly and under similar conditions. In accordance with the AERA, APA, & NCME Standards (1999) and in developing and maintaining tests of the highest quality, CTB has calculated the reliability of each MAP test in a variety of ways: reliability of raw scores, overall standard error of measurement, (Item-response theory) IRT-based conditional standard error of measurement, and decision consistency of achievement-level classifications (2014b, p. 128). CTB/McGraw-Hill (2014b) used Cronbach's coefficient alpha to calculate the reliability of the raw data scores of the MAP tests. The reliability coefficient ratio is 0-1. As seen in Table 9, there is strong consistency on the 4th grade MAP test (p. 129).

Table 9

Reliability Coefficient Data for Communications Arts for Grade 4.

Number of Items	Cronbach's a
51	0.90

Note. Adapted from "Missouri Grade Level Assessment Technical Report 2014," by CTB McGraw-Hill, 2014b, p. 138.

A principal component factor analysis was conducted on each grade/content area MAP. The Standard Error of Measurement (SEM) estimates how repeated measures of a

person's score on an instrument tend to be distributed around their "true score." Although a student's true score are always unknown, "it is expected that 68% of the time a student's score obtained from a single test administration would fall within one SEM of the student's true score and that 95% of the time the obtained score would fall within approximately two standard errors of the true score". (CTB/McGraw-Hill, 2014b, p. 129) Table 10 contains the mean, standard deviation and SEM of Communication Arts content standards for grade 4. This information supports the claim of a strong reliability of the MAP.

Table 10

Technical Properties of MAP Communication Arts Content Standards

Content Standard	Mean	Std. Deviation	SEM*
1	66.49	20.67	13.71
2	77.88	17.75	6.15
3	63.79	23.45	15.02
5	74.75	20.00	10.20

Note. Adapted from "Missouri Grade Level Assessment Technical Report 2014" by CTB McGraw-Hill, 2014b, p. 149. *SEM is reported in percent correct metric as Content Standards are reported in that metric.

Classification accuracy refers to the probability that an examinee's achieved grade classification on an assessment reflects their true grade, while classification consistency refers to the probability that an examinee will be classified into the same grade classification under repeated administrations of an assessment (Wheadon, 2014, p. 1). In the CTB/McGraw-Hill Technical Manual it states; (2014b), "the magnitude of

classification consistency and accuracy measures are influenced by key features of the test design, including, the number of items, the number of cut scores, the reliability, and associated SEM" (p. 132). Further, "…classification accuracy statistics are at or above 0.90 while the classification consistency statistics are at or above 0.87. These results suggest that consistent and accurate performance level classifications are being made for students in Missouri based on the MAP" (CTB/McGraw-Hill, 2014b, p.133).

The MAP assessment items and overall test have been aligned with the Missouri Show-Me Standards and GLE strands, thus demonstrating validity (p. 13) and that "the test items as well as the overall test are functioning appropriately". (p. 74) (CTB/McGraw-Hill, 2014b). In the spring of 2014, 1,382 students took the MAP test. Table 11 shows the number of fourth grade students who scored Below Basic, Basic, Proficient and Advanced (CTB/McGraw-Hill, 2014b).

Table 11

Students Levels of Achievement on Spring MAP Test in 2014 for Grade 4

Below Basic	Basic	Proficient	Advanced
32	164	643	543

Note: Adapted from "MAP District Disaggregate Final Report, 2014", by Missouri Department of Secondary and Elementary Education, n.p.

Data Collection Procedures

Protocols for privacy, confidentiality, and student rights are a concern when conducting research. Therefore, prior to data collection, strict guidelines were in place to maintain these rights. First, Baker University's Institutional Review Board (IRB)

reviewed and accepted a proposal of research concerning this study (see Appendix A). Next, the Research Committee at District XYZ reviewed the IRB findings along with their district data collection request form (see Appendix B). When both Baker University (see Appendix C) and District XYZ personnel (see Appendix D) approved the study, the researcher was able to move forward with data collection.

Once permission was granted by the Associate Superintendent of Instruction and Leadership, this researcher was provided access to AIMSweb Data base system through AIMSweb, Inc. for District XYZ (see Appendix E). The authorization allowed this researcher to access the benchmark assessment and progress monitoring data for students at risk in Communication Arts. Numerical and coded data of the R-CBM oral reading fluency, Maze CBM comprehension, and the MAP assessments were collected. The school district removed the participants' names and student numbers from the data set of the study to protect their identities.

Data Analysis and Hypothesis Testing

This study examined the following research questions to determine which combination of variables best predicts student scores on the MAP assessments. The research questions provide the basis for the data analysis, while the hypotheses were developed to address each question. The research questions, hypotheses and the analysis used.

RQ1. What combination of variables (fall Oral Reading Fluency CBM score, winter Oral Reading Fluency CBM score, spring Oral Reading Fluency CBM score) best predict students Communication Arts MAP scale scores?

H1. A significant combination of variables was identified for predicting student scores on the fourth grade MAP Communication Arts assessment includes fall R- CBM score, winter R-CBM score, and spring R-CBM score.

RQ2. What combination of variables (fall Comprehension Fluency MAZE score, winter Comprehension Fluency MAZE score, spring Comprehension Fluency MAZE score) best predict students Communication Arts MAP Scale scores?

H2. A significant combination of variables was identified for predicting student scores on the fourth grade MAP Communication Arts assessment includes fall MAZE CBM score, winter MAZE CBM score, and spring MAZE CBM score.

Multiple regression analyses were used to find the combination of variables that best predicted MAP scores for fourth grade. A stepwise regression model was conducted using the fourth grade students in district XYZ and subtests (communication arts). Correlations were examined for the strength of the associations between each of the predictor variables and multiple stepwise regression analysis was used to predict student membership on the MAP test. The independent (predictor) variables were oral reading fluency CBM benchmarks (fall, winter, and spring) and the comprehension CBM (MAZE) benchmarks (fall, winter and spring). The Zero variables were included in the

final stepwise regression model. The regression model was significant and the hypothesis was supported.

Stepwise regression essentially performs multiple regression a number of times, each time removing the weakest correlated variable. At the end you are left with the variables that explain the distribution best. The only requirements are that the data are normally distributed (or rather, that the residuals are), and that there is no correlation between the independent variables (known as collinearity).

Limitations

The limitations of a study are the "factors that may have an effect on the interpretation of the findings or on the generalizability of the results" (Lunenburg & Irby, 2008, p. 133). Limitations are not under the control of the researcher. Conclusions from this study could be affected by the following limitations:

1. Student effort on the oral reading fluency CBM benchmark assessments
2. Student effort on the comprehension CBM (MAZE) benchmark assessments
3. Student effort on the Missouri State MAP assessment
4. Screener variance in providing directions for the assessments
5. Screener variance in the time administration of the assessments

Summary

This chapter reviewed the purpose of the study, the research design, population and sample, sampling procedures, instrumentation, data collection procedures, and presented research questions and hypotheses. The research design utilized a multiple

step-wise regression model to determine the relationship between the oral reading fluency and the comprehension CBM interim assessments to the Missouri State MAP score for fourth grade students in District XYZ for the 2013-2014 school year. The results of the data analysis are presented in chapter four.

Chapter Four

Results

The purpose of this study was to determine the extent of the relationship between oral reading fluency reading interim scores and CBM comprehension scores and that of the Communication Arts MAP assessment for students enrolled in District XYZ. Chapter four presents the results of the data analysis for the hypotheses associated with each of the research questions in the study.

Descriptive Statistics

Fourth grade students of District XYZ who participated in the ORF (R-CBM) and reading comprehension (MAZE) interim assessments in the fall, winter and spring and the summative Communication Arts MAP assessment in 2014 comprised the study sample. In the 2013-2014 school year, there were 1,382 students enrolled in District XYZ (MO DESE, 2016c). Students comprised a range of ethnicity (see table 12). Of those students, 1,227 students took the fall, winter, and spring ORF interim assessments, as well as the Communication Arts MAP test, and participated in the study. Therefore, 155 students were not included due to missing scores from one or more assessments. Study participants comprised 55 percent male (*n*=670), 45 percent female (*n*=557).

Table 12

2014 Ethnicity Data of Students in District XYZ for Grade 4

Nationality	*N*	Percent
Asian	22	1.5
Black	205	15
Hispanic	66	5
White	1053	76
Other	36	2

Note: Adapted from District MAP Disaggregate Data, Missouri State Department of Education, 2014, http://mcds.dese.mo.gov/quickfacts/Pages/District-and-School-Information.aspx

In addition, 1,246 students took the fall, winter, and spring CBM reading comprehension assessments, as well as well as the Communication Arts MAP test, and participated in the study. Therefore, 140 students were not included due to missing scores from one or more assessments. Study participants in reading fluency comprised 54 percent male (*n*=678), and 46 percent female (*n*= 569).

Student oral reading fluency and reading comprehension score is based on performance-level percentiles based on AIMSweb national norms. Those students whose score ranges in the ninety-ninth or seventy-fifth percentile have a high probability of meeting proficiency standards. While students whose score ranges in the twenty-fifth percentile have a moderate probability of meeting grade level proficiency, students

performing between the twenty-fifth and tenth percentile are at a moderate to high risk of not meeting grade level proficiencies. Students who score at the tenth or below percentile are at high risk of failure and are recommended to receive instructional support (NCS, Pearson, 2012b, p. 2).

Hypothesis Testing

This section includes the method used to test each hypothesis and the results of the hypothesis testing of the two research questions. Each of the research questions are followed by the corresponding hypothesis statement. The summary results from a multiple stepwise regression model were used determine an archetypal for predicting students score on the state MAP assessment. The level of significance for each multiple regression model was set at .05.

RQ1. What combination of variables [fall Oral Reading Fluency (R-CBM) score, winter Oral Reading Fluency (R-CBM) score, spring Oral Reading Fluency (R-CBM) score] best predicted students Communication Arts MAP scale scores?

H1. A significant combination of variables were identified for predicting student scores on the fourth grade MAP Communication Arts assessment included fall oral reading fluency (R- CBM) score, winter oral reading fluency (R-CBM) score, and spring oral reading fluency (R-CBM) score.

To test H1, a multiple linear regression was calculated predicting subject's MAP score based on students fall, winter, and/or spring oral reading fluency assessment scores. Two significant regression equations were found. First, (F (2, 1225) = 541.8, p <.001),

with a R^2 of .468. The equation predicted MAP is equal to .287 (oral reading fluency fall score) + .310 (oral reading fluency spring score) + 585.00. Subjects' average MAP score increased .376 for an increase of one point on the spring oral reading fluency assessment, and .287 for an increase of one point on the fall oral reading fluency assessment with an SEE of +/- 24.8 points upper level and lower level. In other words, the combination of fall and spring oral reading fluency score predicated a forty-seven percent variability on the MAP score (r = .469) and demonstrated a significance of $p < .001$. The fall and spring interim R-CBM were moderate to strong predictors of the MAP Communication Arts score, which supports hypothesis one.

The multiple regression analysis also determined the spring oral reading fluency interim assessment predicted a forty-five percent variability (F (1,1226) = 1011.4, p <.001), with of .452. The equation predicted MAP is equal to .310 + 582.00 with an SEE of +/- 25.2 points upper level and lower level. In other words, the spring oral reading fluency score significantly predicated a forty-five percent variability on the MAP score (r = .452) and demonstrated a significance of $p < .001$. These results also support hypothesis one, df = 1226, p = .001. The winter oral reading fluency was not included because it did not increase the predictive value of the equation. Table 13 shows the multiple regression analysis of oral reading fluency (R-CBM).

Table 13.

Model Summary, ANOVA & Coefficients Analysis Results for R-CBM

Source	*Beta*	*r. sq.*	*SEE*	*F*	*Sign*
Constant	582.7				
Fall	.287				
Spring	.310	.468	24.84	541.8	.000

Note 1: MAP is the Dependent Variable

Note 2: MAP = .287 (R-CBM Fall Score) + .310 (R-CBM Spring Score)+ 585.00

RQ2. What combination of variables (fall Oral Reading Fluency MAZE score, winter Oral Reading Fluency MAZE score, spring Oral Reading Fluency MAZE score) best predict students Communication Arts MAP scale scores?

H2. A significant combination of variables was identified for predicting student scores on the fourth grade MAP Communication Arts assessment includes fall MAZE CBM score, winter MAZE CBM score, and spring MAZE CBM score.

To test H2, a multiple linear regression was calculated predicting subject's MAP score based on students' fall, winter, and/or spring reading comprehension assessment scores. One significant regression equation was found; (F (3, 1245) = 355.3, p <.001), with a R^2 of .461. The equation predicted MAP is equal to 1.05 (reading comprehension winter score) + 1.15 (reading comprehension fall score) + .765 (reading comprehension spring score) + 602.66. Subjects' average MAP score increased .275 for an increase of one point on the winter reading comprehension assessment, and .246 for an increase of

one point on the fall reading comprehension and .210 for an increase of one point on the spring reading comprehension assessment, with an SEE of +/- 24.3 points upper level and lower level. In other words, the combination of winter, fall and spring reading comprehension scores predicated a forty-six percent variability on the MAP score (r = .462) and demonstrated a significance of $p <$ .001. The combination of winter, fall and spring interim MAZE comprehension assessments were significant predictors of the MAP Communication Arts score *df=1246*, p = .001, which supports hypothesis two. Table 14 shows the multiple regression analysis of reading comprehension.

Table 14.

Model Summary, ANOVA & Coefficients Analysis Results of R-CBM

Source	*Beta*	*r. sq.*	*SEE*	*F*	*Sign*
Constant	602.7				
Winter	1.06				
Fall	1.15				
Spring	.765	.461	24.345	355.3	.001

Note 1: MAP is the Dependent Variable

Note 2: MAP = 1.06 (winter reading comprehension score) + 1.15. (fall reading comprehension score)+ 0.765 (spring reading comprehension score) + 602.66.

Summary

This chapter included descriptive statistics and the results of hypothesis testing for this study. Multiple stepwise regression model were used determine an archetypal for

predicting students score on the state MAP assessment. The data analyses show that both oral reading fluency (R-CBM) and reading comprehension (MAZE CBM) have statistically significant relationships to the state MAP assessment. However, not all the variables were significant predictors of MAP scores. Chapter five includes an overview of the study, major findings, findings related to the literature, implications for action, recommendations for future research, and concluding remarks.

Chapter Five

Interpretation and Recommendations

Chapter five includes a study summary, the findings of this study as they relate to the literature review, conclusions and concluding remarks. This researcher provides an overview of the study, observations, and interpretations of findings, as well as ideas for future research.

Study Summary

Response to Intervention provided an organized environment to use curriculum based measurement interim assessments to tract student reading progress or lack thereof. This study sought to determine the relationship of frequent monitoring of student progress using CBM, and student success on the state-wide summative assessment within an RTI framework.

Overview of the problem. It is a well-known fact that over the last two decades, American students have made little progress in reading. According to the most recent research from the National Center of Educational Statistics, fourth grade student's reading ability remains stagnant. They state,

> In 2015, the percentage of 4th-grade students performing at or above the Basic achievement level (69 percent) was not measurably different from the percentage in 2013, but it was higher than the percentage in 1992 (62 percent). In addition, the percentage of 4th-grade students performing at or above the Proficient achievement level in 2015 (36 percent) was not measurably different from the

percentage in 2013, but it was higher than the percentage in 1992 (29 percent). (2016, p. 143)

In other words, 36 percent of students in our nation are proficient readers at fourth grade. As a result, lawmakers, educational experts, and researchers have established new laws, revised curriculum, updated teacher evaluations, and student assessment practices, to determine school and district effectiveness. These changes often embraced a Response to Intervention structure that comprises the use of frequent monitoring of student progress.

In order to determine whether school districts are meeting the required level of performance, the state of Missouri mandates participation in MAP standardized summative assessment each year. Between the years 2009 and 2013, students in the state and District XYZ were reading 37 to 49 percent below the rate of their age-appropriate peers.

In the 2013-2014 school year, District XYZ initiated Response to Intervention (RTI) in all eighteen elementary schools to meet the needs of the lowest functioning, academic students in math and reading (District XYZ, Five Year Plan, 2013). At the same time, District XYZ utilized Curriculum-based Measurements (CBM) several times during the course of a school year to provide evidence of student achievement (District XYZ, RTI manual, 2013). It is the desire of the district to ensure academic success for all students.

Purpose statement and research questions. The purpose of this study was to determine the extent of the relationship between fourth grade scores on AIMSweb interim benchmarks assessments scores (R-CBM and MAZE CBM) and that of the Communication Arts MAP assessment scores for students enrolled in District XYZ. It is important to ensure that AIMSweb R-CBM and MAZE CBM scores are predictors of the Communication Arts MAP assessment scores. Two research questions were developed to determine if any combination of student scores from the fall, winter, and spring R-CBM oral reading fluency and MAZE comprehension interim assessment scores predicted the Communication Arts portion of the statewide summative MAP assessment.

Review of the methodology. Data analyzed for this quantitative study were collected from a large mid-western suburban school district over one school year, from August, 2013 to May, 2014. Fourth grade students from District XYZ, who participated in the fall, winter, and spring AIMSweb R-CBM and MAZE interim assessments and the MAP Communication Arts test comprised the sample for this study. Multiple regression analyses were used to find the combination of variables that best predicted MAP scores for fourth grade. The hypotheses stated that a large combination of independent variables (fall, winter, and spring CBM and fall, winter, and spring MAZE) scores were predictors of the MAP Communication Arts assessment scores. A stepwise regression model analysis was constructed using the fourth grade students in district XYZ and subtests (communication arts). Correlations were examined for the strength of the associations

between each of the predictor variables and multiple stepwise regression analysis were used to predict student achievement on the MAP test.

Major findings. The evidence provided by the multiple regression model for research question one supported the first hypothesis. Fall and Spring R-CBM oral reading fluency interim assessments were significant predictors of the MAP Communication Arts summative assessments scores. In other words, results from the multiple regression model suggested that fourth grade students who have higher scores on fall and spring oral fluency assessments are more likely to have higher scores on the MAP Communication Arts test.

The evidence provided by the multiple analysis results for research question two supported the second hypothesis. Fall, winter, and spring MAZE comprehension interim scores were significant predictors of the MAP Communication Arts summative data assessment scores. Therefore, the multiple regressions suggests that fourth grade students who have higher scores on fall, winter, and spring comprehension assessments are more likely to have higher scores on the MAP Communication Arts test.

Findings Related to the Literature

The goal of this study was to determine the predictive relationship between the R-CBM ORF and the MAZE comprehension interim scores and the Missouri MAP summative state-wide test scores in a RTI framework. The multiple regression model found evidence that the fall and spring R-CBM ORF, and the fall, winter , and spring MAZE comprehension combination were positive strong predictors of the on the MAP

Communication Arts test scores. This section relates the findings of this study to research presented in chapter two.

Good, Simmons and Kame'enui's (2001) research comprised a wholistic study of emerging reading skills of students in kindergarten through third grade. Good and his colleagues used CBM ORF General Outcome Measures (GOM) to determine the efficacy of student skill at their determined grade level and then as a predictor of performance over time and on the Oregon Statewide Assessment (OSA). Using CBM ORF for first through third grade, Good et al., determined that a strong correlation exists between the CBM ORF at each grade level and the spring R-CBM ORF scores were predictive of the OSA for the third grade students. The students who demonstrated proficiency each year on established benchmark goals were highly likely to be proficient on the OSA. The current study used R-CBM ORF to determine the predictability on the Missouri Communication Arts MAP test. However, instead of using only the spring R-CBM ORF as Good et al., the current study compared a fall, winter and spring RCBM for fourth grade students. While the current study found that the the fall and spring R-CBM ORF were the strongest predictors of the Missouri Communication Arts MAP Assessment, it did agree with Good et al., (2001) that the spring R-CBM had a positive significance. In addition, the current study included only fourth grade students. Overall, the current study supports that R-CBM ORF is a predictive measure of state-wide assessments.

Stage and Jacobsen (2001) conducted a research study to identify the best predictor of the Washington Assessment of Student Learning (WASL) assessment from

September, January, and May oral reading fluency interim assessments. The study comprised 173 fourth grade students. The multiple regression analyses indicate a positive significant level of ORF performance (September, January, and May) predicted the WASL reading performance (Stage & Jacobsen, 2001). The current study indicated a positive significance on the fall and spring R-CBM ORF assessments. The winter R-CBM ORF was found not to increase the predictive validity and therefore was eliminated. The current study supports the conclusion that oral reading fluency is a valid predictor of state-wide summative reading tests.

A study by McGlinchey & Hixon (2004) investigated the predictive value of fourth grade oral reading fluency for performance on a state reading assessment, the Michigan Educational Assessment Program (MEAP). Although this study replicated that of Stage and Jacobsen (2001), McGlinchey et. al study comprised a large sample (n=1,362) across eight years with a diverse population. Results indicate a moderately strong relationship between oral reading rates and MEAP performance. The current study agrees with the findings of McGlinchey and Hixson (2004) that R-CBM scores are predictors of high-stakes summative assessment scores.

Likewise, Shapiro, Keller, Lutz, Santoro, & Hintze, (2006) researched the predictive relationship between CBM and the Pennsylvania System of School Assessment (PSSA) and the MAT-8, a general standardized reading assessment used for fourth grade students. Their study comprised third, fourth and fifth grade students in the context of two school districts with a combined sample size of 1,048 students. One

school district comprised a mixture of urban and suburban school while the other comprised only suburban schools. Shapiro, et. al (2006) report that all correlations (fall, winter, and spring R-CBM ORF) for third and fifth grade predict a moderate to strong relationship with the PSSA, with correlations closely approaching .70 (Shapiro, et al., 2006). Further, the correlation between MAT-8 and the CBM ORF fall, winter, and spring demonstrated a positive moderate to strong relationship for fourth grade students (Shapiro, et al., 2006). The hierarchial regression analysis of their research showed that the winter assessment period the strongest predictor of the PSSA scores, with spring assessments not adding significantly to explanations of variance contributing to outcomes. Comparatively, the current study found the fall and spring R-CBM ORF scores were strong predictors of the MAP assessment. The winter assessment did not increase the predictive value of the equation and was therefore eliminated. However, this study is in agreement with the findings of Shapiro et al., (2006) that R-CBM ORF scores are predictors of state-wide summative assessments.

Researchers Hintze and Silberglitt's (2005) study replicated and extended the research on the predictive validity of R-CBM ORF with first through third-grade students on the Minnesota Comprehensive Assessment (MCA). Hintze and colleagues study resulted in the predictive ability of R-CBM ORF to the MCA as early as first grade using three statistical methodological approaches to standard setting and determine cut scores using R-CBM and performance on high-stakes testing with third-grade students. This study indicated that the use of R-CBM is a strong predictor of more global measures of

reading as in high-stakes standardized state assessments. Further, the R-CBM ORF was more strongly correlated with MCA when the two or more assessments were in close proximately of each other as compared to further apart. Interestingly, the current study found that the strongest predictors of performance on MAP Communication Arts assessment were fall and spring R-CBM ORF. The winter R-CBM ORF was found not to increase the predictive validity and therefore was eliminated. However, the current study supports Hintze and Silberglitt's (2005) study regarding the use of R-CBM is a strong positive predictor of MAP Communication Arts assessment.

Research question 2 focused on the MAZE comprehension scores of fourth grade students. Silberglitt, Burns, Madyun, and Lail (2006) researched the relationship between the Minnesota Comprehensive Assessments-Reading (MCA-R), and the R-CBM ORF and MAZE comprehension assessments for grades 3, 5, 7, and 8 to determine the function of grade level differences. Fall, winter, and spring CBM assessments for both R-CBM and MAZE were provided to a sample of 5,472 students, in five districts, over mulitple years. Silberglitt et al., (2006) found that all correlations between the R-CBM ORF and MAZE comprehension were positively significant to the MCA-R test. The current study also found strong correlation between the both the RCB-M and the MAZE assessments and the Communcication Arts MAP test. Specifically, the current study found that the fall, winter, and spring combination of MAZE tests had positive moderate to strong results to the Missouri MAP Communication Arts test. Therefore, this study

agrees with the findings of Silberglitt et al., (2006) that CBM ORF and MAZE comprehension are predictors of state-wide assessments.

Wiley and Deno (2005) researched in an urban area for both R-CBM ORF and MAZE with English Language Learners (ELL) and non- ELL students as a predictive study of a state test in Minnesota. The results of this study confirm that R-CBM ORF and MAZE are significant predictors of the performance on the Minnesota Comprehensive Assessment (MCA) for non-ELL students. However, for ELL students the MAZE did not add to the prediction of success. The current study, determined that the R-CBM ORF and the MAZE comprehension assessments were moderate to strong predictors of the Communication Arts MAP assessment. Although the current study does not draw out ELL and non-ELL subgroups, Wiley and Deno (2005) does give information related to how students may perform on the assessment when sorted by these subgroups. Therefore, the current study supports Wiley and Deno (2005) for R-CBM ORF and for the MAZE comprehension of non-ELL students. This study contradicts the results of MAZE comprehension with ELL students on state-wide assessments.

Conclusions

As stated in chapter one, it is important to know if benchmark interim assessment scores align with that of state summative assessment scores. If the assessments are in alignment, then benchmark interim assessments would enable educational leaders and teachers to determine a student's academic deficit, in order to provide immediate intervention. The current study's focus included interim and summative assessments of

student achievement in communication arts. Implications for actions and recommendations for future research are included in this section based on the findings of this study.

Implications for action. The data for this research came from District XYZ, a large suburban school district in the mid-west. The district researched may want to segregate students demographically to determine whether there are significant discrepancies among them. Further, this research was performed the first year of implementation of RTI and AIMSweb across all eighteen elementary schools. It would be interesting to see if there are any differences after years of implementation.

However, the present study has much to offer school districts, educational leaders as well as classroom teachers' information working in the Midwest, specifically the state of Missouri with information regarding the influence of CBM oral reading fluency and comprehension interim assessments on individual student progress, instruction, and overall reading growth and development in a RTI framework. Educational leaders and teachers should use the results of this study and the review of literature to make decisions on formative, interim, and summative assessments, to predict individual students MAP scores, and work to remediate him/her before the test is given.

Recommendations for future research. This study added to the research related to the use of CBM oral reading fluency and comprehension interim assessments embedded in a framework of RTI. At the time of this study, there has been new legislation, and action taken to further the academic success of all students. This has

been seen from the reauthorization of IDEA which suggested the use of RTI, adoption of common core national standards, updated teacher evaluations and improvements in teacher preparation. Teachers are asked to work together with each other and administrators to analyze data, use instructional strategies that differentiate, encourage problem-solving, and higher level thinking, as well as to create formative, interim and summative assessments, all as indicators of true student outcomes.

This study is the ground work for future studies for the development of formative, interim and summative assessments that provide on-going information for students, teachers, parents and educational leaders. Ideas to take this study further would be to replicate and extend current finding in more diverse settings, over longer periods of time and with segregated demographics of the student body. The opportunity to apply, extend, replicate, and refine what has been learned in this study is of significant relevance and promise as we continue to determine the elements of an assessment and intervention system necessary to improve reading outcomes for all students.

Further studies that examine the longitudinal predictive power of CBM are needed. This would establish benchmarks for all grade levels that provide prediction on later high-stakes summative assessments. Some research has already begun to establish longitudinal evidence of 1-minute reading samples obtained at the end of first grade are predictive of performance on state-wide standardized testing at the end of third grade (e.g., Hintze & Silberglitt, 2005; Keller & Shapiro, 2005).

Additionally, replicating this study in the context of a RTI framework would continue to provide the evidence of differentiation, use of on-going assessment to remediate students when they first begin to struggle, and to monitor student progress to ensure that all students are provided an enriching education. RTI research has long been established as a format of educators and school leaders integrate assessment and intervention to maximize student achievement and reduce behavior problems and finally to identify those with learning disabilities.

Concluding remarks. This study provided evidence of the importance of CBM oral reading fluency and MAZE comprehension interim assessments as they relate to student reading development for fourth-grade students. A potential implication of this study is the possible use that CBM may have in serving as an effective predictor-outcome on statewide assessment measures. Although the predictive power of the CBM measure did contain a number of false positive and false negatives, the measure could offer districts an inexpensive and efficient mechanism to potentially identify a large group of students who were at risk for not being successful on the statewide assessment measure. Such identification would support an intensive, short-term remediation program focused on teaching students the skills and competencies needed to be successful on the statewide assessment. Given the high-stakes importance of statewide assessments, remediation efforts would be important for students as well as districts.

Also important to note is the efficiency of these CBM measures as compared to norm-referenced achievement tests. Given the expense and time required to administer

norm-referenced achievement tests, CBM offers a potentially inexpensive way for districts to do large-scale screening. In addition, because norm-referenced achievement tests never designed to be responsive to short-term instructional interventions, the use of CBM measures can serve the added purpose of allowing teachers to monitor student performance across time if an intervention plan targeted at students who are at risk for failing statewide assessment measures is implemented. The results of this study suggest that the CBM measures in reading can serve as powerful predictors of students who are likely to be unsuccessful on the statewide assessment. Another implication of this study is the relationship between CBM measures and acquisition of state standards. The moderate to strong correlations between these measures suggests that the acquisition of state standards through the instructional process is reflected in the CBM measure.

References

Bahr, M. W., & Kovaleski, J. F. (2006). The Need for Problem-Solving Teams. *Remedial & Special Education, 27(1),* 2-5.

Barnett, D. W., Daly, E. J., Jones, K. M., & Lentz, F. E. (2004). Response to intervention: Empirically based special service decisions from a single-case design of increasing and decreasing intensity. *The Journal of Special Education, 38,* 66-79.

Barnhardt, J., (2009). *Response to intervention in rhode island: A case study of two elementary schools at different stages of RTI implementation* (Doctoral dissertation). Retrieved from google (3357418)

Batsche, G., Elliott, J., Graden, J. L., Grimes, J., Kovaleski, J. F., Prasse, D., et al. (2005). *Response to intervention: Policy considerations and implementation.* Washington, DC: NASDSE

Berdine, W. H. (2003). The president's commission on excellence in special education: Implications for the special education practitioner. *Preventing School Failure 47(2),* 92-95. Washington: D.C.: Heldref Publications. doi.org/10.1080/1045988 0309604436. Retrieved from http://www.ldonline.org/article/6364/

Bergan, J. R. (1977). *Behavioral Consultation.* Columbus, OH: Charles E. Merrill.

Bergan, J. & Kratochwill, T. R. (1990). *Behavioral Consultation and Therapy.* New York: Plenum Press.

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

Black, P., & William, D. (1998) Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice, 5:1,* 7-74, doi:10.1080/0969595 980050102

Bransford, J.D., & Stein, B.S. (1993). The IDEAL Problem Solver (2nd Edition). W.H. Freeman: New York.

Bradley, M. C., Tamara Daley, Marjorie Levin, Fran O'Reilly, Amanda Parsad, Anne Robertson, and Alan Werner. (2011). *IDEA National Assessment Implementation Study. NCEE 2011-4027.* Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Brigance, A. H. (1999). *Comprehensive Inventory of Basic Skills—Revised*. North Billerica, MA: Curriculum Associates.

Brown-Chidsey, R., Johnson, P. J., & Fernstrom, R. (2005). Comparison of Grade-Level Controlled and Literature-Based Maze CBM Readin Passages. *School Psychology Review, 34(3),* 387-394.

Brown-Chidsey, R., & Steege, M. W. (2010). *Response to intervention; Principles and strategies for effective practice.* New York, NY: Guilford Press.

Buffum, A., Mattos, M., Weber, C. (2009). *Pyramid response to intervention: RTI, professional learning communities, and how to respond when kids don't learn.* Bloomington, IN: Solution Tree Press

Burns, M. K. (2010a). Response-to-intervention research: Is the sum of the parts as great as the whole? *Perspectives on Language and Literacy,36(2),* 5-11. Distributed by the RTI Action Network, Washington, DC. Retrieved from http://www.rtinetwork.org/learn/research/response-to-intervention-research-is-the-sum-of-the-parts-as-great-as-the-whole

Burns, M. K. (2010b). Formative evaluation in school psychology: Fully informing the instructional process. *School Psychology Forum: Research in Practice,4(1), 22-23.*

Burns, M. K., Appleton, J. J., & Stehouwer, J. D. (2005). Meta-analytic review of responsiveness-to-intervention: Examining Field-based and research implemented models. *Journal of Psychoeducational Assessment, 23(4),* 381-394. doi: 10.1177/073428290502300406

Burns, M. K., Coolong-Chaffin, M. (2006). Response to intervention: The role of and the effect on school psychology. *School Psychology Forum [University of Minnesota] 1,* 3-15. Retrieved from http://www.uwec.edu/Psyc/graduate /upload/RTI-Coolong-Chaffin.pdf

Burns, M.K, Deno, S., & Jimerson, S.R. (2007). Toward a unified model of Response to Intervention. In Jimerson, S.R., Burns, M.K, & VanDerHeyden, A.M. (Eds.), The Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention (pp. 428-440). New York: Springer

Burns, M. K., Vanderwood, M. L., & Ruby, S. (2005). Evaluating the readiness of pre-referral intervention teams for use in a problem-solving model. *School Psychology Quarterly, 20,* 89-105. doi:10.1521/sepq.20.1.8964192

Burns, M. K., & Ysseldyke, J. E. (2005). Comparison of existing responsiveness-to-intervention models to identify and answer implementation questions. *The California School Psychologist, 10,* 9-20.

Carney, K. J., & Steifel, G. S. (2008). Long-term results of a problem-solving approach to response to intervention: Discussion and implications. *Learning Disabilities: A Contemporary Journal, 6(2),* 61-75.

Case, L. P., Speece, D. L., & Molloy, D. E. (2003). The Validity of a Response-to-Instruction Paradigm to Identify Reading Disabilities: A Longitudinal Analysis of Individual Differences and Contextual Factors. *School Psychology Review, 32*(4), 557-582

Center on Innovation and Improvement. (n.d.). Monitoring fidelity of implementation. Retrieved from http://www.centerii.org/handbook/resources/5_g_monitoring _fidelity.pdf

Chalfant, J.C., Pysh, M. V., & Moultrie, R. (1979). Teacher assistance teams: A model for within-building problem solving. *Learning Disability Quarterly, 2,* 85-96.

Christ, T. J., Burns, M. K., & Ysseldyke, J. E. (2005). Conceptual confusion within response-to-intervention vernacular: Clarifying meaningful differences. *Communiqué, 34(3),* 1–2

Compton, D. L., Fuchs, D., Fuchs, L., Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of educational psychology, 98(2),* 394-409. doi:10.1037/0022-0663.98.2.394

Connor, C.M., Alberto, P.A., Compton, D.L., O'Connor, R.E. (2014). *Improving reading outcomes for students with or at risk for reading disabilities: A synthesis of the contributions from the institute of education sciences research centers (NCSER 2014-3000).* Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. This report is available on the IES website at http://ies.ed.gov/.

Coulombe, S. & Tremblay, J.F. (2006). Literacy and growth. *The B.E. Journal of Macroeconomics, 8(2), 1-34.* Retrieved from http://www.sba.muohio.edu/davisgk/growth%20readings/10.pdf

Coyne, M. D., Kaméenui, E. J., & Simmons, D. C. (2004). Beginning Reading Intervention as Inoculation or Insulin: First-Grade Reading Performance of Strong Responders to Kindergarten Intervention. *Journal of Learning Disabilities, 37(2),* 90-104. doi:10.1177/00222194040370020101

CTB/McGraw-Hill. (2014a). *Missouri achievement program: Guide to interpreting your results.* Monterey, CA: Author. Retrieved from http://dese.mo.gov/sites/default/files/asmt-gl-gir-spring-2014.pdf

CTB/McGraw-Hill. (2014b). *Missouri grade level assessments technical report 2010.* Monterey, CA: Author. Retrieved from http://dese.mo.gov/sites /default/files/asmt-gl-2014-tech-report.pdf

Daly III, E. J., Martens, B. K., Witt, J. C., & Olson, S. C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology Review, 36,* 562-581.

Del Siegle, D. (2002). *Instrument Reliability.* Neag School of Education-University of Connecticut website. Retrieved from http://researchbasics. education.uconn.edu/ instrument_reliability/

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37,* 184-192. doi:10.1177/0022466 903 0370030801

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

Deno, S. L., Fuchs, L. S., Marston, D., & Shinn, J. (2001). Using curriculum based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 34,* 506-522.

Denton, C. A., Vaughn, S., & Fletcher, J. M. (2003). Bringing research-based practice in reading intervention to scale. *Learning Disabilities Research & Practice, 18(3),* 201-211.

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of tier 3 intervention for students with persistent reading

difficulties and characteristics of inadequate responders. *Journal of Educational Psychology, 105(3).* 633-648. doi: 10.1037/a0032581

Dewey, J. (1910). *How we think.* New York, NY: D.C. Heath and Company.

District XYZ. (2013). District 5 year plan.

District XYZ. (2013). Assistant Superintendent of Elementary Education Meeting.

District XYZ. (2014). District RTI Manual.

District XYZ. (2014). District Intervention Manual.

District XYZ. (2014). District Curriculum and Instruction Page. Retrieved on July, 28 from https://sites.google.com/a/lsr7.net/r7rti/?pli=1

Elementary and Secondary Education Act of 1965, 20 U.S.C. § 6301.

Fletcher, J. M. (1995). *Diagnostic utility of intelligence testing and the discrepancy model for children with learning disabilities: Historical perspectives and current research.* Paper presented at the IQ Testing and Educational Decision Making Workshop, National Research Council, National Academy of Sciences, Washington, D.C.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., Francis, D. J. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86,* 6-23

Flordia Center for Reading Research. (2016). Current research projects. Retrieved from http://www.fcrr.org/about-fcrr/projects.asp

Foorman, B. R., Francis, D. J., & Fletcher, J. M. (1995). *Growth of phonological processing skill in beginning reading: The lag versus deficit model revisited.* Paper presented at the Society for Research on Child Development. Indianapolis, IN.

Fuchs, L. S. (2007, Winter). *NRCLD update on responsiveness to intervention: Research to practice.* National Research Center on Learning Disabilities. Retrieved from files.eric.ed.gov/fulltext/ED543485.pdf

Fuchs, D., & Fuchs, L. S. (1996). Consultation as a technology and the politics of school reform to the isssue. *Remedial and Special Education, 17(6),* 386-392.

Fuchs, L., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice, 13,* 204–219.

Fuchs, D., & Fuchs, L. S. (2001). Responsiveness-to-intervention: A blueprint for practitioners, policymakers, and parents. *Teaching Exceptional Children, 38,* 57-61

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41(1),* 93-99. doi:10.1598/RRQ.41.1.4

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25(1),* 33-45.

Fuchs, D., Mock, D., Morgan P. L., & Young, C. L. (2003). Responsiveness to intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice*, 18, 151-171. doi:10.1111/1540-5826.00072

Fuchs, L. S., & Vaughn, S. R. (2005). Response to intervention as a framework for the identification of learning disabilities. *NASP Communique, 34(1),* 4-6.

Fuchs, D., & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice, 22,* 129 – 136

Gersten, R., & Dimino, J. A. (2006). RTI (Response to Intervention): Rethinking special education for students with reading difficulties (yet again). *Reading Research Quarterly, 41(1),* 99-108. doi:10.1598/RRQ.41.1.5

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to intervention and multi-tier intervention for reading in the primary grades. A practice guide (NCEE 2009-4045).* Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/ practiceguides/

Gleason, T. L. (2013). *The Impact of RTI Practices on Student Reading Achievement.* (Doctoral Dissertation). Retrieved from https://dspace.iup.edu/handle/2069 /2006. (UMI Number: 3558733).

Good, R. H. III, Gruba, J., & Kaminski, R. A. (2002). Best practices in using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an outcomes-driven model. In A Thomas and J. Grimes (Eds.) *Best practices in school psychology IV,* Volume 1 (pp. 699- 720). Bethesda, MD: National Association of School Psychologists.

Good, R. H., III, Simmons, D. C., & Kame'enui, E. J. (2001). The importance of decision making utility of a continuum of fluency-based indicators of foundational reading skills for third grade high stakes outcomes. *Scientific Studies of Reading, 5*, 257–288. doi:10.1207/S1532799XSSR0503_4

Gresham, F. M., (2007). Evolution of the response-to-intervention concept: Empirical foundations and recent developments. In S.Jimmerson, M. Burns, & A. VanderHeyden (Eds), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp.10-24). New York: Springer.

Gresham, F. M., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessments: Principals, procedures, and future directions. *School Psychology Review,* 30*(2),* 156-172.

Hallahan, D. P., Mercer, C. D. (2001). Paper Presented at Learning Disabilities Summit: Building a Foundation for the Future. *Learning disabilities: historical*

perspectives. executive summary. Sponsered by: Special Education Programs Education/Office of Special Education and Rehabilitative Services (ED/OSERS). Washington D.C., 2001. Retrieved from http://www.air.org/ldsummit/

Hanushek, E.A., Jamison, D.T., Jamison, E.A., & Woessmann, L. (2008). Education and economic growth. *Journal of Education Next, 8(2),* 62-70. Retrieved from http://taubcenter.org.il/tauborgilwp/wp-content/uploads/

Heller K. A., Holtzman W. H., Messick S. (1982). *Placing children in special education: A strategy for equity.* Governing Board of the National Research Council. Washington, DC: National Academy Press.

Hermanowicz, H. J. (1961). *Problem solving as teaching method. Educational Leadership.* 18299. Retrieved from http://ascd.com/ASCD/pdf/journals/ed_lead /el_196102_hermanowicz.pdf

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

Hosp, J. L., & Madyun, N. H. (2007). Addressing disportortionality with response to intervention. In S. R. Jimerson, M. K. Burns, & A. VanDerHeyden (Eds.) *Handbook of response to intervention: The science and practice of assessment and intervention* (p 172-181). New York, NY: Springer

Howe, K. B., & Shinn, M. M. (2002). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Eden Prairie, MN: Edformation

Hughes, C. A., & Dexter, D. D. (2008). *Universal screening within a response-to-intervention model.* RTI Action Network Web, National Center for Learning Disabilities, Retrieved from http://www.RTInetwork.org/learn/ research/use-RTI-identify-students-learning-disabilities-review-research

Huitt, W. (1996). Measurement and evaluation: Criterion- versus norm-referenced testing. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from http://www.edpsycinteractive.org /topics/ measeval/crnmref.html

Hursh, D. (2007). Exacerbating inequality: the failed promise of the No Child Left Behind Act. *Race, Ethnicity and Education, 10,* 295-308.

Individuals With Disabilities Education Act of 2004, 20 U.S.C. § 1400 *et seq.* (2004).

International Reading Association. (2010). *Response to intervention: Guiding principles for educators from the international reading association.* Retrieved from International Reading Association: www.reading.org

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for service delivery in an RTI framework: Candidate measures. *School Psychology Review, 36,* 560–582.

Jenkins, J. R., & Johnson, E. S. (2008). *Universal screening for reading problems: Why and how should we do this?* Retrieved from http://www.RTInetwork.org/ Essential/Assessment/Universal/ar/ReadingProblems

Jenkins, J. R., & O'Connor, R. E. (2002). Early identification and intervention for young children with reading/learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 99–150). Mawah, NJ: Erlbaum.

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2007). Response to intervention at school: The science and practice of assessment and intervention. New York, NY: Springer.

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2007). *Handbook of response to intervention: The science and practice of assessment and intervention* (Ed.) New York, NY: Springer

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

Johnson, E. S., Pool, J. L., & Carter, D. (2009). *Implementing a screening process within an RTI framework: A review of methods and instruments K-12.* National Center on Learning Disabilities RTI Action Network. Retrieved from http://www.RTInetwork.org/essential/assessment/screening /screening-for-reading-problems-in-an-RTI-framework

Jones, R.E., Yssel, N., & Grant, C. (2012). Reading instruction in tier 1: Bridging the gaps by nesting evidenced-based interventions within differentiated instruction. *Psychology in the Schools, 49(3), 210-218.* doi: 10.1002/pits.21591

Kaminski, R., & Cummings, K. D. (2007). *Dibels: Myths and facts*. Dynamic Measurement Group. Retrieved from https://dibels.org/papers/Myths_0208.pdf

Kamps, D. & Greenwood, C. (2003, December). Formulating secondary level reading interventions. Paper presented at the Responsiveness-to-Interventions Symposium, Kansas City, MO

Kavale, K. A., Kauffman, J. M., Bachmeier, R. J., & LeFever, G. B. (2008). Response-to-Intervention: Separating the rhetoric of self-congratulation from the reality of specific learning disability identification. *Learning Disability Quarterly, 31,* 135-150. doi:10.2307/25474644

Kim, Y. S., Apel, K., & Al Otaiba, S. (2013). The relation of linguistic awareness and vocabulary to word reading and spelling for first-grade students participating in response to intervention. *Language, Speech and Hearing Services in Schools, 44* 337-347.

Kirk, S., & Kirk, W. (1983). On defining learning disabilities. *Journal of Learning Disabilities, 16(1),* 20-21.

Klotz, M. B., Canter, A., 2007. Response to intervention (RTI): A primer for parents. *National Association of School Psychologists.* Retrieved July 21, 2015 from http://www.nasponline.org/resources/factsheets/RTIprimer.aspx

Kosanovich, M., Ladinsky, K., Nelson, L., & Torgensen, J. (2007). Differentiated reading instruction: Small group alternative lesson structures for all students. Flordia Center for Reading Research. Retrieved August 9, 2015 from http://files.eric.ed.gov/fulltext/ED498777.pdf

Kovaleski, J. F. (2002). Best practices in operating pre-referral intervention teams in Pennsylvania. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-IV* (pp. 645-655). Berthesda, MD: National Association of School Psychologists.

Kovaleski, J. F. (2003). *The three-tier model for identifying learning disabilities: Critical program features and system issues.* Paper presented at the National Research Center on Learning Disabilities (NRCLD) Responsiveness-to-Intervention Symposium, Kansas City, MO. Retrieved October 19, 2015 from http://www.nrcld.org/symposium2003/kovaleski/index.html.

Kovaleski, J. F. (2007a). Response to intervention: Considerations for research and systems change. *School Psychology Review. 36(4)* 638-644. Retrieved from http://blogs.gssd.ca/studentservices/files/2011/06/RTI-Considerations-for-Research-and-Systems-Change.pdf

Kovaleski, J. F. (2007b). Potential pitfalls of response to intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to Intervention: The science and practice of assessment and intervention* (pp. 223-233). New York, NY: Springer.

Kovaleski, J. F., & Glew, M. C. (2006). Bringing instructional support teams to scale: Implications of the Pennsylvania experience. *Remedial and Special Education, 27,* 16-25.

Kratochwill, T. R., & Shernoff, E. S. (2004). Evidence-based practice: promoting evidence-based interventions in school psychology. *School Psychology Review, 33(1)* 34-48. Retrieved from http://www.seattleimplementation. org/wp-content/uploads/2012/05/A_Kratochwill_Shernoff_Promoting EBIsSchoolPsych_EBPAS_2004.pdf

Kratochwill, T. R., Volpiansky, P., Clements, M., & Ball, C. (2007). Professional development in implementing and sustaining multitier prevention models: Implications for response to intervention. *School Psychology Review, 36,* 618-631.

Kruger, L. J., & Struzziero, J. (1995). The relationship between organizational support and satisfaction with teacher assistance teams. *Remedial & Special Education, 16(4),* 203.

Lembke, E. S., McMaster, K. L., & Stecker, P. M. (2010). The prevention science of reading research within a Response-to-Intervention model. *Psychology In The Schools, 47(1),* 22-35. doi:10.1002/pits.20449

Levy, F., & Murnane, R. J. (2012). *New Division of Labor: How Computers Are Creating the Next Job Market.* Princeton, NJ, USA: Princeton University Press. Retrieved from http://www.ebrary.com

Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Lunenburg, F. C., & Irby, B. J. (2008). *Writing a successful thesis or dissertation: Tips and strategies for students in the social and behavioral sciences.* Thousand Oaks, CA: Corwin Press.

Lyon, G. Reid. (1996). Learning disabilities. *The Future of Our Children: Special Education for Children with Disabilities,* (6)1, 54-76. doi:10.2307/1602494

Lyon, G.R., Fletcher, J.M. Shaywitz, S.E., Shaywitz, B.A., Wood, F.B., Schulte, A., Olson, R.K., & Torgesen, J.K. (2001). Learning disabilities: An evidence-based conceptualization. Rethinking special education for a new century (pp. 259-287). Washington, D.C.: Fordham Foundation and Progressive Policy Institute.

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

Marston, D., Muyskens, P., Lau, M., & Canter, A. (2003). Problem-solving model for decision-making with high-incidence disabilities: The minneapolis experience. *Learning Disabilities Research and Practice 18(3),* 187-200. doi: 10.1111/1540-5826.00074

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33(2),* 193-203.

McMaster, K. L., Wagner, D. (2007). Monitoring response to general education instruction. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to Intervention: The science and practice of assessment and intervention* (pp. 223-233). New York, NY: Springer.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.*Exceptional Children, 71(4),* 445-463.

Mellard, D.F., & Johnson, E. (2008). *RTI: A practitioner's guide to implementing response to intervention.* Thousand Oaks, CA: Corwin Press.

Missouri Department of Elementary and Secondary Education, (2008a) State of Missouri RtI Guidelines Appendix A. (June, 2008). Retrieved from http://dese.mo.gov/sites/default/files/RtIGuidelines08.pdf

Missouri Department of Elementary and Secondary Education, (2008b) Special Education Advisory Council Minutes. (September, 2008). Retrieved from http://dese.mo.gov/sites/default/files/AdvisoryPanelMinutes9_08.pdf

Missouri Department of Secondary and Elementary Education, (2014). Missouri School Directory. Retrieved from http://mcds.dese.mo.govguidedinquiry/

District%20and%20School%20Information/Missouri%20School%20Directory.aspx?rp:DistrictCode

Missouri Department of Elementary and Secondary Education. (2015). Missouri School Directory. 2014-2015 Statistics of Missouri Public Schools. Retrieved from http://mcds.dese.mo.gov/quickfacts/District%20and%20School%20Information/Missouri%20School%20Statistics.pdf

Missouri Department of Elementary and Secondary Education. (2016a). Missouri school improvement program: Support and intervention. Retrieved from http://dese.mo.gov/sites/default/files/MissouriSchoolImprovementPlan.pdf

Missouri Department of Elementary and Secondary Education, (2016b). Missouri assessment program grade-level assessments: A guide to interpreting results. Retrieved from http://dese.mo.gov/sites/default/files/asmt-gl-gir-spring-16v1.pdf

Missouri Department of Elementary and Secondary Education. (2016c). History of the Missouri Assessment Program. Retrieved from https://dese.mo.gov /college-career-readiness/assessment/lea-guide-missouri-assessment-program-2015-2016

Missouri Department of Elementary and Secondary Education. (2016d). About Missouri Learning Standards. Retrieved from http://www.missourilearningstandards.com/about/

Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., & ... Shaywitz, B. A. (1998). Subtypes of reading disability:

Variability around a phonological core. *Journal of Educational Psychology*, *90*(3), 347-373. doi:10.1037/0022-0663.90.3.347

Moss, C. M., & Brookhart, S. M. (2009). *Advancing formative assessment in every classroom: A guide for instructional leaders.* Danvers, MA: Association for Supervision and Curriculum Development.

National Association of State Directors of Special Education and Council of Administrators of Special Education & Council of Administrators of Special Education. (2006). *Response to Instruction.* Retrieved from http://www.casecec.org/pdf/RTI/RTI%20An%20Administrator's%20Perspective%201-061.pdf

National Association of State Directors of Special Education. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: National Association of State Directors of Special Education.

National Center for Education Statistics. (2014). *Reading performance.* Retrieved from http://nces.ed.gov/programs/coe/indicator_cnb.asp

National Center for Education Statistics. (2015). *The condition of education.* Retrieved from http://nces.ed.gov/pubs2015/2015144.pdf

National Center on Intensive Intervention at American Institutes for Research (n.d). *Curriculum-based measurement in reading(CBM-R): maze fluency.* Retrieved from http://www.intensiveintervention.org/progress-monitoring/curriculum-based-measurement-reading-cbm-r-maze-fluency

National Center on Response to Intervention. (2010). *Essential components of RTI- a closer look at response to intervention.* Retrieved from http://www.RTI4success.org/sites/ default/files/RTIessentialcomponents_042710.pdf

National Center on Response to Intervention. (2012). *RTI implementer webinar series: What is a mulit-level prevention system.* Retrieved from http://www.RTI4success.org

National Governors Association, the Council of Chief State School Officers, and Achieve, Inc. (2008). *Benchmarking for success: Ensuring u.s. students receive a world-class education.* Retrieved from http://www.corestandards.org/assets/0812BENCHMARKING.pdf

National Institute for Literacy. (2003). Put reading first: The research building blocks for teaching children to read (2nd ed.). Washington, DC: U.S. Department of Education.

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. Reports of the subgroups.* Washington, DC: U.S. Government Printing Office.

National Joint Committee on Learning Disabilities. (2002). *Specific learning Disabilities: finding common ground.* Retrieved from http://www.ldonline.org/article/5720/

National Joint Committee on Learning Disabilities. (2005). Responsiveness to intervention and learning disabilities. Report retrieved from http://www.nasponline.org/advocacy/RTI%20Final%20August%202005.pdf

National Research Center on Learning Disabilities. (2006). *RTI manual.* Retrieved from http://files.eric.ed.gov/fulltext/ED496979.pdf

National Reading Panel. (2000). *Teaching children to read: An evidence based assessment of the scientific research literature on reading and its implication for reading instruction.* Washington, DC: U.S. National Institute for Literacy.

NCS Pearson, Inc., NCS. (2012a). AIMSweb technical manual. Retrieved from http://www.aimsweb.com/wp-content/uploads/aimsweb-Technical-Manual.pdf

NCS Pearson, Inc., NCS. (2012b). Introductory guide. Retrieved from http://www.aimsweb.com/wp-content/uploads/aimsweb_IntroGuide.pdf

No Child Left Behind Act of 2001, Pub. L. 107-110, 115 Stat. 1425 (2001).

O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3* (2), 159–197.

O'Connor, R. E., Fulmer, D., Harty, K., & Bell, K. (2005). Layers of reading intervention in kindergarten through third grade: Changes in teaching and child outcomes. *Journal of Learning Disabilities*, *38,* 440-455.

O'Connor, R. E., Bocian, K. M., Beach, K. D., Sanchez, V., & Flynn, L. J. (2013). Special Education in a 4-Year Response to Intervention (RTI) Environment: Characteristics of Students with Learning Disability and Grade of Identification.

Learning Disabilities Research & Practice (Wiley-Blackwell), 28(3), 98-112. doi:10.1111/ldrp.12013

Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). The role of interim assessments in a comprehensive assessment system: A policy brief. Retrieved from National Center for the Improvement of Educational Assessment, http://www.nciea.org/publication_PDFs/PolicyBriefFINAL.pdf

P.L. 94-142 (1975). *Education for all handicapped children act of 1975*. Washington, DC: U.S. Government Printing Office

Popham, W. J. (2008). *Transformative Assessment.* Danvers, MA: ASCD.

President's Commission on Excellence in Special Education. (2002). *A new era: Revitalizing special education for children and their families.* (U.S. Department of Education Contract No. ED-02-PO-0791). Washington DC: U.S. Department of Education. Retrieved from http://files.eric.ed.gov/fulltext/ED473830.pdf

Reschly, D. J. (1981). Psychological testing in educational classification and placement. *American Psychologist*, *36*(10), 1094-1102. doi:10.1037/0003-066X.36.10.1094

Reschly, D. J. (2003, December). What if LD identification changed to reflect research findings? Consequences of LD identification changes. Paper presented at the Responsiveness-to-Intervention Symposium, Kansas City, MO

Reschly, D. J. (2005). Learning disabilities identification primary intervention, secondary intervention, then what? *Journal of Learning Disabilities, 38(6),* 510-515.

doi: 10.1177/00222194050380060601

Reschly, R. L., Busch, T. W., Betts, J., Deno, S., & 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, doi: 10.1002/pits

Ruhe,V., & Moore, P. (2005). The impact of Reading Recovery on later achievement in reading and writing. *ERS Spectrum, 23(1),* 20-30.

Sanetti, L & Kratochwill, T. (2009). Toward developing a science of treatment integrity: Introduction to the Special Series. *School Psychology Review, 38,* 445-459.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R., (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96(2),* 265-282. doi: 10.1037/0022-0663.96.2.265

Scholin, S. E., & Burns, M. K. (2012). Relationship between pre-intervention data and post-intervention reading fluency and growth: A meta-analysis of assessment data for individual students. *Psychology In The Schools*, *49*(4), 385-398. doi: 10.1002/pits.21599.

Schmitt, M. C., & Gregory,A. E. (2005). The impact of an early literacy intervention: Where are the children now? *Literacy Teaching and Learning: An International Journal of Early Reading and Writing, 7(1-2),* 57-76.

Shapiro, E. S. (2009). *Tiered instruction and intervention in a response-to-instruction model.* Retrieved from http://www.rtinetwork.org/essentials/tieredinstruction /tiered-instruction-and-intervention-rti-model

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

Shinn, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *Journal Of Special Education,34*(3), 164-172. doi:10.1177/002246690003400305

Shinn, M. M., & Shinn, M. R. (2002a). *AIMSweb training workbook: Administration and scoring of reading curriculum-based measurement (R-CBM) for use in general outcome measurement.* Eden Prairie, MN: Edformation, Inc. Retrieved from https://www.researchgate.net/publication/267793797_AIMSweb_Training_Workbook_ Administration_and_Scoring_of_Reading_Curriculum-Based_ Measurement_R-CBM_for_Use_in_General_Outcome_Measurement

Shinn, M. M., & Shinn, M. R. (2002b). *AIMSweb training workbook: Administration and scoring of reading MAZE CBM for use and general outcome measurement.* Eden Prairie, MN: Edformation, Inc. Retrieved from http://www.cnyric.org/tfiles/folder 1052/Administration%20and %20Scoring%20of%20Reading%20MAZE CBM.pdf

Shinn, M. R., (2007). Identifying students at risk, monitoring performance, and determining eligibility within response to intervention: Research on educational need and benefit from academic intervention. *School Psychology Review, 36(4),* 601-617.

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

Speece, D. L., Case, L. P., Molloy D.E. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research and Practice, 18(3),* 147-156, doi/10.1111/154

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*(3), 407-419.

Stanovich, K. E. (1999). The sociopsychometrics of learning disabilities. *Journal of Learning Disabilities, 32(4),* 350-361.

Stanovich, K. E., & Siegel L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86(1),* 24-53.

Stecker, P. M., & Lembke, E. S. (2005). *Advanced applications of CBM in reading: Instructional decision-making strategies manual.* Washington, DC: National Center on Student Progress Monitoring.

Stiggins, R. (2007). Assessment through the student's eyes. *Educational Leadership, 64(8),* 22-26.

Telzrow, C. F., McNamara, K., & Hollinger, C. L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review, 29(3),* 443.

TILSA SCASS Interim Assessment Subcommittee (2008). Interim assessment practices and avenues for state involvement. Washington, DC: Council for Chief State School Officers. Retrieved from http://www.ccsso.org/ Documents/2008/ Interim_Assessment_Practices_and_2008.pdf

Tilly, W. D. (2002). Best practices in school psychology as a problem-solving enterprise. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology IV* (p. 21-36). Bethesda, MD: National Association of School Psychologists.

Tomlinson, C. (2003). Fulfilling the promise of the differentiated classroom: Strategies and tools for responsive teaching. Alexandria, VA: Association for Supervision and Curriculum Development.

Torgesen, J. K. (1998). Catch them before they fall. *American Educator, American Federation of Teachers, Spring/Summer.*

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resistors. *Learning Disabilities Research & Practice, 15*, 55-64.

Torgesen, J. K. (2002). The prevention of reading difficulties. *School Psychology Review, 40,* 7-26.

Torgesen, J. K. (2007). Using an RTI model to guide early reading instruction: Effects on identification rates for students with learning disabilities. *Flordia Center for Reading Research FCRR Technical Report # 7.* Retrieved from http://www.fcrr.org/science/sciencePresentationsTorgesen.shtm

Torgesen, J. K., Morgan, S., & Davis, G. (1992). The effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology, 84,* 364-370. doi: 10.1037/0022-0663.84.3.364

Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. Metsala &L. Ehri (Eds.), *Word recognition in beginning reading* (pp. 161-188). Hillsdale, NJ: Erlbaum.

Torgesen, J., Wagner, R., & Rashotte, C. (1999). *Test of word reading efficiency*. Austin, TX: PRO-ED.

Torgesen, J. K.,Alexander, A.W.,Wagner, R. K., Rashotte, C.A.,Voeller, K.K.S., & Conway, T. (2001). Intensive remedial instruction for children with severe

reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34(1),* 33-58.

Truscott, S. D., Cohen, C. E., & Sams, D. P., Sanborn, K. J., & Frank, A. J. (2005). The current state(s) of prereferral intervention teams: A report from two national surveys. *Remedial & Special Education, 26(3),* 130-140. doi:10.1177/07419325050260030101

University of Oregon. (2016). Center on teaching and learning. Retrieved from retrieved from https://ctl.uoregon.edu/about/ history

U.S. Department of Education. (n.d.) *No child left behind overview.* Retrieved from http://febp.newamerica.net/background-analysis/no-child-left-behind-overview

U.S. Department of Education. (2001). *Ideas That Work. IDEA '97 final regulations: Major issues.* Retrieved from http://www2.ed.gov/offices/OSERS/Policy/IDEA/regs.html

U.S. Department of Education. (2006). *Assistance to states for the education of children with disabilities and preschool grants for children with disabilities; Final rule.* Part II, 34 CFR Parts 300 and 301. Retrieved from http://idea.ed.gov/download/finalregulations.pdf

U.S. Department of Education. (2007). *Archived history: Twenty-five years of progress in educating children with disabilities through IDEA.* Washington, DC: U.S. Office of Special Education Programs.

U.S. Department of Education. (2008). *Reading First.* Retrieved from http://www2.ed.gov/programs/readingfirst/index.html

U.S. Department of Education. (2010a). *Laws and Guidance*. Retrieved from http://www2.ed.gov/policy/landing.jhtml?src=ln

U.S. Department of Education. (2010b). *A blueprint for reform: The reauthorization of the elementary and secondary education act.* Retrieved from http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf

U.S. Department of Education (2012). *State of Missouri ESEA Flexibility Request*. Retrieved from https://www2.ed.gov/policy/eseaflex/approved-requests/mo.pdf

U.S. Department of Education (2015). *Every child succeeds act (ESSA)*Retrieved from http://www.ed.gov/essa?src=rn

U.S. Department of Special Education. (2007a). *Building the legacy: IDEA 2004*. Retrieved from http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C3%2C

U.S. Department of Special Education. (2007b). *Building the legacy: IDEA 2004*. Retrieved from http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C23%2C

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

Vaughn, S., Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research and Practice, 18(3),* Retrieved from http://www.disfor.unict.it/Public/Uploads/links/redefining%20learning%20dis.pd

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

Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in the Schools, 47(5),* 432–444.

Vaughn, S., & Fuchs L. S. (2012). Responsiveness-to-intervention: A decade later. *Journal of Learning Disabilities, 45(3).* 195-203. doi:10.1177/0022219412442150

Vellutino, F. R., Scanlon, D. M., Sipay, E .R., Small, S., Ghen, R., Pratt, A., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88(4),* 601-638.

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39(2),* 157-169.

Wanzek, J., & Vaughn, S. (2010). Tier 3 interventions for students with significant reading problems. *Theory into Practice. 49*, 309-314. doi:10.1080/00405841.2010.510759

Wanzek, J., & Vaughn, S. (2011). Is a three-tier reading intervention model associated with reduced placement in special education? *Remedial and Special Education, 32*, 167–175. doi:10.1177/0741932510361267

Watts-Taffe Watts-Taffe, S., Laster, B. P., Broach, L., Marinak, B., Connor, C. M., & Walker- Dalhouse, D. (2012). Differentiated instruction: Making informed decisions. *Reading Teacher, 66(4),* 303-314.

What Works Clearinghouse (2008). *What works clearinghouse evidence standards for reviewing studies.* Washington, DC: United States Department of Education. Retrieved from https://ies.ed.gov/ncee/wwc/pdf/reference_resources/ wwc_version1_standards.pdf

Wheadon, C. (2014). Classification accuracy and consistency under item response theory models using the package classify. *Journal of Statistical Software, 56(10).* 1-14.

Wiley, H.I. & Deno, S.L. (2005). Oral reading and maze measures as predictors of success for English language learners on a state standards assessment. *Remedial and Special Education 26*(4), 207-214.

Wright, Jim. (2007). *Response to intervention toolkit: A practical guide for schools.* Port Chester, NY: Dude.

Appendices

Appendix A: IRB Form

BAKER UNIVERSITY

Date: 7/30/2016

SCHOOL OF EDUCATION

IRB PROTOCOL NUMBER

GRADUATE DEPARTMENT

(IRB USE ONLY)

IRB REQUEST

Proposal for Research

Submitted to the Baker University Institutional Review Board

I. Research Investigator(s) (Students must list faculty sponsor first)

Department(s) **School of Education Graduate Department**

	Name	Signature	
1.	Dr. Harold Frye	*Harold B. Frye*	Major Advisor
2.	Dr. Phil Messner	***Phillip Messner***	Research Analyst
3.	Dr. Charlsie Prosser		University Committee Member
4.	Dr. Ryan Rostine		External Committee Member

Principal Investigator: Sandra K. Rice
Phone: 816-228-0817
Email: srice1021@sbcglobal.net
Mailing address: 1302 NW 355th Road, Holden, Missouri 64040

Faculty sponsor: Harold B. Frye, Ed.D.
Phone: 913-344-1220
Email: hfrye@bakeru.edu

Expected Category of Review: X Exempt ____Expedited ____Full

II: Protocol Title
Curriculum-Based Fluency and Comprehension Measurements as Predictors of Elementary Student Performance on State-Wide MAP Communication Arts Assessment Scores in an RTI Setting

Summary
In a sentence or two, please describe the background and purpose of the research.

Beginning in the 2013-2014 school year, all students in grades K-6 of District XYZ are given benchmark reading curriculum-based measurements (R-CBM) and MAZE reading comprehension measurements three times each year; fall, winter, and spring. From these criterion assessments, students who score below 25% are provided classroom remediation using the district RTI model (LSSD, 2016). It is important to know if there is a relationship between the formative and summative assessments to justify the usage of the RCBM and MAZE as a source of predictive assessment data for fourth grade students. The intended result of this study is to determine if R-CBM and Maze formative assessments are predictive of the summative ELA MAP assessment in fourth grade.

Briefly describe each condition or manipulation to be included within the study.

What measures or observations will be taken in the study? If any questionnaire or other instruments are used, provide a brief description and attach a copy.

This study will use AIMSweb and MAP archival data for fourth grade from 2013-2014 school year. No other questionnaires or instruments will be used.

Will the subjects encounter the risk of psychological, social, physical, or legal risk? If so, please describe the nature of the risk and any measures designed to mitigate that risk.

No, there will be no risk of psychological, social, physical, or legal risk.

Will any stress to subjects be involved? If so, please describe.

No there will not be any stress to subjects involved.

Will the subjects be deceived or misled in any way? If so, include an outline or script of the debriefing.

No, the subjects will not be deceived or misled in any way.

Will there be a request for information that subjects might consider to be personal or sensitive? If so, please include a description.

No, there will be no request for information that subjects might consider to be personal or sensitive.

Will the subjects be presented with materials that might be considered to be offensive, threatening, or degrading? If so, please describe.

No, the subjects will not be presented with materials that might be considered to be offensive, threatening, or degrading.

Approximately how much time will be demanded of each subject?

There will be no time demanded of each subject.

Who will be the subjects in this study? How will they be solicited or contacted? Provide an outline or script of the information which will be provided to subjects prior to their volunteering to participate. Include a copy of any written solicitation as well as an outline of any oral solicitation.

The subjects in this study are fourth graders from 2013-2014 school year. This study uses archival data and subjects will not be contacted in any manner.

What steps will be taken to ensure that each subject's participation is voluntary? What if any inducements will be offered to the subjects for their participation?

This study uses AIMSweb and MAP archival fourth grade data from the 2013-2014 school year. Student data information will be used without evidence to names or student numbers.

How will you ensure that the subjects give their consent prior to participating? Will a written consent form be used? If so, include the form. If not, explain why not.

No written consent will be needed for participant involvement in this study.

Will any aspect of the data be made a part of any permanent record that can be identified with the subject? If so, please explain the necessity.

No aspect of the data will be made a part of any permanent record that can be identified with the subject.

Will the fact that a subject did or did not participate in a specific experiment or study be made part of any permanent record available to a supervisor, teacher or employer? If so, explain.

No, there is not a need for subject participation in this study.

What steps will be taken to ensure the confidentiality of the data?

The AIMSweb and MAP archival data will be used from fourth grade students in 2013-2014 school year without evidence of student names or numbers. This study will look at the numerical relationship of MAP quartile levels and Tiered levels of students.

If there are any risks involved in the study, are there any offsetting benefits that might accrue to either the subjects or society?

There are no risks involved to students in this study. The intended result of this study is to determine if R-CBM and Maze formative assessments are predictive of the summative ELA MAP assessment in fourth grade.

Will any data from files or archival data be used? If so, please describe.

Yes, this study will use archival data from fourth grade 2013-2014 AIMSweb Benchmark Assessments provided to students three times a year, and fourth grade archival MAP data for the same year.

Appendix B: Baker University IRB Approval

Baker University Institutional Review Board

September 1, 2016

Dear Sandra Rice and Dr. Frye,

The Baker University IRB has reviewed your research project application and approved this project under Exempt Status Review. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

Please be aware of the following:

1. Any significant change in the research protocol as described should be reviewed by this Committee prior to altering the project.
2. Notify the IRB about any new investigators not named in original application.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents of the research activity.
4. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.
5. If the results of the research are used to prepare papers for publication or oral presentation at professional conferences, manuscripts or abstracts are requested for IRB as part of the project record.

Please inform this Committee or myself when this project is terminated or completed. As noted above, you must also provide IRB with an annual status report and receive approval for maintaining your status. If you have any questions, please contact me at emorris@BakerU.edu or 785.594.7881.

Sincerely,

Erin R, Morris PhD
Chair, Baker University IRB

Baker University IRB Committee
Susan Rogers PhD
Nate Poell MA
Joe Watson PhD
Scott Crenshaw

Appendix C: District Approval Letter

Lee's Summit R-7 School District

301 NE Tudor Rd
Lee's Summit, MO 64086-5702
(816) 986-1027 Fax (816) 986-1060

Office of Curriculum and Instruction

August 10, 2016

Sandy Rice
Longview Farm Elementary

Dear Sandy,

After reviewing the changes/updates to your doctorate research study you have my approval to continue with your study.

Please remember you must maintain the confidentiality of all student information at all times.

Sincerely,

Dr. Kevin Daniel
Associate Superintendent Instruction & Leadership

xc: Dr. Katie Collier, Assistant Superintendent of Elementary Instruction
Dr. Ryan Rostine, Principal, Longview Farm Elementary

ar

Appendix D: Data Collection Request Form – District XYZ

Document Provided to Intended Researcher by:

Signature Dr. Ryan A Forster Date: 8/12/16
(Building Principal for District Employee
or SLC Staff for Out-of-district Researchers)

INSTRUCTIONAL OPERATIONS TEAM
Lee's Summit R-7 School District
301 NE Tudor Rd.
Lee's Summit, Missouri 64086

REQUEST FOR PERMISSION TO CONDUCT RESEARCH/GATHER DATA
IN THE LEE'S SUMMIT R-7 SCHOOLS
TO MEET A COURSE REQUIREMENT

DIRECTIONS: The applicant should complete this form, obtain the necessary approval and signatures, and return to:

Associate Superintendent of Instruction & School Leadership
Lee's Summit R-7 School District
301 NE Tudor Rd.
Lee's Summit, Missouri 64086

It may take up to three weeks for requests to be processed; please plan accordingly in order to meet course deadlines.

1. Please describe concisely the basic concepts and goals of your proposed project, and include an explanation of how the project meets a course requirement within the field of education.

Beginning in the 2013-2014 school year, all students in grades K-6 of District XYZ are given benchmark reading curriculum-based measurements (R-CBM and MAZE) three times each year; fall, winter, and spring. From these criterion assessments, students who score below 25% are provided classroom remediation using the district RTI model. It is important to know if there is a relationship between the formative and summative assessments to justify the usage of the RCBM and MAZE as a source of predictive assessment data for fourth grade students. The intended result of this study is to determine if R-CBM and Maze formative assessments are predictive of the summative ELA (Communication Arts) MAP assessment for quartile levels (below basic, basic, proficient and advanced) in fourth grade. This study will help to inform educators and the public of the impact that formative assessments have on student achievement when aligned with academic goals and standards. Also, this study may provide evidence the district needs to meet district and state annual progress as outlined in Missouri's Top 10 by 20 and MSIP 5 goals.

Approved 2008

2. List the names of all data collection instruments you intend to use and enclose a copy of each with this application. Also, enclose a copy of each parent/student consent form. Please describe in detail the distribution, implementation, and collection methods you intend to use in your data collection.

This study will use AIMSweb and MAP archival data for fourth grade from 2013-2014 school year. No other questionnaires or instruments will be used.

3. Give the names of the Lee's Summit R-7 School District public school(s), you intend to involve to meet the project requirements. Are there certain demographics required for the project (ie: grade level, gender, etc.)

For this quantitative study, I would like to use fourth grade benchmark data (AIMSweb RCBM and MAZE CBM) and the state MAP assessment data for student specific quartiles levels (below basic, basic, proficient, advanced) for the 2013-2014 school year. I will not be using student names or information just the district numbers of fourth grade students across all of LSSD elementary schools. I will not be studying specific interventions or focusing on any socio-economic groups of students.

4. What amount of time would be required of staff or students in the R-7 schools in order to meet project requirements?

I will not need additional time of the teachers or district administrators for this study.

5. Are there any other school records you would require (for example, achievement test scores or attendance?). If so, please provide a detailed explanation of your process to code such records to ensure confidentiality.

I would like to use the fourth grade archived data of AimsWeb benchmark (R-CBM and MAZE CBM), and individual student MAP data for the 2013-2014 school year.

6. Give the name of each person who will enter the schools. For nondistrict employees, please provide existing background checks for individuals or a plan to ensure background checks are in place prior to entry in schools.

There will be no individuals entering schools including myself.

7. What is the date you wish to begin? June, 2016 .

8. By what date do you anticipate being finished? December, 2016 .

9. Please obtain the signature of your instructor responsible for this assignment and attach a copy of the assignment guidelines.

Signature: Harold B. Frye EdD

Harold B. Frye
hfrye@bakeru.edu
913-344-1220

Approved 2008

Position: DISSERTATION ADVISOR AND ASSOCIATE PROFESSOR

University/College/School/Department/Division:
GRADUATE SCHOOL – SCHOOL OF EDUCATION
BAKER UNIVERSITY

10. Name of applicant (please print)

Sandra Rice

Sandra K Rice

Signature

1302 NW 355th Road
Holden, Missouri, 64040
Address

5th Grade Teacher at Longview Farm Elementary
Position/Status

6/1/2016
Date

816-830-0812
Phone Number

CRITERIA FOR APPROVAL OR DISAPPROVAL

The approval or disapproval of requests will be made within the following general guidelines.

1. The only projects which will generally be approved are those which:
   a) contribute to the improvement of education in the Lee's Summit R-7 Schools;
   b) contribute to the improvement of education in general.

2. Even within the above categories, studies will generally be disapproved if they:
   a) appear to infringe on the privacy of pupils, parents, or staff members;
   b) present a burden to pupils or staff members;
   c) threaten school-community relations in any way.

3. Research solely for a course requirement will be considered only for Lee's Summit R-7 School District staff.

Approved 2008

4. At any point in the research process, R-7 staff can terminate the study if determined necessary for any reason.

5. The R-7 School District reserves the right to access any results or product created as a result of projects conducted using R-7 students, staff, or facilities.

PARTICIPATION OF THE SCHOOLS

Generally, participation in any research study conducted by an outside agency or individual will be completely voluntary on the part of the principals, teachers, pupils and any other personnel involved.

Approved 2008