The Relationship between Kindergarten and First-Grade DIBELS Next and TRC Scores and Third-Grade TCAP ELA Scores at a Selected School District in Upper East Tennessee

By

JoDee Dotson

A dissertation proposal submitted in partial fulfillment of the requirements for the degree of

Doctor of Education

Education Department

School of Social Sciences and Education

Milligan College, Tennessee

2019

Doctoral Committee:

Dr. Angela Hilton-Prillhart

Dr. Mark Dula

Dr. Patrick Kariuki

Copyright Page

Abstract

The purpose of the research study was to examine the relationship and predictive validity of kindergarten and first-grade students' scores on Dynamic Indicators of Basic Early Literacy Skills Next (DIBELS Next) and Text Reading and Comprehension (TRC) to third-grade students' Tennessee Comprehensive Assessment Program (TCAP) English Language Arts (ELA) scores. Participants included 343 students from eight elementary schools in an Upper East Tennessee school district who were administered the DIBELS Next and TRC in kindergarten in 2014-2015 and first-grade in 2015-2016 and third-grade TCAP ELA in 2017-2018. Eight research questions and eight null hypotheses were analyzed using Pearson correlation coefficient and multiple linear regression. When the variables were measured independently of each other, results showed a significant relationship between kindergarten DIBELS Next composite scores (r=.494, p=.001), kindergarten TRC level of proficiency (r=.468, p=.001), first-grade DIBELS composite scores (*Beta*=.617, p=.001), first-grade DIBELS Next Oral Reading Fluency fluency scores (*Beta*=.5551, p=.001), and first-grade TRC level of proficiency (r=.580, p=.001) to third-grade TCAP ELA. However, there was not a significant relationship between first-grade DIBELS Next Oral Reading Fluency accuracy scores (*Beta*=-.041, p=.544) and first-grade DIBELS Next Oral Reading Fluency retell scores (*Beta*=.070, p=.187) on third-grade TCAP ELA. When variables were measured together, findings indicated that both first-grade DIBELS Next composite scores (*Beta*=.266, p=.001) and first-grade TRC level of proficiency (*Beta*=.268, p=.001) are equally strong predictors of third-grade TCAP ELA followed by kindergarten DIBELS Next composite scores (*Beta*=1.67, p=.004). Kindergarten TRC level of proficiency (*Beta*=.037, p=.547) was not a significant predictor of future performance on third-grade TCAP ELA.

Keywords: Curriculum-Based Measurement (CBM), Dynamic Indicators of Basic Early Literacy Skills (DIBELS Next), DIBELS Next Composite Score, DIBELS Next Oral Reading Fluency, High-Stakes Testing, Response to Intervention (RtI), Text Reading and Comprehension (TRC), Tennessee Comprehensive Assessment Program (TCAP), Universal Screening.

Milligan College Dissertation Defense Approval Form

Candidate Name: JoDee Dotson

Date of Defense: April 23, 2019

Dissertation Title: The Relationship between Kindergarten and First-Grade DIBELS Next and TRC Scores and Third-Grade TCAP ELA Scores at a Selected School District in Upper East Tennessee

Final Approval of Dissertation Defense

A signature below indicates committee members agree with the following:

1) Agreement the dissertation meets with the committee's approval.
2) Agreement the oral defense of the dissertation was successful.

Angela Hilton-Pruitt — Dissertation Chair Signature — Date: 04/23/19

Patricia M. Kariuki — Dissertation Committee Member Signature — Date: 04/23/19

Dissertation Committee Member Signature — Date: 4/23/19

Other Required Signatures

Patricia M Kariuki — Ed.D. Research Director Signature — Date: 4/23/19

Angela Hilton-Pruitt — Area Chair of Education Signature — Date: 4/23/19

Dedication

And whatever you do, whether in word or deed, do it all in the name of the Lord Jesus, giving thanks to God the Father through him.

Colossians 3:17

This research study is dedicated…

To my husband, Kit. I am so blessed by you. You have been my constant encourager, supporter, and personal Excel expert. You now know more about early reading screeners and state assessments than you ever wanted to know! Thank you for loving me through this journey. Now it's time for empty nest! I love you always.

To my daughter, Taylor. Thank you for letting me graduate from Milligan College with you! When I began this journey, I never dreamed that would happen. I am so excited for what your future holds as you soon begin your new roles of wife and graduate student. Continue to work hard, and you will succeed. You are my sunshine, and I love you with all my heart.

To my daughter, Abby. Thank you for not minding that Mom was actually on campus as you began your college career at Milligan. Thanks for laughing when students thought they saw someone in the library that looked like you! I am looking forward to watching you grow spiritually and academically and seeing what God has planned for your future. Continue to study hard and chase your dreams. You are my baby girl, and I love you with all my heart.

To my mother, Dee Ann Whitaker. You were the first teacher I had and the reason I became a teacher. The example you set for me and the passion you had for helping children was truly "in my blood." Thank you for supporting me in so many ways on this journey. I could not have completed it without you. I love you.

To my father, Joe Whitaker. You are my constant supporter and one of my biggest cheerleaders. I am proud to earn this degree from a place that means so much to our family. Thank you for all you have given to this special place and all you have done for me. I love you.

To my in-laws, Denver and Cissie Dotson. I am so grateful to be a part of your family. Thank you for your phone calls and texts inquiring about my classes and dissertation. You have prayed for me and encouraged me so much throughout this process. I love you both.

In memory of my grandparents, Neal and Dorothy Whitaker, and my grandmother, Aleene Cockerham. You have been such a big part of this journey, always in my thoughts and always in my heart. I selfishly wish you were here to share this special time with me. I know you would be so proud. I love each of you and always will.

I am truly blessed.

Acknowledgments

I would like to extend my deepest gratitude to all of my family and friends who have supported and encouraged me throughout this doctoral journey.

I would like to thank my dissertation chair, Dr. Angela Hilton-Prillhart. Thank you for sharing your educational expertise and wisdom with me. I am still in awe of how our topics were so similar. It was such a pleasure to cite your research in mine. Thank you for being my encourager and friend.

I would like to extend my sincere gratitude to my other committee members, Dr. Patrick Kariuki and Dr. Mark Dula. Dr. Kariuki, you never cease to amaze me with your vast knowledge. You made educational statistics interesting and fun for this literacy girl! Thank you for your guidance and encouragement during each and every meeting. I did it! Dr. Dula, welcome to Milligan! The college is blessed to have you as a part of the education faculty. Your class was the first that sparked my interest for quantitative research. Thank you for supporting me through the statistical analysis process and for encouraging me each step of the way.

I would also like to say a special thank you to Dr. Lyn Howell, whose vision began the doctoral program in educational leadership at Milligan College. Thank you for making this possible, Dr. Howell. I am so proud to be part of the inaugural doctoral class!

I would like to thank all Johnson City Schools' administrators, teachers, staff, and students. I am blessed to be a part of such a wonderful school district who puts students first and expects the best. A very special thank you to Dr. Robbie Anderson. You believed in me before I even moved to Johnson City. Thank you for seeing in me what I did not see in myself. I am blessed that you were my professor in this program and continue to be my supervisor, colleague,

and, more importantly, my friend. I learn so much from you daily. Thank you also to Dr. Roger Walk. Your ideas and advice inspired my research topic. I hope my results will be beneficial to the school district. Thank you for your constant support and encouragement. Thank you also to Dr. David Timbs. You are inspirational and a true servant leader. How blessed I have been to know you and learn from you as my mentor and professor. Thank you for encouraging me to begin this journey, to finish this journey, and to follow my dreams.

I would also like to extend my gratitude to my dear friend, Dr. Karin Keith. The words from your dissertation acknowledgements…"your turn!"...inspired me to begin. Thank you for teaching me so much about literacy and about life. You continue to inspire me and support me even though I do not see you on a daily basis. Let's have lunch soon!

Thank you to all of the Johnson City Schools' curriculum coaches whom I work with every day—Tiffany Hogan, Dr. LaDonna Boone, Dr. Lori Church, Dr. Allison Gardenhour, Shannon Suttle, Dr. Sherry Cockerham, Tiffany Hibbitts, Dr. Carleton Lyon, and Dr. Chris Bowen. You have all listened to me and have been a constant support throughout this process. I am forever grateful to each of you. Thank you for all you do because of your love for our students. Tiffany Hogan…"your turn!"

Lastly, I cannot end without thanking Tammy Markland and Jan Zuehlke, my doctoral partners. The P.H. Welshimer Library study rooms will never be the same when we finish this degree! From all the snacks, the laughs, the strange looks from undergraduates (who are those old women and what program are they in?!), the texts, the encouragements, the goal settings, and the dreams, you all carried me through this journey. I am blessed to have gotten to know you and even more grateful to finish with you. Congratulations, Dr. Tammy Markland and Dr. Jan Zuehlke! We did it!

Table of Contents

List of Tables

CHAPTER 1

Introduction

The goal of improving student achievement, especially in reading, unites us all as educators, parents, students, community leaders, and citizens. Using assessments to determine student progress, to measure student achievement, and to guide instruction is a frequent, well-established practice in education. Some assessments are formal and given for summative purposes. Others are informal and given for formative reasons. Ideally, the goal of assessment is to identify the strengths and weaknesses of students through various measures of observing, testing, and interviewing between teacher and student in order to implement effective intervention (Deno & Mirkin, 1977; DeVries, 2011; Fuchs & Fuchs, 2006). Formal state assessments and screeners in math and reading have become more common in schools in response to federal, state, and district accountability demands, specifically involving Response to Intervention (RtI) and identification of students with learning disabilities (Fuchs & Fuchs, 2006; Meisels, 1989). The pressure for students to be on grade level in all subject areas each year has resulted in an "academic trickle-down" as many early literacy programs now also include assessment of students using universal screeners (Meisels, 1989, p. 16). These screeners are often used to evaluate and plan for future instruction and intervention (Invernizzi, Justice, Landrum, & Booker, 2004; Meisels, 1989; Rathvon, 2004; Snow, Burns, & Griffin, 1998). The need to remediate students using universal screeners has gained momentum since these assessments play a decisive role in identifying students who need intervention (D. Fuchs, L. S. Fuchs, & Compton, 2012; Munger, LoFaro, Kawryga, Sovocool, & Medina, 2014). This push has been accelerated by the reauthorizations of the Elementary and Secondary Act of 1964 called No Child Left Behind (NCLB) of 2001 and the Every Student Succeeds Act (ESSA) of 2015 (United States Department of Education, 2001; United States Department of Education, 2015).

The NCLB law of 2001 intensified accountability measures for schools and districts in regards to student achievement, outcomes, and focused attention on early literacy (Invernizzi, et al., 2004). As an update of the 1965 Elementary and Secondary Education Act (ESEA), the NCLB law also put an emphasis on ensuring accountability for all federally-funded schools for student achievement and advancing the academic performance of several subgroups, including English Language Learners, special education students, and socio-economic groups. The law identifies tests as the only measure of outcome. Even though many states already had yearly state assessments prior to this legislation, only half of the states required math and reading for grades 3-8. NCLB mandated that all schools have them in position by 2005-2006 (Huddleston & Rockwell, 2015). In addition, the law demanded that 100% of all students read on grade level by 2014 and mandated intervention if goals were not met (Klein, 2015; Martin, 2016). In 2015, ESSA replaced NCLB. Although it continues to require statewide testing, reporting of subgroups, and school accountability systems, it gives states more authority to decide what is measured and how assessments are utilized in the accountability of schools and districts, which provides more flexibility (Martin, 2016).

In 2007, Tennessee had received a letter grade of "F" on truth in advertising by the United States Chamber of Commerce for low academic expectations (Tennessee Department of Education, n.d.c). There was a 60-point difference in the ratings of proficiency for reading and math between the Tennessee Comprehensive Assessment Program (TCAP) and the National Assessment of Educational Progress (NAEP) (Mansouri, 2018). This resulted in a call to action by the state to ultimately institute more rigorous standards and a new assessment called TNReady (Tennessee Department of Education, n.d.c).

Additionally, as a result of the NCLB education legislation, the Reading First initiative was established to encourage educators to use research-based strategies in grades K-3 instruction and funded professional development in schools with low-performing students (Riedel & Samuels, 2007). Reading First required states to use appropriate tools for assessments that served the purposes of "screening, diagnosis, progress monitoring, and outcome assessment" (Invernizzi, et al., 2004, p. 480). Reading First called for early literacy assessments to include evidence within the five areas of reading, phonemic awareness, phonics, fluency, comprehension, vocabulary, in addition to assessing reading strategies. This recommendation was based on the findings of the National Reading Panel (NRP) report of 1997 that named these five areas as essential for reading success. The panel suggested teaching these areas explicitly in a balanced approach with frequent progress monitoring in order to capture the picture of each reader (National Reading Panel & National Institute of Child Health and Human Development, 2000; Stahl, 2004). Therefore, in 2002, many assessments, including Dynamic Indicators of Basic Early Literacy Skills (DIBELS), gained popularity to assess the five areas of reading (Riedel & Samuels, 2007). When used appropriately, the data from these assessments should provide information at various points in the year and guide instructional decisions for teaching and learning (Invernizzi, et al., 2004).

As more districts began emphasizing literacy assessments, Congress made many changes to the Individuals with Disabilities Education Improvement Act (IDEA) in 2004. As a result of these changes, Response to Intervention (RtI) was established. Under IDEA, school districts were encouraged to substitute RtI for the I.Q. achievement discrepancy model for the identification of learning disabilities. The intent was to identify and remedy learning difficulties early, instead of waiting for students to qualify under the discrepancy model before receiving

specialized instruction (Spear-Swerling, 2011). As a result, a motivation for use of progress monitoring assessments greatly increased. Consequently, Tennessee schools began using information from the RtI model as part of the standards used to diagnose specific learning disabilities (Tennessee Department of Education, 2015). This would eventually be mandated in 2014 as a result of IDEA. The state established RtI frameworks that hoped to close achievement gaps of students. Tennessee's framework included expectations that all students would engage in quality, Tier 1 instruction, and school districts would use universal screeners to assess student learning. The screeners were to be nationally norm-referenced, provide percentile scores, and establish a rate of improvement for reading and math (Tennessee Department of Education, 2015). The school district from this study continues to use DIBELS Next and Text Reading and Comprehension (TRC) as universal screeners for kindergarten and first-grade students, as they are on the state-approved list of assessments for Tennessee.

More recent legislation has included a replacement for NCLB called Every Student Succeeds Act (ESSA) of 2015. Under this law, states are now required to control their own status and improvement goals of academic achievement, graduation rate, and sub-groups from year to year (Marion, 2016). Emphasis on assessments and accountability still exist under this new law, as testing requirements are nearly identical to the requirements of NCLB (Marion, 2016; Martin, 2016; United States Department of Education, 2015).

As a result of the requirements of ESSA, the TCAP changed to a new assessment called TNReady in 2015 with various item types and rigorous questions aligned with the more rigorous and in-depth standards. This assessment evaluates students in grades 3-8 in four subject areas: English Language Arts (ELA), math, science, and social studies. TNReady's new proficiency cut score matches to the letter grade of "B." These cuts are a result of descriptors for achievement

levels that are aligned to the ones used by NAEP. The levels are mastered, on track, approaching, and below. Consequently, because the proficiency cuts were raised, the number of students who scored proficient or advanced on TNReady decreased dramatically (Tennessee Department of Education, 2018).

Background of the Problem

Assessment has long had a well-known and contentious role in American education, and that is never more accurate than it is today (Rothman & Marion, 2016). New incentives, such as national and state legislation, charter schools, vouchers, and private schools, have pressured public schools into a very competitive field where accountability and assessment have become commonplace. The legislation from NCLB and ESSA emphasizes early identification, reading intervention, and intensifies accountability for states and school districts. It places enormous pressure on schools and districts to raise assessment scores (Rothman & Marion, 2016). Consequently, it is challenging for schools to find and utilize practices of assessment that are meaningful, useful, and present a true picture of student progress. The majority of states now require some kind of comprehensive reading assessment beginning in third-grade for district evaluation purposes. Since the majority of state assessments only occur at the end of the school year, the data they provide can be too little too late (McGlinchey & Hixson, 2004). Also, many times state assessments do not provide adequate information to ensure the implementation of instructional steps to improve the performance of students (Medina & Riconscente, 2006). Therefore, it is vital to assess the performance of students throughout the school year and utilize the data to tailor instruction. When this occurs, the likelihood of students' meeting proficiency will increase (Shapiro, 2008). If universal screener instruments, such as DIBELS Next and TRC, can predict performance on the state assessments, districts and schools may have reliable

instruments to progress monitor students toward their long-term goals (McGlinchey & Hixson, 2004).

Statement of the Problem

In the majority of today's schools, students in early childhood grades are likely to be given at least one if not multiple assessments used as universal screeners as they begin their educational journey. Consequently, school districts in Tennessee yearly assess students in grades 3-8 with TCAP TNReady on rigorous state standards in four subject areas, including English Language Arts (ELA). Assessing all students helps educators select who is more than likely to be at risk in reading and assists in determining needed intervention to prohibit reading difficulties from becoming future problems in hopes that all children can become proficient readers (Torgesen, 2000).

The literacy assessment, DIBELS Next, is a common universal screener that thousands of schools in the United States utilize and is designed to provide educators with standards for determining student progress in basic early literacy skills. It is appealing to schools and districts because it is simple, easy to use, provides immediate feedback, administered quickly, and correlates with the curriculum supported by Reading First (Goodman, 2006). However, critics of DIBELS argue that its simplicity, easiness, and efficiency is, in fact, its downfall. Goodman (2006) suggests that DIBELS misrepresents the complexity of the reading process and reading development because each DIBELS measure overemphasizes fluency, accuracy, and speed through one-minute assessments. For example, the DIBELS Next Oral Reading Fluency (ORF) measure is a test of how many words a child can accurately read from a passage in one minute. Goodman argues (2006) that this measure assumes how fast and how well a child can read in just one minute is the same as in all minutes of reading for the child. Strauss (2014) states that

children who self-monitor their reading or reread for understanding while reading a DIBELS Next ORF passage are identified as needing intensive intervention because reading progress with DIBELS is measured by speed. Additionally, speed can be taught; so, what might appear to be an improvement in reading development is actually just faster performance on the test. According to Spear-Swerling (2011), DIBELS' timed-tests cannot serve as substitutes for more thorough diagnostic assessments that might be needed for struggling students. Utilizing screening measures, such as DIBELS, that only assess skill acquisition, can cause some students to be identified for intervention who do not need it or fail to identify others that do (Goodman, 2006).

Goodman (2006) further says that DIBELS is primarily based on the individual skills of reading and leads to classroom instruction that does not emphasize making meaning from texts. Pressley, Hilden, and Shankland (2005) state that the majority of the time DIBELS fails to predict reading performance of other assessments because it measures how quickly students read instead of how a reader comprehends text. Scanlon (2011) points out that DIBELS sends the message to teachers and students that speed is more important than comprehension. This emphasis on speed is incompatible with best practices in reading research and instruction (Goodman, 2006).

As a result of increased accountability related to student achievement, the school district from this study began assessing students in kindergarten, first grade, and second grade in 2005 using DIBELS 6th Edition and DIBELS Next in 2009, in addition to utilizing TRC through mCLASS: Reading 3D Amplify Education's electronic platform three times yearly. These assessments also became the district's universal screeners for RtI in kindergarten and first grade. Teachers continue to use data from these assessments to guide their instruction and to provide

specific strategies for students' strengths and weaknesses as readers. In addition, school and district administrators are able to track students' progress as well as the progress of subgroups.

Consequently, teachers and schools from this study make important decisions regarding classroom instruction and the educational future of students because of the data from the universal screeners, such as DIBELS Next and state assessments (Munger et al., 2014). Additionally, assessments are costly and place a yearly financial burden on states and school districts (VanDerHeyden, 2011). It is apparent that school districts need to be certain in the reliability, validity, and predictability of the utilized assessments (Invernizzi, et al., 2004). Therefore, it is essential to determine the relationship between DIBELS Next and TRC scores to TCAP ELA scores in this school district.

Purpose of the Study

The purpose of this correlational research study is to examine the relationship between kindergarten and first-grade students' DIBELS Next and TRC scores and third-grade students' TCAP ELA scores. The second purpose of this study is to determine the predictive validity of kindergarten and first-grade DIBELS Next and TRC scores on end of third-grade TCAP ELA performance.

Research Questions

Eight important research questions arise to address the purpose of this study:

1. Is there a relationship between kindergarten DIBELS Next composite scores and third grade TCAP ELA scores?
2. Is there a relationship between first-grade DIBELS Next composite scores and first-grade DIBELS Oral Reading Fluency accuracy scores on third-grade TCAP ELA scores?

3. Is there a relationship between first-grade DIBELS Oral Reading Fluency fluency scores and first-grade DIBELS Oral Reading Fluency retell scores on third grade TCAP ELA scores?
4. Is there a relationship between kindergarten Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores?
5. Is there a relationship between first-grade Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores?
6. Is kindergarten DIBELS Next composite score or kindergarten Text Reading and Comprehension (TRC) level of proficiency a better predictor of future performance on third-grade TCAP ELA?
7. Is first-grade DIBELS Next composite score or first-grade Text Reading and Comprehension (TRC) level of proficiency a better predictor of future performance on third-grade TCAP ELA?
8. Is kindergarten DIBELS Next composite score, first-grade DIBELS Next composite score, kindergarten Text Reading and Comprehension (TRC) level of proficiency, or first-grade Text Reading and Comprehension (TRC) level of proficiency the best predictor of future performance on third-grade TCAP ELA?

Significance of the Study

Since many school districts use early reading assessments such as DIBELS Next and TRC as a predictive indicator of students' reading abilities and as RtI universal screeners, it is imperative to gather and analyze data surrounding these assessments. Schools and teachers are making significant decisions about students' interventions and classroom placements as a result of these screening instruments (Torgesen, 2000). This study should add to the body of research and discussion of early reading screeners as a predictor of future reading success in school

districts as demonstrated on state assessments. In addition to contributing to existing research in the field, the researcher hopes the results and discussions from this study stimulate more questions and clarify many answers regarding literacy assessments for schools and districts. If assessments are being utilized within school districts as a screening method for students in kindergarten and first-grade, the results should establish a relationship with the state assessment. However, if the universal screeners do not provide a relationship with state assessment results, district and school administrators may need to obtain new universal screeners to help make more accurate instructional decisions for students.

Definition of Terms

The following definition of terms is provided to ensure uniformity and understanding of these terms throughout the study. The researcher developed all definitions not accompanied by a citation.

Curriculum-Based Measurement (CBM): "…is a set of specific measurement methods for assessing student progress over time and for identifying students in need of additional instructional support and/or further diagnostic testing" (McGlinchey & Hixson, 2004, p. 193). It "is a scientifically valid form of student progress monitoring that incorporates standard methods for test development, administration, scoring, and data utilization" (Strecker & Lembke, 2005, p. 1). DIBELS Next used in this study is a CBM.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS Next): "…is a set of brief, individually administered fluency-based measures for monitoring the development of pre-reading and reading skills for children from preschool through third grade" (Rathvon, 2004, p. 206). They were created by Good and Kaminski (1996) and include assessments for letter and sound recognition, phonemic awareness, nonsense words, oral reading fluency, oral reading accuracy, oral retell fluency and word use fluency. It includes benchmark testing and progress

monitoring (Rathvon, 2004). This school district gives DIBELS Next to all kindergarten and first-grade students three times a year as a universal screener. The DIBELS Next test used for this study was the DIBELS Oral Reading Fluency (DORF).

Dynamic Indicators of Basic Early Literacy Skills Composite Score (DIBELS Next): "…is a combination of multiple DIBELS Next scores and provides the best overall estimate of students' early literacy skills and/or reading proficiency. The DIBELS Next measures (First Sound Fluency, Phoneme Segmentation Fluency, Nonsense Word Fluency, DIBELS Oral Reading Fluency) that are used to calculate the DIBELS Next Composite Score vary by grade and time of year" (Dynamic Measurement Group, Inc., 2016, p. 4). This study utilized the DIBELS Next composite score for kindergarten and first grade for the end of year.

Dynamic Indicators of Basic Early Literacy Skills Oral Reading Fluency (DORF): a DIBELS Next measure where the teacher gives a student a leveled text and asks him/her to read it aloud for one minute. The assessor records errors in accuracy. The student then retells what he/she read. This measure is repeated three times with different leveled passages and three individual DORF scores are generated based on students' accuracy, fluency, and retell. It is usually administered from mid-first grade to sixth grade. This measure demonstrates students' "accurate and fluent reading of connected text" (Dynamic Measurement Group, Inc., 2011a, p. 1). The end of the year data for first grade was utilized in this study.

High-stakes testing: "Tests are often considered 'high-stakes' when their results are perceived by students, teachers, administrators, parents, or the general public, as being used to make important decisions that immediately and directly affect them" (Madaus, 1988, p. 87).

Response to Intervention (RtI): a framework built around a tiered intervention model in the areas of reading, math, and writing that helps to provide opportunities for small group

instruction and intervention to students who fall at or below the 25th percentile to help them grow academically. This is also a data-based decision-making process in which the collected data informs and guides instructional decisions for the benefit of each student (Tennessee Department of Education, 2015).

Text Reading and Comprehension (TRC): a running record reading assessment administered with a technology platform that utilizes texts from leveled book sets to determine a student's instructional reading level. The text complexity of the book sets follow the criteria of Fountas and Pinnell's guided reading leveling system (Amplify Education, Inc., 2013a). TRC is administered to all kindergarten and first-grade students three times a year as a universal screener in this school district. The TRC reading level for kindergarten and first grade for the end of year was utilized in this study.

Tennessee Comprehensive Assessment Program (TCAP): a program of state-mandated assessments since 1988 for Tennessee students in grades 3-8 in English Language Arts (ELA), math, science, and social studies. The end-of-year assessments also include End of Course (EOC) tests for high school students and tests designed specifically for students with disabilities and English Language Learners (Tennessee Department of Education, n.d.b). The data from the third-grade TCAP ELA were used in this study.

Universal Screening: a process of administering an assessment of academic skills in reading and math to all students to identify strengths and areas of need. It provides school districts with accurate data to assist in making decisions about interventions, remediation, and enrichment for each child (Tennessee Department of Education, 2015). DIBELS Next and TRC were the assessments used for universal screening in the study.

Limitations of the Study

This study is limited because the results may only be generalized to the population within this one school district.

Organization of the Study

Chapter 1 presents the introduction, statement of the problem, research questions, significance of the study, definition of terms, and limitations of the study. Chapter 2 contains the review of related literature and research that addresses the historical impact of educational assessments, early reading measures, Curriculum-Based Measurement, Response to Intervention, development of oral reading fluency, DIBELS Next, TRC, ELA TCAP, and validity and reliability. The methodology and the procedures used to gather data are presented in Chapter 3. The results of analyses and findings to emerge from the study are contained in Chapter 4. Chapter 5 contains a summary of the study and findings, a discussion, conclusions drawn from the findings, and recommendations for further study.

CHAPTER 2

Review of Literature

In the era of high-stakes testing and accountability in our nation's schools, educators are concerned with finding and utilizing the most effective reading assessments possible that provide accurate information for instruction, intervention, and students' reading success. This literature review focuses on the impact and importance of early reading screeners and state assessments based on previous research studies. The early reading screeners presented in the review are utilized for identifying and serving students in RtI in the school district where this study was conducted. A history of reading and assessments is also included to provide background and historical context to the review. Essential components of the early reading screeners, including phonemic awareness, phonics, fluency, and comprehension, are also provided. Additionally, an explanation of school district measures, such as DIBELS Next and TRC, and state measures, such as TCAP TNReady, are included in the review.

A History of Reading and Assessments

According to Huddleston and Rockwell (2015), standardized testing began in the middle 1800s. Prior to this time, the primary form of assessment and reading instruction was oral recitation until Horace Mann, a public education advocate, called for standardized written essay exams. Mann's supporters believed this tool would effectively assess the numerous amount of students who flooded schools because of the compulsory education law, which changed from educating the "elite to educating the masses" (Haladyna, Haas, & Allison, 1998, p. 262). The exams were utilized to identify the best instructional strategies, divide students into ability groups, and determine who would be promoted or admitted to colleges (Readence & Moore, 1983). Mann also proposed a curriculum change. He advocated that reading should be taught through a whole-word approach instead of exclusively with sound-symbol relationships, which

had been earlier promoted by the Colonists (Pearson, 2000; Rasinski, 2003). In addition, the first IQ test was one of the earliest forms of assessment developed by Binet and Simon in 1904, which was able to identify if a child's mental functioning was typically developing (Huddleston & Rockwell, 2015; Marks, 1989; Shepard, 2016). So, in the early 1900s, the shift from achievement testing to ability testing took hold with the goal of organizing and ranking students. Schools were looking to recognize students who were academically unsuccessful. However, despite criticism of standardized ability testing, it continued in the United States (Haladyna, et al., 1998; Marks, 1989).

Reading was one of the last subject areas assessed by a standardized test when William S. Gray, a famous reading theorist, developed an oral reading assessment in 1914 (Pearson, 2000). The complex process of reading, disagreements over the definition of reading, and how to assess comprehension impeded its development (Readence & Moore, 1983). In addition, reading instruction had always historically included reading aloud, which was difficult to assess. Since oral reading required students to be assessed individually, it was time-consuming to test using standardized assessments favoring oral reading. Therefore, reading assessments and reading instruction began to shift from oral to silent reading while oral reading was discouraged and discontinued in many schools (Pearson, 2000; Rasinksi, 2003). As silent reading became a more common instructional practice in schools because of its speed, comprehension, and efficiency, teachers began to focus on it during classroom instruction. Therefore, teaching to the test started the shift from oral reading to silent reading in schools (Huddleston & Rockwell, 2015; Pearson, 2000).

There were several forms of standardized test questions developed during the early 1900s in an attempt to assess comprehension (Huddleston & Rockwell, 2015). These included

passages, puzzles, and answering questions (Readence & Moore, 1983). Many of these formats became too subjective, took too much time to administer, and narrowed the breadth of comprehension behaviors (Huddleston & Rockwell, 2015; Readence & Moore, 1983). Eventually, the answering questions format came from Thorndike who created Scale Alpha, a standardized reading test, for grades three through eight. It was comprised of short paragraphs with increasingly difficult literal-level questions. There was no time limit, and answers were graded on a scale from 0-4 (Huddleston & Rockwell, 2015; Readence & Moore, 1983). This answering questions format became the model because it was a more objective, precise, and convenient measure (Readence & Moore, 1983). Answering questions were modified to the multiple-choice format that extended into the 1930s, the computerized scoring in the 1960s, and still exists today (Readence & Moore, 1983). This was significant because the multiple-choice format could provide data about many students at a very low cost, increasing comparison and competition between states, school districts, schools, and teachers (Haladyna, et al., 1998).

In the 1950s and 1960s, politicians once again became interested in using standardized testing for political reasons (Huddleston & Rockwell, 2015). In addition, taxpayers raised concerns that schools were spending tax monies without reporting what was being accomplished with it (Marks, 1989). Consequently, the ESEA Act was passed in 1965 and gave economic support to schools who served economically disadvantaged students (Huddleston & Rockwell, 2015). The law also required the expansion and accountability of standardized testing for schools in order to lessen the achievement gap of students of various economic backgrounds (Shepard, 2016). This opened the door for new and better uses of standardized tests to assess programs (Alcocer, n.d.). Additionally, ESEA was revolutionary because the federal government, which had historically left education up to the states and local leaders, began to give federal money to

educational institutions and had a part in decisions regarding policy (Duffy, Giordano, Farrell, Paneque, & Crump, 2008).

In 1969, NAEP was established to assess student achievement in reading and mathematics. The NAEP standards would be used as the national standards because each of the state's educational expectations was so diverse. Since the NAEP scores were divided by region, schools' and states' assessment results could not be compared (Ardoin & Christ, 2008).

The Commission on Excellence in Education published its report in 1983 called *A Nation at Risk*, which noted the bleak status of education in the United States because of worsening SAT scores, weak performance on comparison tests internationally, and low standardization scores of high school students. As a result, educators were urged by the Commission to go back to teaching basic skills, to implement standardized tests at many grade levels, and to compare American schools with other progressive nations (Huddleston & Rockwell, 2015; Sheninger & Murray, 2017). The report had a significant effect on the policies in education and on the spread of standardized testing (Huddleston & Rockwell, 2015). Consequently, four years later, 45 states, including the District of Columbia, were using statewide assessments to measure student achievement (Marks, 1989).

In response to *A Nation at Risk* report, the Commission on Reading Report, which is also known as *Becoming a Nation of Readers*, emphasized a balanced approach with both whole-word instruction and phonics (Alvermann, 1986). The study proposed that comprehension stemmed from automatic word recognition based on fluency, which came from letter and sound knowledge (phonics). The report insisted that phonics be taught by the end of second grade, in order for children to have an opportunity to read earlier and be able to comprehend quicker. This was a constructivist approach to reading instruction as teachers were urged to improve students'

comprehension by using a balance of phonics and authentic literature (Alvermann, 1986; Cowen, 2003; Farstrup, 2002). This led to the whole language movement in the 1980s, which continued to emphasize authentic literature, activities, and writing across content areas while de-emphasizing skills (McLaughlin, 2008).

The debate between whole language and phonics approaches gained momentum. As the debate continued, Congress requested a report on phonics and how it was related to reading instruction (Cowen, 2003). Marilyn Jager Adams was a lead researcher for the proposal, which was submitted by The Center for the Study of Reading. She discovered that students who utilized both phonics and meaningful text had exceptional results in reading achievement (Pearson, 2000). Therefore, the proposal suggested that reading instruction needed to include phonemic awareness, explicit phonics, authentic literature reading, and reading aloud. The downfall of whole language was further fueled by low reading scores on a California assessment by students who had been taught reading using whole language (Cowen, 2003).

As states continued to develop standards and high-stakes assessments, legislators continued to make decisions on educational issues and disperse funds. In the late 1990s, the NRP was created to review research on reading and to identify effective practices. The panel concluded that the five areas of reading, which are phonemic awareness, phonics, fluency, vocabulary, and comprehension, are essential in producing effective readers (National Reading Panel & National Institute of Child Health and Human Development, 2000). The report also reinforced the need for literacy professional development for educators (Cowen, 2003).

In 2001, after the NRP conclusions, NCLB ordered assessments and accountability for all states. Its accountability system was high-stakes because it either rewarded or punished schools and districts for the academic achievement of their students. Critics said that the accountability

system produced a narrowing of the curriculum, encouraged teaching to the test, placed too much importance on one test grade, and hurt the economically disadvantaged which could all result in the widening of achievement gaps among specific student groups (Dworkin, 2005). In 2015, ESSA, the new law, returned power from NCLB to the states and local school districts from the federal government. The ESSA continues to promote standards-based education, prepare students to be college and career-ready, and decrease the load of testing while still reporting yearly student achievement measures (Sheninger & Murray, 2017).

Standardized testing is engrained in American education. From our standardized testing history, we can conclude that testing has historically been and will always be a way to know how schools affect students (Haladyna, et al., 1998). As our understanding about reading assessments advances, so does our understanding of what makes reading instruction and reading assessments effective (Afflerbach, P., Kim, J. Y., Crassas, M. E., & Cho, B. Y., 2011). Reading instruction has settled on a more balanced approach of effective instructional strategies, including phonemic awareness, phonics, fluency, and comprehension. Current issues in assessing reading involve misalignment between standards and assessment, use of assessment to make instructional decisions, misinterpretations of scores, and the effects of assessments on students and teachers (Haladyna, et al., 1998; Pearson, 2000).

Early Reading Measures

Phonemic Awareness

From early encounters with oral language, reading, and writing, children realize spoken words have sounds. Phonemic awareness, according to the International Reading Association (IRA) and the National Association for the Education of Young Children (NAEYC) (1998), is “a child’s understanding and conscious awareness that speech is composed of identifiable units,

such as spoken words, syllables, and sounds" (p. 4). Yopp and Yopp (2000) further stated that phonemic awareness is a category under phonological awareness, which refers to a more general attentiveness of language's sound structures. Additionally, phonemic awareness more specifically includes being aware of the "speech stream" which consists of a sequence of phonemes, or small units of sound (Yopp & Yopp, 2000, p. 6). For example, students who are phonemically aware can identify the three sounds in the spoken word, book (/b/-/oo/-/k/) and can also blend these phonemes together to make a word (book). There are many aspects to phonemic awareness, according to Lane and Pullen (2004) and Yopp and Yopp (2006), but there is not a specific sequence to acquiring sounds (Cunningham, J., Cunningham, P., Hoffman, J., & Yopp, H., 1998). The aspects of phonemic awareness are the ability to hear syllables in a word, to hear initial letter sounds, to distinguish between rime and rhyme, to blend sounds together to make a word, to orally segment words, and to manipulate phonemes to make new words.

There are differing opinions regarding the necessity of phonemic awareness in becoming a proficient reader. The findings from the NRP (2000) indicate that one of the best predictors in learning how to read is phonemic awareness. Blevins (2006) and IRA/NAEYC (1998) claim children must be phonemically aware to become capable readers; therefore, instruction in phonemes is necessary. Cunningham, et al. (1998) proposes that phonemic awareness is also necessary but is not all it takes to become a good reader. Yopp and Yopp (2000) argue that phonemic awareness is only a part of a balanced literacy program in classrooms of pre-kindergarten and kindergarten students. In addition to phonemic awareness, expert readers also need to have an understanding of how phonemes are associated with words in print (Lane & Pullen, 2004). On the other hand, Smith (1999) claims that it is nearly impossible to separate individual sounds in words and calls phonemic awareness a false idea.

A quantitative study conducted by the NRP evaluated the relationship between phonemic awareness, reading, and spelling. Results showed phonemic awareness does influence students' comprehension and decoding (Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh, & Shanahan, 2001). Another study by Blachman, Tangel, Ball, Black, and McGraw (1999) followed students in a phonemic awareness study from kindergarten through second grade. For the treatment group, the phonemic awareness program in kindergarten was supplemented with 30 minutes of small group reading instruction in first and second grade, which included explicit, systematic instruction in phonics and phonemic awareness instruction previously received in kindergarten. Conversely, the control group received 30 minutes of basal reader instruction in addition to the basal workbook and a phonics workbook. At the end of second grade, the treatment group significantly outperformed the control group using a standardized measure of word reading and word identification. Students who had instruction in both phonemic awareness and phonics for reading in grades kindergarten through second grade where significantly above their peers in reading at the end of second grade (Blachman, et al., 1999).

Additionally, a study was conducted by Aspiranti, Hilton-Prillhart, Bell, and McCallum (2017) using the CBM Kindergarten-Monitoring Instructional Responsiveness: Reading (K-MIR:R) which measures crucial beginning reading skills. In the study, two administrations of the K-MIR:R and the Discovery Education Assessment (DEA) were given to 99 kindergarten students to determine concurrent validity. The researchers concluded that a screener for kindergarten should include multi-faceted items for letter matching, letter sound identification, and onset and rimes, as these are crucial skills needed for successful early readers. This supports research that says the most effective kindergarten assessments contain letter naming fluency, letter sound identification, blending onset and rimes, phoneme segmentation, and sound

repetition (O'Connor & Jenkins, 1999). As students enter first grade, screeners should assess phonemic awareness, decoding, and word identification, and reading text (Foorman, et al., 1998). The results of CBM Kindergarten-Monitoring Instructional Responsiveness: Reading (K-MIR:R) study also demonstrated that group administration is a valid measure and helps to decrease the testing time for the teacher as the K-MIR:R is accessible and economical.

Invernizzi, Justice, Landrum, and Booker (2004) argue that early literacy screening is important in identifying students with reading delays and help reduce future reading problems. They state the screener should look at reading skills across early literacy, which includes phonemic awareness, alphabet knowledge, concept of word, and phonics. Teachers must have knowledge about students among these four dimensions. The researchers also add that the screener should be easy and efficient to administer, valid and reliable, and the results should be transparent and easy to understand. Consequently, DIBELS Next, the early reading and universal screener used in this study, utilizes measures called First Sound Fluency (FSF) and Phoneme Segmentation Fluency (PSF) that assess phonemic awareness in kindergarten and PSF again at the beginning of first grade. The measures are included in the DIBELS Next composite scores used in the study (Dynamic Measurement Group, Inc., 2016).

Phonics

As students progress in their understanding of oral language and phonemic awareness, instruction in phonics and the alphabetic principle begins. Johnson and Keier (2010) define phonics as "the relationship between written and spoken sounds and letters" (p. 120). Phonics is a way of teaching reading which emphasizes the attainment of letter-sound relationships and uses the knowledge of those relationships to read (decode), spell (recode), and write (encode) words (Harris & Hodges, 1995; National Reading Panel & National Institute of Child Health and

Human Development, 2000). Research concludes that to read and comprehend connected text, students must demonstrate automaticity of the alphabetic principle, which is the understanding that alphabet letters and their corresponding phonemes can be utilized to read words (Harn, Stoolmiller, & Chard, 2008). Cunningham (2004) stated that children learn to read by concentrating on onset, rime, initial letter, and by looking at the length of words. Additionally, she emphasized that students learn these skills when they are taught, practiced, and supported in a meaningful context.

The role of phonics in literacy instruction has been an ongoing debate for many years. In her book, *Learning to Read: The Great Debate*, Jeanne S. Chall (1966) advocated for systematic and explicit phonics instruction because she said this results in better word reading, spelling, vocabulary, and comprehension. However, she concluded that the most effective approaches include phonics in conjunction with reading for meaning and reading connected text. As a result of her work, many reading commercial reading programs during the 1960s and 1970s changed their content from isolated phonics instruction to using more challenging stories and characters. More recently, Adams (1990) also found phonics knowledge essential, but it should be combined with reading instruction focused on meaning and purpose. Johnson and Keier (2010) studied beginning readers and discovered that they tend to rely too much on phonics without enough emphasis on meaning as they make substitutions for words; hence, their errors do not make sense. Beginning readers have a tendency to read on when it does not make sense and lack self-monitoring strategies (Johnson & Keier, 2010). Therefore, it is important to join phonics with meaning.

Phonics is measured by the DIBELS Next assessment included in the Nonsense Word Fluency (NWF) and the DIBELS Oral Reading Fluency (DORF). The measures are also included

in the DIBELS Next composite scores (Dynamic Measurement Group, Inc., 2016). Both of these measures are included in this study. Although phonics is important, studies continue to show the DORF measure as the best predictor of reading comprehension in first and second-grade students from all of the DIBELS Next measures. A study was conducted by Riedel and Samuels (2007) with 1,518 first grade students who were given the DIBELS measures during three assessment windows during the 2003-2004 school year. Students were also given the GRA+DE at the end of first grade to assess comprehension and the TerraNova was given at the end of second grade. The study examined the relationship between the DIBELS Next scores compared to the GRA+DE and TerraNova. Results showed that the ORF correctly classified 80% of the students. Furthermore, the NWF score was a better predictor of comprehension than PSF, classifying 68% of the students on the end of year assessment correctly compared to 53% for the PSF measure (Riedel & Samuels, 2007).

Goodman (2006) criticizes the use of these measures to predict reading comprehension. He disagrees with the idea that each skill must be mastered before moving onto the next skill as suggested by DIBELS Next. He also proposes that if students score poorly on these measures, classroom instruction will focus on these isolated skills, such as phoneme segmentation and reading nonsense words, without implementing other strategies that help overall reading. In addition, because the NWF measure is timed, he says this could affect proficient readers who might try to make sense out of nonsense words but are penalized for doing so (Riedel & Samuels, 2007).

Fluency

For years, reading fluency has been called the missing element in reading instruction (Rasinski, 2003). However, in 2000, the NPR recognized fluency as one of the five areas

essential for reading success. Fluency is the ability to read quickly, without effort, and with efficiency and meaningful expression. It is more than just reading accurately (Rasinski, 2003). Although many readers know how to accurately decode words, they may not be fluent or as automatic in their word recognition (Rasinski, 2003). Thus, they have to spend more time decoding, which takes away from comprehension. LaBerge and Samuels (1974) proposed the theory of automaticity that states readers have a limited amount of attention they can give to cognitive demands, such as reading unless they can alternate or can use one automatically (Rasinski, Reutzel, Chard, Linan-Thompson, 1996). Reading commands readers to be able to decode the words and comprehend the text. However, given the limited amount of cognitive resources available while reading and if attention is given primarily to decoding, then comprehension suffers (Griffith & Rasinski, 2004). Even if readers can decode words effortlessly, their comprehension is compromised if they are unable to give sufficient attention to making meaning from the text (Griffith & Rasinski, 2004). The ability to read words with automaticity is a feature of fluent reading. The theory of automaticity in reading gives a theoretical account for the importance of a reader to be able to read fluently (LaBerge & Samuels, 1974). One goal for reading fluency is to develop automatic decoding so that it requires minimal attention. The theory predicts a positive alignment between decoding automaticity and comprehension. When word identification and decoding become automatic as a result of practice, less attention has to be given to basic reading skills; thus, readers are able to give maximum attention to the real purpose of reading, which is comprehension (Walczyk, 2000). A study, sponsored by the United States Department of Education, showed the association between oral reading fluency and silent reading comprehension by asking fourth graders to read a text aloud which was rated on a rubric. They were then given a test to demonstrate their

comprehension of texts read silently. The most fluent readers in the study were also the ones who had the best comprehension. Every time the oral reading fluency declined, so did their silent reading comprehension (Rasinski, 2003). Therefore, we can conclude that fluency influences the efficiency of reading and comprehension (Griffith & Rasinski, 2004). When readers focus on oral reading fluency, they realize that words are not the only thing that gives the text meaning. Meaning is also made through intonation, phrasing, expression, pausing, and prosody (Rasinski, 2003).

Prosody or expressive reading is another important dimension of reading fluency. It refers to the ability to make reading orally sound like oral speech (Rasinski, et al., 1996). Fluent readers not only read with automaticity but segment text into appropriate and meaningful chunks or phrases. They also read with expression and attend to punctuation while reading. As they read, they are able to extract meaning from their oral interpretation of the text. Therefore, as readers learn to read with prosody, their expressive reading enables them to construct meaning from the text (Kuhn & Rasinski, 2011; Kuhn & Stahl, 2000). Beginning readers often have difficulty applying prosodic elements to phrases in texts. Recent research shows that fluency is an important factor in reading and is also associated with comprehension and student achievement (Kuhn & Stahl, 2000). Rasinski (1985) measured the qualitative use of prosody in oral reading using a quantitative rubric and found significant relationships between prosody according to the rubric and performance on a silent reading standardized test for third grade (r=.74) and fifth grade (r=.73) students. Likewise, the study of prosody and performance of fourth graders on a silent reading comprehension test by Daane, Campbell, Grigg, Goodman, and Oranje (2005) resulted in a substantial correlation and similar results.

Since reading fluency encompasses word recognition accuracy and automaticity in addition to prosody, this essential reading element can be assessed in various ways. Informal reading inventories and running records assess accuracy by determining the percentage of words the reader reads correctly when orally reading (Rasinski, et al., 1996). A student should be able to read a grade-level text with 90 to 95 percent accuracy. Any percentage below 90 may be a sign of weakness in decoding (Rasinki, 2003). Research has shown this assessment is a valid measure of reading proficiency (Fuchs, Fuchs, & Deno, 1982).

Reading rate is another way reading fluency can be assessed, assuming that quickly reading is an indication of automaticity with word recognition (Rasinski, 2003; Rasinski, et al., 1996). Deno (1985) created an approach called Oral Reading Fluency (ORF), which is a Curriculum-Based Measurement (CBM). His hope was that it would be an assessment that would be efficient, reliable, and valid. The ORF requires readers to read grade-level passages in one minute. The score from the ORF is the number of words read correctly in one minute. These scores are then compared to grade level norms to analyze. Readers who score at or close to the 50 percentile norms are said to be progressing satisfactorily in automaticity. Readers who are below the norm for their grade level and time of year may be referred to as at risk for developing reading fluency (Rasinski, 2003; Rasinski, et al., 1996). DIBELS Next is one assessment that utilizes the DORF as a test of oral reading fluency and is factored into each students' composite score in the middle of first grade and above. However, there are critics of reading rate as an assessment of reading because it emphasizes speed instead of comprehension. This has led some to the belief that the goal of reading instruction is reading quickly when, in fact, comprehension is the true meaning of reading (Kuhn & Rasinski, 2011; Rasinski, et al., 1996; Routman, 2003). Fluency without comprehension is merely "calling words" (Routman, 2003, p. 128). Routman

(2003) suggests that fluency is more of a predictor of comprehension for students in early grades and less in students who are in intermediate grades. Many students are able to read every word but cannot talk about their reading. If oral reading assessments and reading rate are going to be assessed, critics suggest comprehension must be assessed by using oral reading (Kuhn & Rasinski, 2011; Rasinski, Reutzel, et al., 1996). In addition, Rathvon (2004) states that early reading assessments need to include both fluency and accuracy to account for those students who read accurately but slowly. The information gleaned from such assessments will assist in, not only identifying those students who are at risk but also for planning purposes for intervention. Thus, the DIBELS Next DORF assessment includes a fluency score, an accuracy score, and a retell score (Dynamic Measurement Group, Inc., 2016).

Comprehension

Comprehension was also identified as one of the crucial areas of reading by the NRP (2000). The ultimate goal of reading is comprehension, which is critical to being a successful reader (Almasi & Hart, 2011; Rathvon, 2004). Comprehension is defined as the ability to obtain meaning with oral or written texts and is a constructive process (Duke & Carlisle, 2011; Rathvon, 2004). Children learn to read, so they can develop an understanding of what occurs in the text (Pressley, 2000). In order to comprehend, learn from, make connections, and remember ideas from texts, a reader must draw on his or her schema or previous knowledge to make sense or understand the message or text (Anderson, 2013). This is called schema theory. The schema of a reader is organized into repositories into which experiences are placed. The repositories assist readers in understanding reading, assimilating new learning and connecting it to previous learning, and remembering specific information by making connections to others (Pearson, 2000). Schema theory is closely related to Piaget's two types of learning: assimilation and

accommodation. In assimilation, new information is stored in schema that already exists. Accommodation is adapting the organization of our schemata to fit the new information (Pearson, 2010). When readers are reading new texts, such as on a reading assessment, the schemata activates. This allows strategic readers to make inferences about the text and use their schema to connect previous learning with new learning (Anderson, 2013).

Schema theory also states that since every reader's background knowledge is varied, more than one interpretation of a text is completely possible (Anderson, 2013). Frederic Barlett, who used schema to explain the interpretation of texts by using culture in the 1930s, said that readers actively make meaning of texts they read instead of just getting meaning from the texts themselves (Pearson, 2010). This aligns to the modern-day constructivists' view of reading comprehension that all readers have to create a model of reading for what they read (Pearson & Stephens, 1994).

Louise Rosenblatt (1994), one of the most influential educators on reading, also believed in the act of readers making meaning from texts. Her transactional theory explained that meaning does not exist in the text but is instead a transaction between the reader and the text at a particular time and context (Pearson, 2010). She proposed that meaning is not already in the reader or in the text, but meaning happens during the transaction between the two. The meaning is influenced by the prior knowledge of the reader, the experience of the reader, the reader's interest, and the reader's attitude. Therefore, it is possible for readers to arrive at various understandings and responses of the same, identical text (Squire, 1994). Rosenblatt (1994) further explains efferent reading and aesthetic reading. Efferent reading, which focuses on the meaning that is carried out of the text (informational), and aesthetic reading, which forms a

connection to the text through feelings, senses, and intuitions (literature), are on a continuum of how readers read and interpret a text (Rosenblatt, 1994).

The ability to decode and fluently read is vital for reading and for comprehension, but decoding itself is not enough to ensure accurate comprehension (Almasi & Hart, 2011). When word calling, automaticity, and fluency is emphasized in the early grades, it is often at the detriment to developing understanding (Allington, 2006; Routman, 2003). This can create "word callers" who have acquired sufficient skills for decoding but do not comprehend what they are reading (Cartwright, 2010; Stanovich, 1986). When working with struggling readers, teachers have a tendency to focus on decoding processes at the expense of instruction on meaning. In addition, when students read fluently, they are not as likely to be recognized as a struggling reader, as demonstrated in a study by Applegate, Applegate, and Modla (2009), even though they may not be comprehending (Allington, 1980; Cartwright, 2010). In the study, the teacher recognized 171 highly fluent readers in grades 2 through 10 to observe the relationship between fluency and comprehension. The researchers' study was motivated by a classroom teacher who said the best reader in her class was a word caller. The researchers gave students the Critical Reading Inventory-2, which included narrative passages at each students' grade level followed by a series of comprehension activities including retelling and open-ended questions. Results showed that one-third of the fluent and strong readers greatly struggled with comprehension. It is hard not to conclude that these students were chosen as strong readers only because of their fluency, accuracy, and prosody. Word callers also made up at least one-third of struggling readers in a study conducted by Riddle Buly and Valencia (2002). This study examined struggling readers who had failed state assessments. Results found six types of struggling

readers, which were automatic word callers, struggling word callers, slow word callers, word stumblers, slow and steady comprehenders, and disabled readers (Cartwright, 2010).

Some studies have concluded that teacher classification of word callers are undependable, such as the study by Hamilton and Shinn (2003). In addition, Meisinger, Bradley, Schwanenflugel, Kuhn, and Morris (2009) studied the relationship between teacher identification and the occurrence of word callers in elementary students. Word callers for the study were described as those with a normal fluency score of a minimum of 95 on the Gray Oral Reading Test-4 fluency scale (Meisinger, et al., 2009). The students also had below average comprehension skills (standard score of 85 on the Wechsler Individual Achievement Test–Reading Comprehension). This study concluded that the existence of word callers are found more in the late elementary grades than in the early elementary grades. These findings support the idea that fluency may become increasingly detached from comprehension as readers grow because as older students read more complex text, they need skills other than fluency to read. Also, some teachers may focus more on comprehension instruction than other teachers. Lastly, older students may comprehend better while they are silently reading as opposed to orally reading and so fewer word callers might have been identified if silent reading was used in the study (Meisinger, et al., 2009). We cannot presume all students who are word callers have comprehension difficulties as a result of the instruction they have received. Some students do struggle more with processing meaning than their peers (Cartwright, 2010).

Oral reading fluency measures from curriculum-based assessments could easily miss word callers, especially those in elementary grades (Meisinger, et al., 2009). Assessments of fluency without comprehension measures and instruction that separates fluency and word recognition from comprehension are misrepresentations of students' true reading ability

(Cartwright, 2010; Kuhn & Rasinski, 2011). Therefore, assessment of comprehension and its related skills is critical. For students who are beginning to read, word identification and phonemic decoding measures are the best at predicting reading comprehension. The ability to decode explains approximately 80% of the variance in reading comprehension for first graders (Foorman, et al., 1997). Therefore, decoding deficiencies are the main limitations on reading comprehension in early elementary students (Badian, 2001; Shankweiler, et al., 1999). Inaccurate and slow, laborious decoding limits reading comprehension (Rathvon, 2004). Since most beginning reader texts are rich in high-frequency words, students with poor decoding skills may be successful with accuracy and comprehension by memorizing high-frequency words. However, these students could be at risk for future comprehension difficulties because of their limited decoding ability (Rathvon, 2004). Byrne, Freebody, and Gates (1992) researched changes in reading skills of second and third-grade students that had discrepancies in their decoding and comprehension skills. Second-grade students who were weak in decoding, but strong in high-frequency word recognition demonstrated better comprehension than those who were stronger in decoding but weaker in high-frequency word recognition. However, by third grade, the pattern had transposed. The students who were the early better decoders now had better comprehension skills and the early weaker decoders had fallen behind in comprehension skills. Decoding continues to support comprehension skills throughout a students' school career (Rathvon, 2004).

Another assessment of comprehension is retelling. Retelling is a foundational reading skill that asks students to orally recall what happens in the text. If students are able to accurately talk about what they have read, they should be able to answer more advanced comprehension questions involving main idea, summarizing, and making connections (Dynamic Measurement Group, 2011b; Roberts, Good, & Corcoran, 2005). Retell also applies to both fiction and non-

fiction texts, and prevents students from guessing the answer based on background knowledge (Dynamic Measurement Group, Inc., 2011b). Reading comprehension research concerning retell confirms that it is a reliable and valid measure for assessing students' understanding (Marcotte & Hintze, 2009; Roberts, et al., 2005). DIBELS Next ORF includes a Retell Fluency (RF) measure and score. Studies concerning the relationship between DIBELS Next RF and comprehension are scarce. In two of the three studies that have been completed, there was a significant statistical relationship between RF and comprehension, but it was considerably weaker than the ORF (McKenna & Good, 2003; Roberts, et al., 2005). In another study by Pressley, Hilden, and Shankland (2005), there was no statistical relationship between RF and comprehension and the researchers expressed concern about scoring procedures for the RF. Riedel and Samuels (2007) suggest more research is needed to examine the effectiveness of administering the RF. In addition, the early text levels (levels A-E) in the TRC assessment includes retelling, which assesses students' foundational understanding of each book. Both DIBELS Next DORF and TRC assessments are utilized in this research study.

Curriculum-Based Measurement (CBM)

Accountability and assessment have become the mantra of education as a result of high stakes testing and increasing pressures at the national and state levels. The challenge for schools is to develop or provide assessments that are meaningful to teachers and parents, in addition to being useful for guiding classroom instruction. Curriculum-based measurement (CBM), developed by Deno (1985), is a group of measures that assess student progress and identifies students who might need extra instructional support and/or more testing (McGlinchey & Hixson, 2004). CBM measures growth over time as a progress monitoring measurement that uses standardized methods for test administration and scoring (Strecker & Lembke, 2005). Deno's

intent in the development of CBM was to test the effectiveness of data-based program modification (DBPM), which was an intervention model for special education. It was also designed to give special education teachers a tool that could be utilized to repeatedly assess students' academic interventions and use the data to evaluate the effectiveness of classroom instruction. To establish the validity and reliability of DBPM, a research and development program was managed for six years through the University of Minnesota Institute for Research on Learning Disabilities (Deno, 2003). CBM indicates student progress in reading, writing, spelling, and math and has only recently been used by classroom teachers as a result of RtI (Clarke, 2009).

CBM differs in many ways from other testing formats. With CBM, measures are fluency-based, which means that it is time efficient. CBM measures are one to three minutes long depending on what is being measured and many samples are given to increase validity (Deno, 2003). Also, the score from CBM indicates overall competence in that subject. The passages used are developed based on an end of year text complexity level. Another important feature of CBM is that the measures are given repeatedly across time. Students are given different, but equivalent materials that are from the same source. The equivalent materials hold task difficulty consistent and allow the teacher to make generalizations about student proficiency based on unfamiliar but similar texts (Deno, 2003). Since CBM is given across time, it gives students a chance to improve within the school year. Unlike standardized state assessments, which are only given once a year and do not provide data until the next school year, CBM is given regularly, so instruction and intervention can be quickly modified as needed (McGlinchey & Hixson, 2004). Therefore, both formative and summative assessment data can be gleaned from CBM. As a formative assessment, CBM offers repeated measures to direct instructional decisions.

Summatively, CBM provides data at a specific point in time to determine the student's level of mastery. The power of CBM is that it samples the entire, yearlong curriculum within its easy and quickly administered and scored measures (Clarke, 2009; Fuchs, 2016). Time is the biggest hindrance to implementing CBM. Teachers need training and practice in giving the assessments and in using the data to make instructional changes, which all take time (Clarke, 2009).

Several studies have shown a relationship between CBM reading scores and student proficiency on state assessments. On high-stakes assessments and in a variety of research studies, high correlations (r=.65-.85) have been achieved between CBM scores for reading and math and student performance (Deno, 2003). For example, a study in Oregon concluded that students, at the end of first grade, who could accurately read at least 40 words in one minute were on a path for reading success. In addition, students at the beginning of third grade were most likely to show reading proficiency on Oregon's state assessments if they can read more than 110 words correctly in one minute (Deno, 2003). Moreover, Stage and Jacobson (2001) studied the relationship between CBM ORF measure and the Washington Assessment of Student Learning (WASL). The ORF measure, which was given throughout the school year in September, January, and May, reliably predicted the WASL reading performance in May. The September oral reading fluency had positive predictive power and predicted WASL failure at r=.41. In addition, the negative predictive power from the September oral reading fluency predicted success on WASL at r=.90. These results confirm Deno's (1985) statement that oral reading fluency can be used as a "vital sign of reading achievement" (p. 224). However, he warned that the average number of accurate words read aloud in a minute does not uncover the complete reading processes of a student (Deno, 1985). McGlinchey and Hixson (2004) replicated the Stage and Jacobson (2001) study with the Michigan Educational Assessment Progam (MEAP) with fourth-time grade

students over a period of eight years with a larger sample of students and more diverse population. Their findings were similar in that a positive correlation was found between the two assessments. McGlinchey and Hixson (2004) concluded the positive relationship between CBM and certain state assessments may provide school districts with motivation to embrace and purchase CBM. This research study hopes to provide the school district used in the study with evidence to continue their use of CBM measures provided there is a positive correlation with third grade TCAP ELA scores.

Response to Intervention (RtI)

CBM can be beneficial as an effective measure within classrooms to provide information regarding the RtI process for students. RtI is a multi-tiered intervention process that is designed to help academically struggling students (Fuchs & Fuchs, 2006). In Tennessee, it is also a problem-solving model tailored to student individual needs in reading, math, and/or writing. Its emphasis is on high-quality instruction where instructional decisions are driven by student data. If a student does not respond to intensive interventions, it may be suspected that the student has a Specific Learning Disability (SLD). In this case, the student may be referred for special education interventions (Tennessee Department of Education, 2015). Traditionally, the only option available to students who did not succeed in the regular classroom was a special education placement. Most often these students did not demonstrate important discrepancies between achievement and intellectual ability until they were third graders. The discrepancy model has been a “wait to fail” model (Fuchs & Fuchs, 2007; Tennessee Department of Education, 2015, p. 7). As a result of the reauthorization of IDEA, an emphasis was placed on early intervention for students who are at risk for problems related to academics and/or behavior (Fuchs & Fuchs, 2007). Currently, schools must be proactive in providing intervention for students instead of

waiting for them to fail. The RtI framework in Tennessee represents intervention services that involve general education and special populations' staff collaboratively working together to meet the individual needs of every student (Tennessee Department of Education, 2015).

Tennessee's RtI framework is a visual representation of an inverted triangle that spans from general to special education. The foundation of RtI is Tier I where core instruction is received and grade-level expectations are provided to all students (Tennessee Department of Education, 2015). Tier I is designed to meet the needs of 80-85% of all students. Tier II and Tier III instruction is provided in reading, writing, and math depending on the needs of the student. Extra assistance is provided within Tier II to students who score between the 11th and 25th percentile in reading and/or math on the universal screener. Generally, 10-15% of students will need Tier II interventions, which occur in a 30-45 minute intensive group designed to meet the needs of the students as determined by the universal screener. Students who have not made significant progress in Tier II, in addition to Tier I, receive Tier III services. These students are either more than 1½ -2 years behind grade level or have fallen below the 10th percentile on universal screeners in reading and/or math. Tier III meets the needs of 3-5% of students in a 45-60 minute group comprised of three to five students where skills taught are more intensively and explicitly targeted to each individuals' area of need (Berkeley, Bender, Gregg-Peaster, & Saunders, 2009; Tennessee Department of Education, 2015).

Prior to RtI, there were many successful and unsuccessful attempts at providing intervention for students. The unsuccessful interventions helped put in place some of the foundational parts of the current model of RtI (Jimerson, Burns, & VanDerHeyden, 2007). Teacher referral, instead of proof of need, was the primary avenue to identify students who needed intervention. Because teacher beliefs would vary, some students had a greater chance or

less of a chance to be served in intervention. Students who needed extra instructional support may not have received it and those who did may have been served longer than necessary. Formative assessments were needed as evidence of student growth and ongoing progress monitoring was necessary to ensure intervention strategies were helpful to students. Prior to RtI, intervention became re-teaching of classroom content. However, re-teaching content, in the same way, was unhelpful to a struggling student (Kimmel, 2008).

Current RtI programs address these challenges to be sure students grow during their intervention experience. First, assessment is a chief component of RtI. The data that is obtained from the assessment drives the decision-making process. It is recommended that all Tennessee students in grades 3-8 be given a universal screening assessment three times yearly (beginning, middle, and end) in one or more academic area. The universal screener must be a brief, nationally normed, skills-based assessment of academic skills (Tennessee Department of Education, 2015). It determines which students need intervention and what skills need to be mastered. This process ensures students who need intervention are not overlooked and also prevents students who have mastered their targeted skills from having to stay in intervention programs longer than necessary (Fuchs & Fuchs, 2006). Students in Tier II and Tier III are assessed every five to ten days with progress monitoring assessments to determine the academic gaps that exist. The data also provides information to classroom teachers to ascertain if they need to change their strategies, materials, or instructional practices. In addition, data from progress monitoring gives diagnostic information that helps RtI teachers make intervention and placement decisions or if students need to be referred to special education (Fuchs & Fuchs, 2006; Tennessee Department of Education, 2015). Universal screeners and progress monitoring are the major forms of assessments required by federal policy for RtI (Tennessee Department of Education,

2015). The frequency with which these assessments are given is consistent across the nation with most states requiring universal screening three times yearly and progress monitoring in Tier II and Tier III two to four times monthly (Berkeley, et al., 2009). The administration, training, and fidelity of implementation of the universal screener and progress monitoring is the responsibility of the Local Education Agency (LEA) (Tennessee Department of Education, 2015).

Many kinds of universal screeners are available for RtI. Tennessee requires that school districts utilize universal screeners that have been nationally-normed reference in conjunction with percentile rankings (Tennessee Department of Education, 2015). In the context of RtI, CBM is a powerful, valid, and reliable tool as a universal screener and progress monitoring because it helps to identify students who need intervention, directs the most appropriate level of intervention, and evaluates if the intervention was successful (Mellard & Johnson, 2008). The literature cites DIBELS as an appropriate CBM measure because it includes a benchmark score with a predictive value of future student reading success and can determine students most in need of intervention (Abbott & Wills, 2012; Barnett, Daly, Jones, & Lentz, 2004). The school district in the research study utilizes DIBELS Next as the universal screener and progress monitoring assessment for RtI purposes for students in grades kindergarten and first grade. Researchers argue that a systematic monitoring system for intervention purposes, such as CBM, far exceeds the benefits of summative assessments (Shinn, 2010).

Measures Used in the Northeast Tennessee School District

Dynamic Indicators of Early Basic Literacy Skills (DIBELS) Next

The Dynamic Indicators of Early Basic Literacy Skills (DIBELS) Next assessment, utilized extensively in the United States to assess early reading skills, was developed by Roland

Good III and Ruth Kaminski from the University of Oregon (Riedel & Samuels, 2007). Its subtests or measures assess reading skills accentuated in the National Reading Panel report, which includes phonemic awareness, phonics, fluency, vocabulary, and comprehension (National Reading Panel & National Institute of Child Health and Human Development, 2000). The authors state the five measures are indicators of critical skills every student must learn for reading proficiency. The purposes of the assessment are to early identify students who need intervention, to predict future reading difficulties, to assist teachers in identifying areas of needed instructional support, and to appraise school districts' instructional supports systems (Dynamic Measurement Group, 2011a.; Good, Simmons, & Kame'enui, 2001; Rathvon, 2004). Similar to CBM, DIBELS Next are fluency-based measures designed to be concise, (one to three minutes each), useful, cost-effective, individually administered, formative, and normalized (Dynamic Measurement Group, 2011a). While being used in over 13,000 schools nationwide, the DIBELS Next subtests intend to measure phonemic awareness and phonics of kindergarten and first-grade students. The subtests include First Sound Fluency (FSF), Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF). Another test, DIBELS Oral Reading Fluency (DORF), is given to students beginning in the middle of first grade through sixth grade, which assesses students accuracy, fluency, and comprehension (retell) in reading connected texts (Riedel & Samuels, 2007).

A combination of raw scores for each individual subtest and an overall composite score is provided by DIBELS Next. The scores are interpreted in relation to DIBELS Next benchmark goals and cut points to conclude if students' data are above, at, below, or well below the benchmark (Dynamic Measurement Group, 2016). The benchmark goals are "empirically derived, criterion-referenced target scores that represent adequate reading skill for a particular

grade and time of year" (Dynamic Measurement Group, Inc., 2016, p. 1). The benchmark goals and composite scores were developed using collected data from a 2009-2010 study involving Group Reading and Diagnostic Evaluation (GRADE) as the external standard. The study was conducted with 3,816 on and below grade level students in grades kindergarten through sixth grade in 5 states and 13 elementary and middle schools. DIBELS Next states that if students reach benchmark goals, it is likely they will meet the next DIBELS Next benchmark goal if they receive effective classroom reading instruction. For students who do not reach benchmark goals, it is likely that intervention or support will be needed of them to achieve later reading goals (Dynamic Measurement Group, Inc., 2016; Kaminski & Cummings, 2007).

The DIBELS Next composite score is a grouping of multiple scores on subtests and gives the best overall estimate of students' early reading skills and proficiency. The composite score should always be interpreted first. If students are at or above the benchmark goal on the DIBELS Next composite score, they are more than likely to reach later important reading milestones. However, even if students met the goal of the composite score, they still may need additional support in early reading skills as evidenced by the benchmark score on an individual subtest. This is especially true for students with a composite score close to the benchmark goal. The measures that are used to compute the composite score are different by grade level and time of year, so the composite score does not indicate growth across grade levels or within a grade (Dynamic Measurement Group, Inc., 2016; Kaminski & Cummings, 2007).

Assessment occurs three times a year with all students using the grade-level DIBELS Next benchmark measures. The initial screening occurs in the fall and identifies students who may be considered at risk for reading difficulty and end-of-year reading goals. DIBELS Next can also be used for progress monitoring with students who are at risk on the benchmark assessment.

The subtests are given to students more frequently in the skill area in which the students are receiving intervention or instruction in between assessment windows. This helps to make sure students are making satisfactory progress as learning and intervention are occurring. When progress monitoring, students can be given materials at grade-level or out-of-grade level, depending on what is needed for each student (Dynamic Measurement Group, 2011a). The data from DIBELS Next measures can be beneficial to determine the effectiveness of intervention and instruction in RtI (Dynamic Measurement Group, 2011a; Hagan-Burke, Burke, & Crowder, 2006; Rathvon, 2004). The other two assessment windows are the middle and end of year.

The school district in this study chose DIBELS Next as a universal screener for kindergarten and first-grade students in 2006-2007 because it was included in a suite of assessments called mCLASS: Reading 3D provided by Wireless Generation, now Amplify Education. The school district was aspiring to collect and store assessment data digitally and the assessment suite offered those capabilities. The suite also includes the TRC assessment and a Word Recognition (WR) assessment to provide assessments that demonstrate students' behaviors while reading connected text and high-frequency words. Beginning with Palm Pilots and now with iPads, teachers and RtI interventionists from the school district administer mCLASS: Reading 3D to students three times a year as described above.

Many of the research studies that have been conducted involve DIBELS 6th Edition. The assessment used in this research study is DIBELS Next. DIBELS Next differs from DIBELS 6th Edition in that the FSF subtest on DIBELS Next replaced the Initial Sound Fluency (ISF) subtest from the 6th edition. The NWF subtest now contains revised directions and scoring rules, which include whole words read (WWR) correctly. In addition, the DORF utilizes new passages, which were field-tested and are leveled empirically (Dynamic Measurement Group, Inc., 2011a.).

DIBELS Criticism

DIBELS has drawn criticism from many researchers in the field of reading. The most common criticism is that it does not adequately assess comprehension (Goodman, 2006). However, both supporters and opponents agree comprehension is the main goal of reading. If DIBELS subtests are aligned to comprehension, they can be utilized to identify those students in need of comprehension intervention. However, if they are not properly aligned, resources and intervention services could unnecessarily be given to students (Crowell, 2015; Riedel & Samuels, 2007). Goodman (2006) states that fluency is what rules in the DIBELS assessment, which places an over-emphasis on speed. Furthermore, speed can be taught to students, so that what appears to be making progress in reading, may only be evidence of quicker performance on the test. Thus, we now have students who can read quickly and accurately, but do not comprehend what they are reading (Allington & Pearson, 2011). Samuels (2006) claimed that ORF is not a true test of fluency since fluency involves both decoding and comprehension. On the ORF subtest, the focus is on speed and decoding words; therefore, it does not sufficiently assess comprehension even through its retell measure. In addition, there are concerns about both the validity of the retell portion of the subtest and the reliability of its scoring procedures (Riedel & Samuels, 2007).

Critics of DIBELS say it is unclear if reading comprehension is related to the PSF or NWF subtests since few studies have been conducted or failed to find a correlation (Riedel & Samuels, 2007). One of the few studies involved Cook (2003) who found a substantial correlation between the PSF and NWF and the Woodcock-Johnson Total Reading Cluster. On the other hand, two studies found no sufficient correlation between PSF scores of first graders and the Stanford Diagnostic Reading Test (Kaminski & Good, 1996). Additionally, Goodman

(2006) disagrees with the theory of DIBELS that specific reading skills (reading nonsense words or segmenting phonemes) must be learned before moving onto other skills, such as fluency and comprehension. He warns that students who are shown to be at risk with these specific skills may experience classroom reading instruction that is void of authentic reading involving making meaning from text and strategy-building (Goodman, 2006; Pearson, 2006).

Since DIBELS Next is used for progress monitoring purposes in RtI, critics address the tendency for some teachers to teach to the subtests, which gives them too much credibility as reliable reading measures (Alllington & Pearson, 2011; Manzo, 2005). Since teachers are asked to provide intervention for low-performing students and the subtests are so specific, it is not surprising that teachers teach to the test, which helps improve their students' scores. Furthermore, because there is such an emphasis on testing accountability, many teachers teach what is tested, which leads educators to teach to students' weakness on the DIBELS assessment. By primarily focusing on early reading skills, time spent with stronger literacy practices and independently reading is greatly reduced in the classroom (Allington & Pearson, 2011). Critics claim this has led to a narrowing of the curriculum (Tierney & Thome, 2006). Tierney and Thome (2006) state, "Indeed, DIBELS may be perpetuating the literacy gap it has promised to eliminate" (p. 53).

Studies Involving DIBELS ORF

DIBELS authors maintain that the subtests are helpful in predicting future reading difficulty and are also useful in timely and accurately identifying students who need intervention (Good, Simmons, & Kame'enui, 2001). Despite the few studies involving PSF and NWF, there have been many studies conducted that link the formative value of oral reading fluency to student performance on high-stakes assessments. Buck and Torgesen (2003) studied third-grade

students' curriculum-based measures of oral reading fluency scores and their performance on the Florida Comprehensive Test-Sunshine State Standards (FCAT-SSS) from thirteen schools in one school district. The median score for the ORF passages was used to evaluate the relationship. A significant correlation occurred between the assessments (r=.70, p<.001).The results show that 91% of students who had an oral reading fluency score at or above 110 correct words per minute (cwpm) also achieved a level 3 or above on the FCAT (Buck & Torgesen, 2003). This section will describe studies that specifically compared DIBELS ORF to performance on summative state assessments.

A research study by Shaw and Shaw (2002) studied the predictability of DIBELS ORF to third-grade placement levels on the Colorado State Assessment Program (CSAP). They concluded that the DIBELS was a strong predictor of the CSAP score with correlation coefficients ranging from .73 on the fall and winter assessments to .80 on the spring assessment. Similarly to Buck and Torgesen (2003), Shaw and Shaw (2002) found that 90% of the students, who scored 110 or above on the spring ORF, scored proficient or advanced on the CSAP.

The median scores from thirty-eight third grade students on the DIBELS ORF were compared with their reading achievement on the North Carolina End of Grade test (EOG). The DIBELS ORF included students reading three passages orally for one minute each. The EOG consisted of reading passages with multiple-choice questions. The assessments were given a short time apart, so the results were considered a correlation instead of a prediction. The correlation was high (r=.73). Of the students in the study, 63.2% of the students scored at Level III or higher on the North Carolina EOG and reached the goal of 110 cwpm or better on the DIBELS ORF. The correlation was weak for the students who scored less than 69 cwpm on the ORF (Barger, 2003).

Likewise, Wilson (2005) compared the median scores of third-grade students on the DIBELS ORF to determine a correlation with the Arizona Instrument to Measure Standards (AIMS), which is a multiple-choice comprehension assessment. The results indicated there was a positive correlation (r=.741) between DIBELS ORF and AIMS. Students who had higher fluency levels on the DIBELS ORF scored higher on the AIMS. Students who were identified as at-risk on the DIBELS ORF did not meet proficiency on the AIMS. Therefore, the study concluded that the DIBELS ORF scores can predict student proficiency levels on the AIMS (Wilson, 2005).

Roehrig, Petscher, Nettles, Hudson, and Torgesen (2008) studied the predictive and concurrent validity of the DIBELS Oral Reading Fluency (ORF) to predict performance on the Florida Comprehensive Assessment Test (FCAT-SSS) and Stanford Achievement Test (SAT-10), which is a reading comprehension assessment. The three assessments were given during the same time period. The study consisted of 16,539 third grade students in schools that profited from Reading First funding. The third administration of the DIBELS ORF in February/March resulted in the strongest correlation with both the FCAT-SSS and SAT-10 (r=.70, r=.71). The fall administration also showed moderately strong correlations (r=.66, r=.68). The researchers from the study noted it is vital that educators have accurate student assessment data if rigorous state grade-level reading standards are to be achieved by all students as a result of accountability measures and policies. Since oral reading fluency has been shown to predict student achievement on state assessments, the validity and utilization of such measures are justified (Roehrig, et al., 2008).

Research studies indicate that oral reading fluency is more highly correlated with comprehension than other DIBELS subtests. Riedel and Samuels (2007) present a number of

hypotheses for why this may be true. The structure of the DIBELS ORF itself more closely mirrors a typical reading comprehension assessment since students are reading connected text than the other DIBELS subtests. Also, being able to read a connected text quickly and with accuracy possibly plays a role in comprehension, which might explain the relationship between comprehension and oral reading fluency. In addition, students who are considerably behind their peers by the middle of first grade may have learned lower-level reading skills, such as the PSF, which makes the subtest less likely to separate good readers from poor readers (Samuels & Riedel, 2007). For these reasons, there may be more of a relationship between DIBELS ORF and reading comprehension than any other DIBELS subtests.

Text Reading and Comprehension (TRC)

Text Reading and Comprehension (TRC) is a standardized, formative reading assessment designed to assist educators in monitoring and observing reading development in students and to help with the identification of reading difficulties. This early literacy formative assessment, which is aligned to Common Core State Standards (CCSS) and includes DIBELS Next, is part of mCLASS: Reading 3D created and published by Amplify. The TRC assessment permits educators to use a running record to analyze the student's ability to accurately read and comprehend grade-level texts. The student reads a text aloud while the teacher records the student's oral reading behaviors, which determines the student's reading accuracy. The comprehension portion of the TRC assists teachers in determining the student's understanding of the text by asking a combination of retelling, open-ended, and written comprehension questions. The written comprehension questions, which are optional, are not used by the school district in the study because of the administration time. The accuracy rate and comprehension scores determine the overall instructional reading level of the student at three different administration

periods throughout the year, beginning, middle, and end of year (Gushta, Parisi, Richards, & York, n.d.).

TRC assessment was originally developed by Wireless Generation led by Larry Berger and Gregory Gunn. After observing teachers give students oral reading tests by using paper and pencil, they believed that kind of scoring was an outdated practice. They thought a mobile device that scored running records would be easier and faster. Additionally, the device could sync to a software system on a secure website, so data would be readily available and archived for future use. Consequently, they developed the TRC, which is a computerized running record assessment. It was developed, tested in the field, and joined DIBELS Next in the suite of resources known as mCLASS: Reading 3D (Gushta, et al., n.d.; Snow, et al., 2018; W.K. Kellogg Foundation, n.d.).

TRC assessment uses 76 fiction and non-fiction texts published by Harcourt Publishing within the Atlas book set. These books include print concepts and reading behaviors and correspond to levels A to Z of a gradient of text difficulty (Fountas & Pinnell, 1996). Amplify conducted an analysis of the texts, both qualitatively and quantitatively, to examine the text complexity across reading levels. TRC measures growth in accuracy, fluency, and comprehension at consecutively harder levels of text complexity (Gushta, et al., n.d.). Data acquired from the TRC assist educators in determining students' reading level and better match students to texts. Research shows students should read with at least a 90-94% rate of accuracy in order to comprehend text, which means students can read these texts independently with some help from a teacher or peer. This is called a student's instructional reading level. Texts at this level appropriately challenge students' reading abilities without impacting their comprehension. The independent level of a text is easier for students and gives them an opportunity to build fluency and confidence. Independent reading levels are texts students can read with 95% or

higher accuracy and above 80% comprehension. In addition, the frustrational level of a text is one that is too challenging for students to read with appropriate accuracy and comprehension. These are texts that students read with less than 90% accuracy and below 80% comprehension. Students need extra support in reading these texts in a shared reading atmosphere (Clay, 2002; Fountas & Pinnell, 1999; Fuchs, Fuchs, & Deno, 1982).

Knowing and understanding the independent, instructional, and frustrational text levels of students is extremely valuable. It assists teachers in differentiating Tier I instruction so they can group students who read at comparable levels, recognize students who need intervention, and identify the specific instruction students need to become proficient readers (Clay 2002; Snow, et al., 1998). The TRC also helps school districts decide which students are making progress toward the end of the year goals and those who are making progress to meet the end of the next school years' goals (Zhao & Von Secker, 2008). A study conducted by Ross (2004) concluded that schools that implement running records as an assessment system, such as TRC, have students who exhibit higher achievement on reading and writing tests than schools that utilize other classroom assessment programs.

TRC is a running record assessment based on an approach by Marie Clay (2005) that gives teachers a way to assess how students perform on foundational skills needed to become a fluent reader and how to utilize those skills in reading more complex texts (Gushta, et al., n.d.). Running record assessments are a way for teachers to capture and analyze the miscues of readers by focusing on a student's oral reading behaviors using a standardized coding system, which provides a visual representation of the reading (Afferbach, Kim, Crassas, & Cho, 2011; Johnson, 2006). In the 1960s, Kenneth Goodman coined the term "miscue" when he recognized that errors made while reading are actually "windows" into the process of reading (Pearson & Stephens,

1994, p. 27). The miscues made by a reader actually reveal the reader's ability to make sense of what he or she is reading (Crowell, 2015). Miscues give insight into how students decode print, involve prior knowledge, fluently read, construct meaning, and self-monitor their comprehension. This insight can enlighten the teacher about the reader's sound-symbol correspondences and literal and inferential comprehension (Afferbach, et al., 2011). Goodman also discovered readers use cueing systems, which are syntactic cues, semantic cues, and graphophonemic cues, to make sense of text (Pearson & Stephens, 1994). These cues can be used for identifying specific strengths and weaknesses of the reader, so precise needs can be addressed during reading instruction (Afferbach, et al., 2011; Clay 2002).

Comprehension is also a vital component of the TRC assessment. The comprehension of students is assessed in three formats: retell, oral comprehension, and written comprehension (optional). The retell is scored on a rubric from 0 to 3 according to the number of details from the beginning, middle, and end of the text that the reader provides. The retell questions are included in early level texts (Levels A through E) that contain less content for high-quality text-dependent questions. Five oral comprehension questions are assigned to books at levels D through Z. The questions align to the CCSS for English Language Arts, and each book contains one question or more questions aligned to the CCSS substrands. Bloom's Taxonomy and Webb's Depth of Knowledge were also used in the creation of questions to ensure the appropriate level of cognitive complexity (Gushta, et.al., n.d.). A score of 80% on the comprehension questions equates to proficient while lower percentages equate to non-proficient. The school district in the study does not administer the optional written comprehension questions to students. Snow, et al., (2018) found poor performance on the written comprehension questions from TRC led to the underestimating of the instructional reading level of primary students and resulted in a longer

administration time for first graders (*M*=40.1) as compared to an informal reading inventory (*M*=13.2).

TRC utilizes cut points, which were developed by eleven leveled-reading experts representative of different educational backgrounds and geographical locations for various performance levels (far below proficient, below proficient, proficient, and above proficient). The CCSS for ELA steered the specific proficiency levels for each assessment window (beginning of year, middle of year, end of year) and for each grade K-6. These were developed based on the Item Descriptor Matching Method, which is appropriate for assessments that produce results by categories. Far below proficient identify students who do not demonstrate any of the reading behaviors, which are contained in the CCSS ELA. Below proficient students show some or few of the expected behaviors but not all. Proficient students exhibit minimal CCSS ELA expectations. Students who show reading behaviors that exceed expectations of CCSS ELA are above proficient. The literacy experts ensured the cut points for the beginning and middle of year would allow students to be ready to meet the end of year expectations (Gushta, et al., n.d.).

In addition, the Atlas book set participated in field-testing to demonstrate the validity and reliability of TRC and the texts as an oral reading accuracy and comprehension assessment. The progression of the complexity of the text levels (which was theorized from levels A to Z) was approved by the field study based on how the students performed on the texts and "the statistical estimates of text difficulty" (Gushta, et al., n.d., p. 22). From the student field test data, Amplify was able to acquire qualitative feedback in order to enhance texts and comprehension questions (Gushta, et al., n.d.).

Studies involving TRC

Every kindergarten through third grade-student (approximately 400,000 students) in North Carolina is administered the Amplify mCLASS: Reading 3D assessment, which includes the TRC, three times per year since 2013. Therefore, several research studies in North Carolina have included TRC. A study published by Amplify Education investigated the validity evidence on mCLASS: Reading 3D (DIBELS Next and TRC) with student outcomes on the North Carolina End-of-Grade Reading Comprehension Test (NC EOG). The sample of third-grade students (N=53,890) at 850 schools were matched with their mCLASS: Reading 3D data and their NC EOG data. The studies showed that as student performance rises on DIBELS Next and TRC, so does their performance on the NC EOG. There was a strong correlation with NC EOG for the final instructional reading level for TRC (r=0.71, p<0.05) and for DIBELS Composite score (r=0.74, p<0.05). Students who were proficient on both DIBELS Next and TRC were also proficient or above on the NC EOG (68%). Likewise, students who were not proficient on both DIBELS Next and TRC were also usually not proficient on the NC EOG (96%). The overall statistical analysis showed that DIBELS Next and TRC together provide 79% accuracy in predicting proficiency on the NC EOG Grade 3 reading assessment (Amplify, 2013b).

Bowles (2015) studied the relationship and the predictability between the mCLASS Reading 3D assessment (DIBELS Next and TRC) and the NC EOG reading assessment from 143 students in grades 3-5 in North Carolina. Results showed that all grade levels had the same proficiency levels on mCLASS Reading 3D and NC EOG except TRC scores of fifth graders. An analysis of the multiple regression test indicates that the DIBELS Next ORF and TRC scores significantly predicted NC EOG scale scores, $F(2, 55)=38.728$, $p<.05$, adj. $r^2=.570$. TRC statistically significantly predicted scale scores on the NC EOG at all grade levels in the study,

with fifth-grade students showing the strongest correlation with *r*=.616. As a result of increased accountability in school districts, it is vital that assessments, which are being used for instruction and intervention, are aligned (Bowles, 2015).

Tennessee Comprehensive Assessment Program (TCAP) TNReady

Since 1988, the Tennessee Comprehensive Assessment Program (TCAP) has operated as the state's testing program for second through eleventh-grade students. Tennessee began a new assessment in 2015-2016 for students in third through eighth grades called TNReady, which is aligned to Tennessee's more rigorous standards. This statewide, timed, criterion-referenced assessment surpasses the requirements of ESSA and requires students to take yearly assessments in mathematics, English language arts, science, and social studies (Tennessee Department of Education, 2017; Tennessee Department of Education, 2018; Tennessee Department of Education, n.d.b.). The TCAP serves these main objectives: provides feedback about students' academic progress and how it aligns with grade-level expectations; gives parents and teachers a big-picture perspective about how a student is progressing compared to peers across the district and state, including a student's strengths and growth opportunities; builds confidence and transparency about students' readiness for postsecondary and the workforce among Tennessee colleges, universities, and employers; helps educators strengthen instruction and reflect on their practice; holds us accountable to serving all students fairly; and highlights schools where students are excelling, so we can learn from those who are doing well (Tennessee Department of Education, n.d.c.).

TCAP TNReady was created to measure critical thinking and higher expectations for the students of Tennessee (Tennessee Department of Education, 2017). Item types on TCAP TNReady ELA extend beyond multiple-choice items, and include evidence-based selected

responses (ESBR), multiple select, writing prompts, and editing tasks (Tennessee Department of Education, n.d.e). The questions created for the TCAP TNReady experience a three-step evaluation process. During the first step, the Tennessee Department of Education develops expectations for item development aligned to the state standards. Teachers and the test vendor create test questions. Next, the questions are examined for content and bias by the Tennessee Department of Education and by a group of Tennessee educators, administrators, and supervisors who are represented by grade level or grade band. Educators are able to accept test items, revise, or reject items. Then, the Educational Testing Service (ETS) (2012) reviews and revises items based on that feedback. Each question is then field tested and reviewed for statistical validity. Once the validity is established, the testing vendor adds the questions to the operational assessment, and the entire test is reviewed for accuracy by the Tennessee Department of Education. Finally, the final assessment is administered in Tennessee schools and classrooms (ETS, 2012; Tennessee Department of Education, n.d.b).

TCAP TNReady ELA assesses the Tennessee Academic State Standards by using both literary and informational texts for students to demonstrate their knowledge of close reading, analyzing texts, answering text-based questions, present a response to a writing prompt, and demonstrate an understanding of the English language. The assessments are given to students in four subparts and evaluate students' understanding of comprehension, foundational literacy, fluency, language, spelling, and writing. TCAP TNReady scores are classified into these performance levels: mastered (Level 4), on track (Level 3), approaching (Level 2), and below (Level 1) (Tennessee Department of Education, n.d.e).

Testing times for TCAP TNReady assessments are released each year by the Tennessee Department of Education. Grade 3 ELA includes four subparts, which total 216 minutes. Overall,

the state-mandated testing for 2018-2019 will include 345 total minutes, including ELA, Math, and Science tests. This time is lower than in previous years because students will not take a TCAP TNReady in Social Studies as new standards are implemented in 2019-2020 (Tennessee Department of Education, n.d.e).

In addition to yearly state assessments, most school districts, like the one in this study, also give students universal screeners, benchmark tests, and other district assessments (Nelson, 2013). Nelson (2013) found students in grades 3-8 spending 80 hours per year on test preparation, which involved completing practice tests and learning test-taking strategies. Furthermore, students spent 15 hours or more per year taking the required assessments (Nelson, 2013). The State Collaborative on Reforming Education (SCORE) (2015) conducted a survey, interviews, and focus groups of Tennessee's principals, teachers, and superintendents to give these stakeholders voice in the improvement of assessment. When asked if too much time, about right, or too little time is currently spent on assessments, 51% of school leaders, 56% of principals, and 74% of teachers responded with too much time. Additionally, teachers responded that the top challenge faced with assessments is reduced instructional time in the classroom (SCORE, 2015).

Since there is a plethora of time spent on preparing and administering assessments to students in this high-stakes educational environment, educators need to ensure assessments are utilized for appropriate purposes. Popham (2001) advocates against using state assessments for evaluating schools and teachers, grading and promoting students, and utilizing the data for daily instructional decisions in the classroom since the items on the assessments only represent a small sample of standards-aligned questions. Instead, he acknowledges state assessments should be used for informing parents and teachers about student achievement, identifying students for

special programs (gifted, RtI), and allotting optional instructional resources. In the SCORE report, Tennessee teachers stated that state and district assessments are useful for teaching and learning. They used assessments mostly to diagnose skill deficits of students, to improve instruction, to set goals and group students, and to predict future student achievement. School leaders added that state and district assessments are also useful for predicting student performance on state end-of-year assessments (SCORE, 2015).

Chapter Summary

As a result of recent national legislation presented in the literature, educators have focused their efforts like never before on beginning reading and the importance of students' early success. Early reading measures, such as phonemic awareness, phonics, fluency, and comprehension, are essential pieces of reading instruction and necessary to produce proficient readers (Aspiranti, et al., 2017; Invernizzi, et al., 2004). Furthermore, it is imperative that school districts have broad, efficient, valid, and reliable early reading screeners to assess students' knowledge of important literacy concepts (Invernizzi, et al., 2004). The knowledge gained from the data of the reading screeners assists educators in identifying students who are in need of extra reading instruction and the use of specific intervention strategies for those students. They also help to provide purposeful classroom instruction. Moreover, the early reading screeners utilized in school districts should have a significant relationship to other high-stakes reading assessments as students grow as readers. The literature supports significant relationships between curriculum-based assessments as early reading screeners and standards-based state assessments. Even though the DIBELS assessment is highly criticized, there is a strong relationship between DIBELS ORF and state assessments as presented in the literature. In addition, the text levels of TRC also boast a relationship to state assessments when it is utilized. Consequently, the significant relationships between early reading screeners and state assessments could potentially assist educators in more

accurately identifying what students need in early reading instruction, lessen the need for ongoing intervention, and, most importantly, increase students' overall reading success.

CHAPTER 3

Research Design and Methods

The purpose of this non-experimental quantitative research study was to examine the relationship between kindergarten and first-grade students' Dynamic Indicators of Basic Early Literacy Skills Next (DIBELS Next) and Text Reading and Comprehension (TRC) scores and third-grade students' Tennessee Comprehensive Assessment Program (TCAP) English-Language Arts (ELA scores). The study was a quantitative predictive correlational research design to determine if there was a relationship between two or more variables. The second purpose of this study was also to determine the predictive validity of kindergarten and first-grade DIBELS Next and TRC scores on end of third-grade TCAP ELA performance. This chapter detailed the design, methods, and procedures used to answer the six research questions used in this study. The population, the sample of students, and descriptions of the instrumentations used in the study were examined. The explanation of the data collection and data analysis, including descriptive statistics, were also explained in this chapter.

Research Questions and Null Hypotheses

Eight important research questions arise to address the purpose of this study:

1. Is there a relationship between kindergarten DIBELS Next composite scores and third- grade TCAP ELA scores?

   $H_0$1: There is no relationship between kindergarten DIBELS Next composite scores and third-grade TCAP ELA scores.

2. What is the relationship between first-grade DIBELS Next composite scores and first-grade DIBELS Oral Reading Fluency accuracy scores on third-grade TCAP ELA scores?

$H_0$2: There is no relationship between first-grade DIBELS Next composite scores and first-grade DIBELS Oral Reading Fluency accuracy scores on third-grade TCAP ELA scores.

3. What is the relationship between first-grade DIBELS Oral Reading Fluency fluency scores and first-grade DIBELS Oral Reading Fluency retell scores on third-grade TCAP ELA scores?

   $H_0$3: There is no relationship between first-grade DIBELS Oral Reading Fluency scores and first-grade DIBELS Oral Reading Fluency retell scores on third-grade TCAP ELA scores.

4. Is there a relationship between kindergarten Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores?

   $H_0$4: There is no relationship between kindergarten Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores.

5. Is there a relationship between first-grade Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores?

   $H_0$5: There is no relationship between first-grade Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores.

6. Is kindergarten DIBELS Next composite score or kindergarten Text Reading and Comprehension (TRC) level of proficiency a better predictor of future performance on third-grade TCAP ELA?

   $H_0$6: There is no relationship between kindergarten DIBELS Next composite score or kindergarten Text Reading and Comprehension (TRC) level of proficiency on third-grade TCAP ELA.

7. Is first-grade DIBELS Next composite score or first-grade Text Reading and Comprehension (TRC) level of proficiency a better predictor of future performance on third-grade TCAP ELA?

   $H_0$7: There is no relationship between first-grade DIBELS Next composite score or first-grade Text Reading and Comprehension (TRC) level of proficiency on third-grade TCAP ELA.

8. Is kindergarten DIBELS Next composite score, first-grade DIBELS Next composite score, kindergarten Text Reading and Comprehension (TRC) level of proficiency, or first-grade Text Reading and Comprehension (TRC) level of proficiency the best predictor of future performance on third-grade TCAP ELA?

   $H_0$8: There is no predictor of future performance by kindergarten DIBELS Next composite score, first-grade DIBELS Next composite score, kindergarten Text Reading and Comprehension (TRC) or first-grade Text Reading and Comprehension (TRC) on third-grade TCAP ELA.

Population and Sample

The population of this study was 608 students who were third graders in 2017-2018 from eight elementary schools in the Upper East Tennessee school district. There were 323 males and 285 females. Of the population, 67% were white, 15% black or African-American, and 18% represented other ethnicities.

The sample of this study was limited to 343 students from eight elementary schools in the Upper East Tennessee school district who were in third grade in 2017-2018. There were 189 males and 154 females. The sample was comprised of 70% of students who were white, 13% who were black or African-American, and 17% who were other ethnicities. The sample of students was selected because they were given the TCAP ELA third-grade assessment in 2017-

2018 and were also administered the DIBELS Next and TRC assessments as kindergarten students in 2014-2015 and first-grade students in 2015-2016. The demographic profile for the participants is in Table 1.

The names of third-grade students originated from the school district's student information service, PowerSchool. The students' scores on kindergarten and first-grade DIBELS Next and TRC scores were paired with their score on the TCAP ELA third-grade assessment. Transient students who had incomplete data were excluded from the study. In addition, students who participated in special education services or were given accommodations on the third-grade TCAP ELA were also omitted.

Table 1

Demographic Profile for the Participants

Schools	*(N)*	*(n)*	Males	Females
School 1	67	36	22	14
School 2	97	50	32	18
School 3	107	71	40	31
School 4	76	42	21	21
School 5	40	19	11	8
School 6	72	44	22	22
School 7	79	43	21	22
School 8	70	38	20	18

Instrumentation

DIBELS Next

Description

The data used in this study were collected using DIBELS Next, TRC, and TCAP ELA assessments. DIBELS Next assessment measured early literacy skills, such as phonemic awareness, phonics, fluency, and comprehension, of students in kindergarten through sixth grade. Six measures were included that specify the critical skills that every student must learn and understand in order to be a proficient reader. The one-minute measures consisted of First Sound Fluency (FSF), Letter Naming Fluency (LNF), Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and DIBELS Oral Reading Fluency (DORF). The measures were designed to identify students who are struggling with basic early literacy skills and to provide specific support for these students.

Administration

The DIBELS Next assessment was administered three times yearly with students using the grade-level measures. The initial screening occurred in the fall and identifies students who may be considered at risk for reading difficulty and end-of-year reading goals. DIBELS Next may also be used for progress monitoring students who are at risk on the benchmark assessment (Dynamic Measurement Group, 2011a). The school district in the study used DIBELS Next as its universal screener for kindergarten and first-grade students as a part of an assessment suite called mCLASS: Reading 3D provided by Amplify Education. Teachers and RtI interventionists administered DIBELS Next one-on-one to students using iPads in August, December, and May and stored the data on the mCLASS website.

Scoring

The DIBELS Next composite score was a grouping of multiple scores from each of the measures and gives the best overall estimate of a student's early reading skills and proficiency. If students are at or above the benchmark goal on the DIBELS Next composite score, they are more than likely to reach later important milestones in reading. However, if a student's composite score is close to the benchmark goal, he/she may need additional support in early reading skills as evidenced by the benchmark score on an individual measure. The levels for the DIBELS Next measures included Above Benchmark or At Benchmark (80-90% of achieving early literacy goals and likely to only need core support), Below Benchmark (40-60% of achieving early literacy goals and likely to need strategic support), and Well Below Benchmark (10-20% of achieving early literacy goals and likely to need intensive support) based on specific cut scores for each measure (Dynamic Measurement Group, Inc., 2011a). The DIBELS Next cut points for end-of-year kindergarten are included in Table 2. Additionally, the first-grade composite score, DORF accuracy score, and DORF retell score are included in Table 3.

Table 2

DIBELS Next Cut Points for Grade K End of Year Composite Score

Benchmark Status	Composite Score
Well Below Benchmark	$\leq$ 88
Below Benchmark	89-118
Benchmark	119-151
Above Benchmark	$\geq$ 152

Table 3

DIBELS Next Cut Points for First Grade End of Year

Benchmark Status	Composite Score	DORF Accuracy	DORF Fluency	DORF Retell
Well Below Benchmark	≤ 88	≤ 81%	0-31	---
Below Benchmark	89-118	82%-89%	32-46	0-14
Benchmark	119-151	90%-96%	47-66	15-16
Above Benchmark	≥ 152	≥ 97%	67+	≥ 17

Reliability and Validity

Numerous studies have proven the reliability and validity of DIBELS Next. Of these studies, many linked the formative value of DIBELS oral reading fluency to student performance on high-stakes assessments. Buck and Torgesen (2003) studied oral reading fluency scores and student performance on the Florida Comprehensive Assessment Test-Sunshine State Standards (FCAT-SSS). A significant correlation occurred between the assessments (r=.70, p<.001). Likewise, Shaw and Shaw (2002) studied the predictability of DIBELS Next ORF to the third grade Colorado State Assessment Program (CSAP). DIBELS was a strong predictor of the CSAP score with correlation coefficients ranging from .73 on the fall and winter assessments to .80 on the spring assessment. Other studies involving oral reading fluency and state assessments have demonstrated similar results. Since oral reading fluency has been shown to predict student achievement on state assessments, the validity and reliability of such measures are justified (Roehrig, et al., 2008). Additionally, experts at the Tennessee Department of Education regularly studied the DIBELS Next and included it as part of an approved list of RtI measures for school districts in Tennessee.

TRC (Text Reading and Comprehension)

Description

TRC was a standardized, formative reading assessment aligned to Common Core State Standards (CCSS) and published by Amplify Education, a digital education company that includes DIBELS Next as a part of its assessment suite called mCLASS: Reading 3D (Gushta, et al., n.d.; Snow, et al., 2018; W.K. Kellogg Foundation, n.d.). The TRC assisted educators in monitoring reading behaviors, observing reading development, and identifying reading difficulties in students. The assessment served as a component of the district's Response to Intervention (RtI)'s universal screener. It utilized 76 fiction and non-fiction texts published by Harcourt Publishing within the Atlas book set. These books include print concepts, reading behaviors, and corresponded to levels A to Z of a gradient of text difficulty (Fountas & Pinnell, 1996).

Administration

TRC was administered individually three times yearly (August, December, and May) to all kindergarten and first-grade students in the school district. Teachers and RtI interventionists administered TRC one-on-one to students using iPads and stored the data on the mCLASS website. Administrators use a running record to analyze a student's ability to read and comprehend grade-level texts. Students read a text aloud while the teacher records and codes the student's oral reading miscues and behaviors, which determines the student's reading accuracy. The comprehension portion of the assessment assisted teachers in determining a student's understanding of a text by asking oral retell, oral open-ended (implicit and explicit) questions, and written comprehension questions. The optional written comprehension questions are not utilized by the school district in the study because of the length of administration time.

Scoring

Data acquired from the TRC assisted educators in determining students' reading level and better-matched students to texts. The overall instructional reading level of the student was determined by the accuracy rate and comprehension ability of the student when and after reading the texts. Research shows students should read with at least a 90-94% rate of accuracy in order to comprehend text, which means students can read these texts independently with some help from a teacher or peer. This is called a student's instructional reading level. Texts at this level appropriately challenge students' reading abilities without impacting their comprehension. The independent level of a text is easier for students and gives them an opportunity to build fluency and confidence. Independent reading levels are texts students can read with 95% or higher accuracy and above 80% comprehension. In addition, the frustrational level of a text is one that is too challenging for students to read with appropriate accuracy and comprehension. These are texts that students read with less than 90% accuracy and below 80% comprehension. Students need extra support in reading these texts in a shared reading atmosphere (Clay, 2002; Fountas & Pinnell, 1999; Fuchs, Fuchs, & Deno, 1982).

The TRC provided an instructional reading level based on Fountas and Pinnell's leveling system (A through Z) which included 93% or more accuracy rate and 80% or more oral comprehension. Additionally, TRC utilized cut points, which were developed by eleven leveled-reading experts representative of different educational backgrounds and geographical locations for various performance levels (far below proficient, below proficient, proficient, and above proficient). The CCSS for ELA steered the specific proficiency levels for each assessment window (beginning of year, middle of year, end of year) and for each grade kindergarten through sixth grade. These were developed based on the Item Descriptor Matching Method, which was

appropriate for assessments that produce results by categories. Far below proficient identifies students who did not demonstrate any of the reading behaviors, which were contained in the CCSS ELA. Below proficient students showed some or few of the expected behaviors but not all. Proficient students exhibited minimal CCSS ELA expectations. Students who showed reading behaviors that exceeded expectations of CCSS ELA were above proficient. The literacy experts ensured the cut points for the beginning and the middle of year allowed students to be ready to meet the end of year expectations (Gushta, et al., n.d.).The cut points for kindergarten and first-grade end of year TRC were included in Table 4 and Table 5.

The nominal cut points for TRC were converted to ordinal values to facilitate analysis based on research by Snow, et al., (2018). Their research studied the difference in reading levels of 196 North Carolina students (grades 1-3) between TRC (a state-mandated assessment) and an Informal Reading Inventory (IRI). Table 6 shows how the ordinal levels correspond to the reading levels of Fountas and Pinnell (Print Concepts through Z) and to traditional grade levels (Pre-Primer to fifth grade) (Snow, 2014). Print Concepts (PC) and Reading Behaviors (RB) and levels X, Y, and Z were not included in their study, but the researcher included PC and RB in the emergent category and levels X, Y, and Z in the fifth grade category as this is where they should naturally fall being on either extreme of this scale.

Table 4

TRC Cut Points for Grade K End of Year

Performance Level	Book Level
Far Below Proficient	≤ B
Below Proficient	C
Proficient	D
Above Proficient	≥ E

Table 5

TRC Cut Points for First-Grade End of Year

Performance Level	Book Level
Far Below Proficient	≤ G
Below Proficient	H-I
Proficient	J-K
Above Proficient	≥ L

Table 6

TRC Ordinal Value Levels

Level	Grade	Book Level
0	Emergent	PC, RB, A, B, C
1	Pre-Primer	D, E
2	Primer	F, G
3	Late-First	H, I, J
4	Second	K, L, M, N
5	Third	O, P, Q
6	Fourth	R, S, T, U
7	Fifth	V, W, X, Y, Z

The comprehension retell was scored on a rubric from 0 to 3 according to the number of details given by the student from the beginning, middle, and end of the text read by the student. The retell questions were included in early level texts (Levels A through E) that contained less content for high-quality, text dependent questions. Five oral comprehension questions were assigned to books at levels D through Z. The questions aligned to the CCSS for English Language Arts (ELA), and each book contained one or more questions aligned to the CCSS substrands. Bloom's Taxonomy and Webb's Depth of Knowledge were also used in the creation of questions to ensure the appropriate level of cognitive complexity (Gushta, et al., n.d.). A score of 80% on the comprehension questions equated to proficient while lower percentages equated to non-proficient. The school district in the study did not administer the optional written comprehension questions to students.

Reliability and Validity

TRC utilizes the Atlas book set which participated in field-testing to demonstrate the validity and reliability of TRC and the associated texts as an oral reading accuracy and comprehension assessment. Analysis of the texts, which included qualitative and quantitative, examined the text complexity across grade levels and was approved by a field study based on how the students performed on the texts and the "statistical estimates of text difficulty" (Gushta, et al., n.d., p. 22). Amplify Education was able to acquire qualitative feedback in order to enhance texts and comprehension questions from the student field test data (Gushta, et al., n.d.).

In addition, several research studies have investigated the reliability and validity of the TRC assessment. Amplify Education (2013b) found a strong correlation with the third grade NC EOG for the final instructional reading level for TRC (r=0.71, p<0.05). Bowles (2015) found the TRC significantly statistically predicted scale scores on the NC EOG at all grade levels in the study, especially with fifth-grade students who showed the strongest correlation with r=.616.

TCAP (Tennessee Comprehensive Assessment Program) TNReady ELA

Description

TCAP TNReady ELA (English Language Arts) was a criterion-referenced, standards-based assessment given to all students in grades 3-8. This state-wide timed assessment surpassed the requirements of ESSA and required students to take yearly assessments in mathematics, ELA, science, and social studies (Tennessee Department of Education, 2017; Tennessee Department of Education, 2018; Tennessee Department of Education, n.d.b.). The ELA portion of the TCAP assessed the Tennessee Academic State Standards, utilized literary and informational texts, and required students to closely read, analyze text, answer text-dependent

questions, respond to a text-based writing prompt, and demonstrate knowledge of the English language.

TCAP TNReady was created to measure critical thinking and higher expectations for the students of Tennessee (Tennessee Department of Education, 2017). Item types on TCAP TNReady ELA extend beyond multiple-choice items, and include evidence-based selected responses (ESBR), multiple select, writing prompts, and editing tasks (Tennessee Department of Education, n.d.e).

Administration

TCAP TNReady ELA was administered to students during a two-week assessment window in April-May each school year. The ELA assessments were given to students in four subparts, which totaled 48-84 items, and evaluated students' understanding of written expression, reading literature, reading informational text, listening to literary text, listening to informational text, foundational literacy, fluency, language, and conventions. The test totaled 216 minutes. Subpart 1 totaled 80 minutes, subpart 2 and 3 each totaled 43 minutes, and subpart 4 totaled 50 minutes. Testing times for TCAP TNReady assessments were released by the Tennessee Department of Education. Students that had accommodations for TCAP testing were excluded from this research study.

Scoring

On the TCAP TNReady ELA, there were 74-112 score points possible. TCAP TNReady ELA scale score was provided and correlated to these achievement performance levels: Level 4 – Mastered (mastered grade level); Level 3 – On Track (on grade level); Level 2 – Approaching (approaching grade level); and Level 1 – Below (below grade level). Cut points were utilized to

assign students to these different achievement levels. The cut points for the 2017-2018 TCAP TNReady ELA assessment for third grade are included in Table 7.

Table 7

TCAP TNReady ELA Cut Scores for Grade 3

Performance Level	Cut Score
Level 1: Below	200-321
Level 2: Approaching	322-358
Level 3: On Track	359-390
Level 4: Mastered	391-450

Reliability and Validity

The questions created for the TCAP TNReady experience a three-step evaluation process. During the first step, the Tennessee Department of Education develops expectations for item development aligned to the state standards. Teachers and the test vendor create test questions. Next, the questions are examined for content and bias by the Tennessee Department of Education and by a group of Tennessee educators, administrators, and supervisors who are represented by grade level or grade band. Educators are able to accept test items, revise, or reject items. Then, the Educational Testing Service (ETS) (2012) reviews and revises items based on that feedback. Each question is then field tested and reviewed for statistical validity. Once the validity is established, the testing vendor adds the questions to the operational assessment, and the entire test is reviewed for accuracy by the Tennessee Department of Education. Finally, the final assessment is administered in Tennessee schools and classrooms (ETS, 2012; Tennessee Department of Education, n.d.b).

Data Collection

Prior to beginning this research study, the researcher obtained permission from the Institutional Review Board (IRB) at Milligan College. Upon approval, the researcher acquired permission from the school district to conduct the quantitative study through a letter to the Superintendent, a meeting with the Supervisor of Accountability and School Improvement, and a written proposal to the school district. Once permission was granted, the sample for the study was selected by the researcher and the school district. After the sample was identified, a letter was sent by the researcher to the school district's Supervisor of Testing, Educator Evaluation, and Response to Intervention and to the Director of Accountability and School Improvement, requesting access to TCAP TNReady ELA scores and data. The Supervisor of Testing, Educator Evaluation, and Response to Intervention provided the single file of TCAP TNReady ELA scale scores without identifiable information by creating a unique identifier for each student. The researcher, by virtue of her position, has complete access to district DIBELS Next and TRC scores from the mCLASS website. Data were transferred into Microsoft Excel without identifiable information. The researcher grouped each student's data from 2014-2015, 2015-2016, and 2017-2018 in an Excel file. Transient students who had incomplete data were excluded from the study. Demographic data, involving gender, race, ethnicity, special education services, and testing accommodations, were accessed through the school district's PowerSchool database. Students who participated in special education services or were given accommodations on the third-grade TCAP TNReady ELA were omitted from the study.

Data Analysis

Data analyses were conducted using IBM Statistical Package for Social Sciences (SPSS) v25. Analyses were managed for each research question as follows:

1. In response to question one, regarding the relationship between kindergarten DIBELS Next composite scores and third-grade TCAP ELA scores, a correlation coefficient (Pearson *r*) was utilized to determine the strength and direction of the relationship between the two variables.
2. Multiple linear regression was used in question two to determine the relationship of the independent variables, first-grade DIBELS Next composite scores, and first-grade DIBELS ORF accuracy scores, on the dependent variable, third-grade TCAP ELA scores. The coefficient of determination will explain, by the predictor variables, the amount of variance and the ANOVA will indicate significance (Holcomb, 2017).
3. Multiple linear regression was also used in question three to establish the relationship of the independent variables, first-grade DIBELS Next ORF fluency scores and first-grade DIBELS ORF retell scores, on the dependent variable, third-grade TCAP ELA scores. The amount of variance, explained by the predictor variables, will be determined by the coefficient of determination. Significance will be indicated with ANOVA.
4. In response to question four regarding the relationship between kindergarten TRC level of proficiency and third-grade TCAP ELA scores, a correlation coefficient (Pearson *r*) was utilized to determine the strength and direction of the relationship between the two variables.
5. A correlation coefficient (Pearson *r*) was used for question five to determine the strength and direction of the relationship between first-grade TRC level of proficiency and third-grade TCAP ELA scores.

6. In response to question six, a multiple linear regression was utilized to determine which independent variable, kindergarten DIBELS Next composite scores or kindergarten Text Reading and Comprehension (TRC), was a strong predictor of future performance on third-grade TCAP ELA, the dependent variable.
7. Multiple linear regression was also used in question seven to determine which independent variable, first-grade DIBELS Next composite scores or first-grade Text Reading and Comprehension (TRC), predicts future performance on the dependent variable, third-grade TCAP ELA.
8. Question 8 was analyzed using multiple linear regression with zero-order correlations to establish if kindergarten DIBELS Next composite scores, first-grade DIBELS Next composite scores, kindergarten Text Reading and Comprehension (TRC) level of proficiency, or first-grade Text Reading and Comprehension (TRC) level of proficiency was the best predictor of future performance on third-grade TCAP ELA.

All data were analyzed at the .001 significance level. The analysis results for each question are included in chapter 4.

Chapter Summary

This chapter contains the methodology used in this non-experimental quantitative research study. After a brief introduction, research questions, including null hypotheses, and the population and sample were presented. Additionally, the instrumentations used in the research study and the processes for data collection and data analysis were described.

CHAPTER 4

Data Analysis and Findings

The purpose of this quantitative study was to examine the relationship between kindergarten and first-grade students' DIBELS Next and TRC scores and third-grade students' TCAP ELA scores. The second purpose of this study was to determine the predictive validity of kindergarten and first-grade DIBELS Next and TRC scores on end of third-grade TCAP ELA performance. Participants in this study included third-grade students in 2017-2018 from eight elementary schools in the Upper East Tennessee school district. The students were given the TCAP ELA third-grade assessment in 2017-2018 and the DIBELS Next and TRC assessments as kindergarten students in 2014-2015 and first-grade students in 2015-2016. Transient students who had incomplete data and those who had participated in RtI services were excluded from the study. This chapter details the findings resulting from the data analysis of the six research questions used in this study.

Demographic Data

The population of this study was 608 third-grade students in 2017-2018 from eight elementary schools in the Upper East Tennessee school district. Of this population, there were 323 males and 285 females. Within the population, 67% were white, 15% black or African-American, and 18% represented other ethnicities.

The sample included 343 students who were third graders in 2017-2018 from eight elementary schools in the Upper East Tennessee school district. Of this sample, 189 were males and 154 were females. The sample contained 70% of students who were white, 13% who were black or African-American, and 17% who were other ethnicities. The sample of students was selected because they were also administered the DIBELS Next and TRC assessments as kindergarten students in 2014-2015 and first-grade students in 2015-2016 and were given the

TCAP ELA third grade assessment in 2017-2018. Each student's score on kindergarten and first-grade DIBELS Next and TRC scores were paired with their score on the TCAP ELA third grade assessment. Students who participated in special education services or were given accommodations on the third-grade TCAP ELA were omitted. Additionally, transient students who had incomplete data were excluded from the study.

Findings

Research Question 1

Research Question 1: Is there a relationship between kindergarten DIBELS Next composite scores and third-grade TCAP ELA scores?

$H_0$1: There is no relationship between kindergarten DIBELS Next composite scores and third-grade TCAP ELA scores.

A correlation coefficient (Pearson's *r*) was conducted to determine if there was a relationship between kindergarten DIBELS Next composite scores and third grade TCAP ELA scores. The results of the analysis indicate a direct, moderate, statistically significant relationship (r=.494, p=.001) between kindergarten DIBELS Next composite scores (M=154.73, sd=36.045) and third grade TCAP ELA scores (M=365.55, sd=27.829). In order to determine how much influence the kindergarten DIBELS Next composite scores had on third grade TCAP ELA scores, a coefficient of determination was calculated. The results indicate r^2=.244 which means 24% of the variance in third grade TCAP ELA scores can be explained by the kindergarten DIBELS Next composite scores. 76% of the variance in third grade TCAP ELA scores can be explained by other variables. The null hypothesis was rejected. Table 8 shows the correlation between the two groups.

Table 8

Pearson Correlation Coefficient and Means for Kindergarten DIBELS Next Composite Score and Third-Grade TCAP ELA

Source	Mean	*r*	r^2	*p*
DIBELS Next	154.73	.494	.244	.001*
TCAP ELA	365.55	.494	.244	.001*

Note: *indicates significance at *p*=.001.

Research Question 2

Research Question 2: What is the relationship between first-grade DIBELS Next composite scores and first-grade DIBELS Oral Reading Fluency accuracy scores on third-grade TCAP ELA scores?

$H_0$2: There is no relationship between first-grade DIBELS Next composite scores and first-grade DIBELS Oral Reading Fluency accuracy scores on third-grade TCAP ELA scores.

In response to research question two, a multiple linear regression was calculated to predict third-grade TCAP ELA scores based on first-grade DIBELS Next composite scores and first-grade DIBELS Next Oral Reading Fluency accuracy scores. A significant regression equation was found [F (2, 340)=88.943, p=.001] with an R^2 of .343. This suggests that 34% of the variance in third-grade TCAP ELA scores could be explained by the predictor variables. Results also suggest that 66% of the variance in third-grade TCAP ELA scores could be explained by other variables other than the predictor variables. To determine which predictor variables were significant predictors of third-grade TCAP ELA, Beta scores were examined. First-grade DIBELS Next composite score had a Beta score of .617, p=.001. First-grade DIBELS

Oral Reading Fluency accuracy score had a Beta score of -.041, *p*=.544. Therefore, first-grade DIBELS Next composite score was a significant predictor of third-grade TCAP ELA scores. First-grade DIBELS Next ORF accuracy score was not significant. The null hypothesis was rejected. The results are displayed in Table 9.

Table 9

Coefficients for Each Predictor Variable and the Dependent Variable

Variable	*B*	*Beta*	*t*	*Significance*
First-grade DIBELS Next composite score	.257	.617	9.092	.001*
First-grade DIBELS Next ORF accuracy score	-.171	-.041	-.608	.544

Note: *indicates significance at *p*=.001.

Research Question 3

Research Question 3: What is the relationship between first-grade DIBELS Next Oral Reading Fluency fluency scores and first-grade DIBELS Next Oral Reading Fluency retell scores on third-grade TCAP ELA scores?

$H_0$3: There is no relationship between first-grade DIBELS Next Oral Reading Fluency fluency scores and first-grade DIBELS Next Oral Reading Fluency retell scores on third-grade TCAP ELA scores.

A multiple linear regression was calculated to predict third-grade TCAP ELA scores based on first-grade DIBELS Next Oral Reading Fluency fluency scores and first-grade DIBELS Next Oral Reading Fluency retell scores. A significant regression equation was found [F (2, 315)=73.538, p=.001] with an R^2 of .318. This suggests that 32% of the variance in third-grade

TCAP ELA scores could be explained by the predictor variables. Results also suggest that 68% of the variance in third-grade TCAP ELA scores could be explained by other variables other than the predictor variables. To determine which predictor variables were significant predictors of third grade TCAP ELA, Beta scores were examined. First-grade DIBELS Next Oral Reading Fluency fluency scores had a Beta score of .551, p=.001. First-grade DIBELS Next Oral Reading Fluency retell scores had a Beta score of .062, p=.187. Therefore, first-grade DIBELS Next Oral Reading Fluency fluency scores were a significant predictor of third-grade TCAP ELA scores. DIBELS Next Oral Reading Fluency retell scores were not a significant predictor of third-grade TCAP ELA scores. The null hypothesis was rejected. The results are displayed in Table 10.

Table 10

Coefficients for Each Predictor Variable and the Dependent Variable

Variable	*B*	*Beta*	*t*	*Significance*
First-grade DIBELS Next ORF fluency score	.494	.5551	11.714	.001*
First-grade DIBELS Next ORF retell score	.093	.070	1.323	.187

Note: *indicates significance at p=.001.

Research Question 4

Research Question 4: Is there a relationship between kindergarten Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores?

$H_0$4: There is no relationship between kindergarten Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores.

A correlation coefficient (Pearson's *r*) was conducted to determine if there was a relationship between kindergarten TRC level of proficiency and third-grade TCAP ELA scores. The results of the analysis indicate a direct, moderate, statistically significant relationship (*r*=.468, *p*-=.001) between kindergarten TRC level of proficiency (*M*=1.11, *sd*=1.384) and third-grade TCAP ELA scores (*M*=365.55, *sd*=27.829). In order to determine how much influence the kindergarten TRC level of proficiency had on third-grade TCAP ELA scores, a coefficient of determination was calculated. The results indicate r^2=.219 which means 22% of the variance in third grade TCAP ELA scores can be explained by the kindergarten TRC level of proficiency. 78% of the variance in third-grade TCAP ELA scores can be explained by other variables. The null hypothesis was rejected. Table 11 shows the correlation between the two groups.

Table 11

Pearson Correlation Coefficient Summary for Kindergarten TRC Level of Proficiency and Third-Grade TCAP ELA

Source	Mean	*r*	r^2	*p*
TRC	1.11	.468	.219	.001*
TCAP	365.55	.468	.219	.001*

Note: *indicates significance at *p*=.001.

Research Question 5

Research Question 5: Is there a relationship between first-grade Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores?

$H_0$5: There is no relationship between first-grade Text Reading and Comprehension (TRC) level of proficiency and third-grade TCAP ELA scores.

A correlation coefficient (Pearson's *r*) was conducted to determine if there was a relationship between first-grade TRC level of proficiency and third-grade TCAP ELA scores. The results of the analysis indicate a direct, strong, statistically significant relationship (*r*=.580, *p*=.001) between first-grade TRC level of proficiency (*M*=4.37, *sd*=1.559) and third-grade TCAP ELA scores (*M*=365.55, *sd*=27.829). In order to determine how much influence the first-grade TRC level of proficiency had on third-grade TCAP ELA scores, a coefficient of determination was calculated. The results indicate r^2=.336 which means 34% of the variance in third-grade TCAP ELA scores can be explained by the first-grade TRC level of proficiency. 66% of the variance in third grade TCAP ELA scores can be explained by other variables. The null hypothesis was rejected. Table 12 shows the correlation between the two groups.

Table 12

Pearson Correlation Coefficient Summary for First-Grade TRC Level of Proficiency and Third-Grade TCAP ELA

Source	Mean	*r*	r^2	*p*
TRC	4.37	.580	.336	.001*
TCAP	365.55	.580	.336	.001*

Note: *indicates significance at *p*=.001.

Research Question 6

Research Question 6: Is kindergarten DIBELS Next composite score or kindergarten Text Reading and Comprehension (TRC) level of proficiency a better predictor of future performance on third-grade TCAP ELA?

$H_0$6: There is no relationship between kindergarten DIBELS Next composite score or kindergarten Text Reading and Comprehension (TRC) level of proficiency on third-grade TCAP ELA.

A multiple linear regression was calculated to predict third-grade TCAP ELA scores based on kindergarten DIBELS Next composite score and kindergarten TRC level of proficiency. A significant regression equation was found [F (2, 340)=69.374, p=.001] with an R^2 of .290. This suggests that 29% of the variance in third-grade TCAP ELA scores could be explained by the predictor variables. Results also suggest that 71% of the variance in third-grade TCAP ELA scores could be explained by other variables other than the predictor variables. To determine which predictor variables were significant predictors of third grade TCAP ELA, Beta scores were examined. Kindergarten DIBELS Next composite score had a Beta score of .334, p=.001. Kindergarten TRC level of proficiency had a Beta score of .267, p=.001. This indicates that although both variables were significant of third-grade TCAP ELA, kindergarten DIBELS Next composite score was a stronger predictor. The null hypothesis was rejected. The results are displayed in Table 13.

Table 13

Coefficients for Each Predictor Variable and the Dependent Variable

Variable	*B*	*Beta*	*t*	*Significance*
Kindergarten DIBELS Next composite scores	.258	.334	5.836	.001*
Kindergarten TRC level of proficiency	5.362	.267	4.660	.001*

Note: *indicates significance at p=.001.

Research Question 7

Research Question 7: Is first-grade DIBELS Next composite score or first-grade Text Reading and Comprehension (TRC) level of proficiency a better predictor of future performance on third-grade TCAP ELA?

$H_0$7: There is no relationship between first-grade DIBELS Next composite score or first-grade Text Reading and Comprehension (TRC) level of proficiency on third-grade TCAP ELA.

A multiple linear regression was calculated to predict third-grade TCAP ELA scores based on first-grade DIBELS composite score and first-grade TRC level of proficiency. A significant regression equation was found [F (2, 339)=108.295, p=.001] with an R^2 of .390. This suggests that 39% of the variance in third-grade TCAP ELA scores could be explained by the predictor variables. Results also suggest that 61% of the variance in third-grade TCAP ELA scores could be explained by other variables other than the predictor variables. To determine which predictor variables were significant predictors of third-grade TCAP ELA, Beta scores were examined. First-grade DIBELS composite score had a Beta score of .346, p=.001. First-grade TRC level of proficiency had a Beta score of .323, p=.001. This indicates that although both variables were significant of third-grade TCAP ELA, first-grade DIBELS Next composite score was a stronger predictor than first-grade TRC level of proficiency. The null hypothesis was rejected. The results are displayed in Table 14.

Table 14

Coefficients for Each Predictor Variable and the Dependent Variable

Variable	*B*	*Beta*	*t*	*Significance*
First-grade DIBELS Next Composite score	.145	.346	5.428	.001*
First-grade TRC	5.793	.323	5.066	.001*

Note: *indicates significance at p=.001.

Research Question 8

Research Question 8: Is kindergarten DIBELS Next composite score, first-grade DIBELS Next composite score, kindergarten Text Reading and Comprehension (TRC) level of proficiency, or first-grade Text Reading and Comprehension (TRC) level of proficiency the best predictor of future performance on third-grade TCAP ELA?

$H_0$8: There is no predictor of future performance by kindergarten DIBELS Next composite score, first-grade DIBELS Next composite score, kindergarten Text Reading and Comprehension (TRC) or first-grade Text Reading and Comprehension (TRC) on third-grade TCAP ELA.

A multiple linear regression was calculated to determine the best predictor of future performance on third-grade TCAP ELA scores between kindergarten DIBELS Next composite score, first-grade DIBELS Next composite score, kindergarten TRC level of proficiency, and first-grade TRC level of proficiency. A significant regression equation was found [F (4, 338)=58.945, p=.001] with an R^2 of .411. This suggests that 41% of the variance in third-grade TCAP ELA scores could be explained by the predictor variables. Results also suggest that 59% of the variance in third-grade TCAP ELA scores could be explained by variables other than the

predictor variables. To determine which predictor variables were significant predictors of third-grade TCAP ELA, Beta scores were examined. Kindergarten DIBELS Next composite score had a Beta score of .164, p=.004. First-grade DIBELS Next composite score had a Beta score of .266, p=.001. Kindergarten TRC level of proficiency had a Beta score of .037, p=.547. First-grade TRC level of proficiency had a Beta score of .268, p=.001. Due to the fact that many of these tests measure for similar constructs, tests for collinearity were conducted, and no variables had a VIF (variance inflation factor) of more than 5; thus, collinearity was not an issue with this regression. Therefore, it was apparent that both first-grade DIBELS Next composite score and first-grade TRC level of proficiency were equally strong predictors on third-grade TCAP ELA followed by kindergarten DIBELS Next composite score. Kindergarten TRC level of proficiency did not significantly influence third-grade TCAP ELA. The null hypothesis was rejected. The results are displayed in Table 15.

Table 15

Coefficients for Each Predictor Variable and the Dependent Variable

Variable	*B*	*Beta*	*t*	*Significance*
Kindergarten DIBELS Next composite scores	.129	.167	2.940	.004*
First-grade DIBELS Next composite scores	.112	.266	4.001	.001*
Kindergarten TRC level of proficiency	.735	.037	.603	.547
First-grade TRC level of proficiency	4.781	.268	3.873	.001*

Note: *indicates significance at p=.001.

Chapter Summary

In this chapter, kindergarten and first-grade DIBELS Next, kindergarten and first-grade TRC, and third-grade TCAP ELA data from a cohort of 343 students from an Upper East Tennessee school district were analyzed and presented. These students had been given the DIBELS Next and TRC assessments as kindergarten students in 2014-2015 and first-grade students in 2015-2016 and were also given the TCAP ELA third-grade assessment in 2017-2018. Eight research questions and eight null hypotheses were addressed. Results showed there was a significant relationship between kindergarten DIBELS Next composite score, kindergarten TRC level of proficiency, and first-grade TRC level of proficiency to third-grade TCAP ELA. First-grade DIBELS composite score and first-grade DIBELS Next Oral Reading Fluency fluency score all had a significant relationship on third-grade TCAP ELA, whereas there was not a significant relationship between first-grade DIBELS Next Oral Reading Fluency accuracy score and first-grade DIBELS Next Oral Reading Fluency retell score on third-grade TCAP ELA. Kindergarten DIBELS Next composite scores, first-grade DIBELS Next composite scores, and first-grade TRC level of proficiency were all predictors of third-grade TCAP ELA with first-grade DIBELS composite scores and first-grade TRC level of proficiency equally strong predictors of third-grade TCAP ELA. Kindergarten TRC level of proficiency was not a significant predictor of third-grade TCAP ELA.

CHAPTER 5

Summary of Findings, Discussions, Conclusions, and Recommendations

This chapter contains a summary of findings, discussions, and conclusions as well as recommendations for readers who may use the result of this research study to inform their school district's literacy assessment practices. The purpose of this correlational research study is to examine the relationship between kindergarten and first-grade students' DIBELS Next and TRC scores and third-grade students' TCAP ELA scores. The second purpose of this study is to determine the predictive validity of kindergarten and first-grade DIBELS Next and TRC scores on end of third-grade TCAP ELA performance.

Summary of Findings

The statistical analysis reported in this research study was based on eight research questions and eight null hypotheses. When each variable was measured independently of each other, results showed there was a significant relationship between kindergarten DIBELS Next composite scores, kindergarten TRC level of proficiency, and first-grade TRC level of proficiency to third-grade TCAP ELA. First-grade DIBELS composite scores and first-grade DIBELS Next Oral Reading Fluency fluency scores both had a significant relationship on third-grade TCAP ELA. However, there was not a significant relationship between first-grade DIBELS Next Oral Reading Fluency accuracy scores and first-grade DIBELS Next Oral Reading Fluency retell scores on third-grade TCAP ELA. When all variables were measured together, findings indicated that both first-grade DIBELS Next composite scores and first-grade TRC level of proficiency are equally strong predictors of third-grade TCAP ELA followed by kindergarten

DIBELS Next composite scores. Kindergarten TRC level of proficiency was not a significant predictor of future performance on third-grade TCAP ELA.

Discussion of Findings

As the era of high-stakes testing and accountability in schools persists, it is increasingly important that school districts locate and utilize the best early reading screeners available to guide reading instruction and intervention for students. Since the academic success of schools and school districts are judged solely based on the results from state assessments, it is equally essential that early reading screeners used by schools and school districts are predictive of future performance on high-stakes tests, such as state assessments. Results of this research study provide strong evidence for using DIBELS Next with kindergarten students and TRC with kindergarten and first-grade students for early reading screening. As a CBM assessment, the subtests within DIBELS Next are concise, skills-driven, fluency-based and individually administered. The subtests measure phonemic awareness, phonics, fluency, and comprehension. TRC, a running record assessment, provides students with authentic instructional texts to read, in addition to standards-based comprehension questions in a one-on-one setting and measures reading level, fluency, and comprehension. When utilized together, these assessments provide information regarding essential skills every student needs for reading proficiency based on the NRP report (National Reading Panel & National Institute of Child Health and Human Development, 2000). Findings of this research study are consistent with studies involving other early reading screeners which suggest phonemic awareness and phonics should be included in these assessments (Aspiranti, et al., 2017).

Although both DIBELS Next and TRC measure different reading skills, when used with first-grade students, they are equally strong predictors of third-grade TCAP ELA. Numerous

research studies show oral reading fluency is a strong predictor of reading comprehension and of success on state assessments (Barger, 2003; Buck & Torgesen, 2003; Deno, 2003; Roehrig, et al., 2008; Samuels & Riedel, 2007; Shaw & Shaw, 2002; Wilson, 2005). Therefore, many early reading screeners and CBMs, such as DIBELS Next, include oral reading fluency as subtests to predict reading comprehension. Critics of DIBELS Next find the subtests concerning because of the emphasis on fluency and speed when the ultimate goal of reading is comprehension; thus, the need for a comprehension component within an early reading screener is emphasized by numerous researchers (Rasinski, et al., 1996; Routman, 2003; Kuhn & Rasinski, 2011).

Although first-grade DIBELS Next Oral Reading Fluency fluency scores are significant to third-grade TCAP ELA, first-grade DIBELS Next Oral Reading Fluency accuracy scores are not significant. This finding can be explained with the difference between oral and silent reading. Students are assessed on their oral reading on the first-grade DIBELS Next Oral Reading Fluency assessment, whereas students are assessed on their silent reading on the third-grade TCAP ELA. When students read orally, especially beginning first-grade readers, they read more carefully to read all of the words accurately because they not only see the word, but they also hear it. These readers are typically more concerned with reading all of the words correctly than they are with understanding what they have read, especially when they are reading aloud to someone, as the case with the first-grade DIBELS Next Oral Reading Fluency. As readers transition to silent reading, which is typical of second and third graders, they are able to skim and still retain the meaning of the text. When students are reading the third-grade TCAP ELA, they are only listening to themselves read in their head because they are more independent readers, so they do not feel pressured to read every word accurately.

Even though DIBELS Next Oral Reading Fluency does contain a retell subtest designed to measure comprehension, it is not statistically significant to third-grade TCAP ELA in the research study. This is consistent with previous findings that question the reliability of the DIBELS Next Oral Reading Fluency retell scoring procedures (McKenna & Good, 2003; Pressley, et al., 2005; Roberts, et al., 2005). The DIBELS Next Oral Reading Fluency retell is administered orally based on oral readings of texts, whereas the comprehension section of the third-grade TCAP ELA contains multiple-choice and multiple-select responses on standards-aligned questions based on students' silent reading of the texts. An oral retell is an appropriate measure of comprehension for kindergarten and first-grade students as it supports their oral language development and follows a familiar story structure. Likewise, an appropriate measure for a more advanced reader is silent reading of comprehension questions. Students who have DIBELS Next Oral Reading Fluency retell scores that are below benchmark will tend to have a below grade level performance level on third-grade TCAP ELA. The students who struggle with oral retell may struggle with silent reading and answering comprehension questions. More research may also be needed to examine the effectiveness and administration of DIBELS Next Oral Reading Fluency retell (Riedel & Samuels, 2007).

The research study showed a moderate relationship between kindergarten TRC and third-grade TCAP ELA, but TRC was the most significant predictor of third-grade TCAP ELA in first-grade. When all four variables were analyzed, first-grade TRC level of proficiency was just as strong of a predictor as first-grade DIBELS Next scores. These findings are consistent with Ross (2004) who concluded that schools, which implement running records as a system of assessment, have students who exhibit higher achievement on reading and writing tests than schools that utilize other classroom assessment programs. Findings from other studies indicate running

records as a valid measure of reading proficiency because the assessment addresses accuracy and comprehension (Fuchs, et al., 1982). Running records immerse early readers in real, meaningful texts and are designed to give teachers powerful insight into reading processes by examining a reader's errors or miscues and then determining what cueing systems (meaning, visual, syntax) are being utilized by the reader. This information is helpful to RtI interventionists when working with struggling readers so they can teach students other strategies to use. Ideally, proficient readers utilize all cueing systems when reading a text.

The TRC also involves oral comprehension questions which are aligned to the ELA Common Core State Standards and involve various question-types from Bloom's Taxonomy and Webb's Depth of Knowledge (Gushta, et al., n.d.). Both TRC and TCAP require students to read whole texts and answer comprehension questions related to reading standards. They both provide students with picture support and the opportunity to find evidence in the text for comprehension questions. The assessments expect fluent reading and reading flow that require decoding and comprehension simultaneously. If students are "word callers," those who can read the words in the passage accurately and fluently but still may not be able to comprehend or retell the text, students' reading level or performance level is negatively affected (Cartwright, 2010; Riedel & Samuels, 2007). Early reading screeners for first grade students should include a running record component, such as the TRC since it is the most significant predictor for future performance on the third-grade TCAP ELA.

Limitations of the Study

Results of this study provide useful information regarding the validity of early reading screeners to state assessments. However, several limitations exist from the current research study. One possible limitation involves administration training and fidelity. Although detailed

training was provided to school personnel who were responsible for administering the DIBELS Next and TRC assessments, the researcher did not observe the administration of the testing. Also, there is an inconsistent procedure for testing administrations in each school regarding DIBELS Next and TRC universal screener assessments. The administrator may be a RtI teacher or a classroom teacher. It is also not guaranteed that the same testing administrator administrated the assessments to the same students each year.

In addition, the researcher did not observe fidelity of scoring. Detailed instructions for scoring were provided during the training and the researcher was available for troubleshooting as needed. However, testing administrators with more experience might be more fluent administrators, which would make a difference in scoring on timed assessments, such as DIBELS Next. In addition, although the TRC texts are designed to be a "cold read" for students, the researcher cannot guarantee that all students had not read the text before administration.

Additionally, a possible limitation would be the types of assessments used in the research study. The researcher studied DIBELS Next, which has a skills-based emphasis, and TRC and TCAP ELA which focus on standards. By utilizing early reading screeners that are skill-specific, students could be chosen for intervention that do not need it or those students who do need it might be overlooked. Furthermore, DIBELS Next and TCAP ELA are timed assessments, whereas TRC is not timed, which could be problematic and negatively affect the reading level of students.

Conclusions

This research study shows that although first-grade DIBELS Next and first-grade TRC measure different skills and strategies, they are strongly related to each other, especially in relationship to third-grade TCAP ELA. Students who have proficient scores on DIBELS Next

and TRC tend to demonstrate proficiency on the third-grade TCAP ELA. The findings suggest that if school districts are currently using DIBELS Next in first grade as a universal screener for RtI, this early reading screener should be continued. Since it consists of one-minute, timed subtests, it is more time-efficient even though it is administered individually to students and has a strong correlation to state assessments.

If school districts do not utilize a running record assessment in first grade, such as TRC, they should consider it. Even though both early reading screeners had equally strong relationships to third-grade TCAP ELA, TRC demonstrated the strongest relationship. TRC provides teachers with a students' instructional reading level and boasts a comprehension component consisting of standards-based, literal and non-literal questions. However, school districts should consider the time the TRC takes to administer individually. Although the Amplify mCLASS: Reading 3D manual promotes it as a 20-minute assessment per student, other research studies have not supported these claims. Some studies state that it takes as long as 60 minutes to individually administer, which can negatively affect students' reading achievement (Snow, et al., 2018).

The first-grade DIBELS Next and TRC level of proficiency have a stronger relationship to third-grade TCAP ELA than kindergarten DIBELS Next or kindergarten TRC level of proficiency. This is because first graders have an additional year of reading instruction and practice in order to become proficient readers compared to readers in kindergarten who are just beginning to read. If students are not proficient readers in kindergarten, they have first grade to continue to grow as readers. Goodman (2006) said that as students' reading ability increases, students will do better on all assessments that involve reading and writing. Therefore, first-grade students who are proficient with multiple early literacy skills, in addition to reading and

comprehension of connected text, have a greater chance at performing successfully on the third-grade TCAP ELA than kindergarten students. An early reading screener may be necessary for kindergarten students for the purposes of providing early warning signs of deficient skills, but there is a weak relationship to third-grade TCAP ELA if students are administered the DIBELS Next and no significance to third-grade TCAP ELA if kindergarten students are administered the TRC. Other kindergarten early reading screeners could be considered as a future study to see if relationships exist between those screeners and state assessments.

Recommendations for Practice

It is recommended that school districts and schools adopt "an empirically sound, highly effective assessment practice (i.e. CBM)" and ensure their assessments correlate with the summative state tests (McGlinchey & Hixson, 2004, p. 202). The results of this study support the use of curriculum-based measurements in kindergarten and first-grade as early reading screeners that are predictive of state assessments. Additionally, a running record assessment, such as TRC, should be utilized as well since they are powerful tools in predicting student scores on state assessments. Schools, school administrators, and teachers should regularly evaluate and progress monitor the reading progress of students to provide them with the appropriate reading instruction in Tier I and in RtI. In addition, school districts, school administrators, and teachers should use DIBELS Next and TRC appropriately for assessing students and collecting data to guide instructional decisions. DIBELS Next should not be utilized to promote an over-emphasis on speed reading in the classroom. In the same way, school districts, school administrators, and teachers should also assess second-grade students with early reading screeners, such as DIBELS Next and TRC. Second grade assessments and third grade assessments could have an even stronger relationship and predictability given the shorter length of time between the early reading

screener and the administration of the state assessment beginning in third-grade (Riedel & Samuels, 2007).

Teachers should regularly meet in professional learning communities to discuss data and next steps for reading instruction (Clarke, 2009). This ensures students are receiving the targeted instruction they need. Also, literacy materials and programs purchased by school districts and utilized in classrooms should reflect the five big ideas of reading. This study highlights the importance of high-quality, balanced reading instruction in kindergarten and first grade that involves phonemic awareness, phonics, comprehension, vocabulary, and fluency.

Recommendations for Further Study

The results of this study indicate that first-grade DIBELS Next and first-grade TRC were equally significant predictors of third-grade TCAP ELA. A replication of this study should be conducted in similar schools and school districts since "the widespread use of DIBELS for measuring progress and guiding instructional decisions makes it imperative for researchers to continue to examine the validity of the instrument" (Riedel & Samuels, 2007, p. 549). Also, an expansion of this study which would include subgroups, such as gender, ethnicity, English language proficiency, low-income status, and RtI students, would be beneficial. A similar study that includes DIBELS Next and TRC benchmark status categories and TCAP ELA performance levels could also yield interesting findings. Additionally, a comparable study is recommended using other curriculum-based measurement assessments that are presently utilized as universal screeners throughout other school districts, such as AIMSWeb and EasyCBM. Furthermore, a similar research study involving TRC should be conducted as there is a lack of research on this assessment in peer-reviewed journals. An expansion of this study should include a focus on

comprehension, including student responses to literal and non-literal comprehension questions, from TRC and state assessments.

References

Abbott, M., & Willis, H. (2012). Improving the upside-down Response-to-Intervention triangle with a systematic, effective, elementary school reading team. *Preventing School Failure, 56*(1), 37-46. doi:10.1080/1045988X.2011.555793

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Afflerbach, P., Kim, J. Y., Crassas, M. E., & Cho, B. Y. (2011). Best practices in literacy assessment. In L. M. Morrow and L. B. Gambrell (Eds.), *Best Practices in Literacy Instruction* (pp. 319-340). New York, NY: The Guilford Press.

Alcocer, P. (n.d.). History of standardized testing in the United States. Retrieved from https://www.nea.org/home/66139.htm

Allington, R. L. (1980). Teacher interruption behaviors during primary grade oral reading. *Journal of Educational Psychology, 72*(3), 371-77. doi:10.1037/0022-0663.72.3.371

Allington, R. L. (2006). Fluency: Still waiting after all these years. In S. J. Samuels and A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (94-105). Newark, DE: International Reading Association.

Allington, R., & Pearson, P. D. (2011, September). Conversation currents: The casualties of policy on early literacy development. *Language Arts, 89*(1), 70-74.

Almasi, J. F., & Hart, S. J. (2011). Best practices in comprehension instruction. In L .M. Morrow and L. B. Gambrell (Eds.), *Best Practices in Literacy Instruction* (319-340). New York, NY: The Guilford Press.

Alvermann, D. (1986). Becoming a nation of readers: The report of the commission on reading—a critical review. *Georgia Journal of Reading, 11*(2), 24-27.

Amplify Education, Inc. (2013a). *mCLASS: Reading 3D text reading and comprehension: Using*

the Common Core State Standards for English language arts to revised performance standards. Retrieved from https://ode.state.or.us/wma/teachlearn/testing/resources/mclass_reading_3d_blueprints.pdf

Amplify Education, Inc. (2013b). Validity evidence for mCLASS: Reading 3D and student performance on the 2012-2013 North Carolina end of grade reading comprehension test. Retrieved from www.amplify.com

Anderson, R. C. (2013). Role of the reader's schema in comprehension, learning, and memory. In D. E. Alvermann, N. J. Unrau, & R. Ruddell (Eds.). *Theoretical Models and Processes of Reading* (6th ed) (pp. 469-482). Newark, DE: IRA.

Applegate, M. D., Applegate, A. J., & Modla, V. (2009). 'She's my best reader; she just can't comprehend': Studying the relationship between fluency and comprehension. *The Reading Teacher, 62*(6), 512-521. doi: 10.1598/RT.62.6.5

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

Aspiranti, K., Hilton-Prillhart, A., Bell, S. M,. & McCallum, S. (2017). Kindergarten-Monitoring Instructional Responsiveness: Reading (KMIR:R): Examination of an authentic curriculum-based measure of beginning reading skills. *Research and Practice in the Schools, 5*(1), 1-10.

Badian, N. A. (2001). Phonological and orthographic processing: Their roles in reading prediction. *Annals of Dyslexia, 51*, 179-202.

Barger, J. (2003). *Comparing the DIBELS Oral Reading Fluency indicator and the North Carolina end of grade reading assessment* (Technical Report). Asheville, NC: North

Carolina Teacher Academy.

Barnett, D. W., Daly, E. J., Jones, K. M., & Lentz, F. E. (2004). Response to intervention: Empirically-based special service decisions from single-case designs of increasing and decreasing intensity. *The Journal of Special Education, 38*, 66-79. doi: 10.1177/00224669040380020101

Berkeley, S., Bender, W., Gregg-Peaster, L., & Saunders, L. (2009). Implementation of Response to Intervention: A snapshot of progress. *Journal of Learning Disabilities, 42*(1), 85-95. doi: 10.1177/0022219408326214

Blachman, B. A., Tangel, D. M., Ball, E., Black, R., & McGraw, C. K. (1999). Developing phonemic awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing: An Interdisciplinary Journal, 11*, 239-273. doi: 10.1023/A:1008050403932

Blevins, W. (2006). *Phonics from A to Z: A practical guide* (2nd ed.). New York: Scholastic.

Bowles, A. S. (2015). Does mCLASS Reading 3D predict student reading proficiency on high-stakes assessments? *Journal of Organizational & Educational Leadership, 1*(1), 1-27.

Bryne, B. Freebody, P., & Gates, A. (1992). Longitudinal data on the relations of word-reading strategies to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly, 27*, 141-151. doi: 10.2307/747683

Buck, J., & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test* (FCRR Technical Report #1). Tallahassee, FL: Florida Center for Reading Research.

Cartwright, K. B. (2010). *Word callers: Small-group and one-to-one interventions for children who "read" but don't comprehend*. Portsmouth, NH: Heinemann.

Chall, J. S. (1966). *Learning to read: The great debate* (3rd ed.). New York: McGraw-Hill.

Clarke, S. (2009, January/February). Using curriculum-based measurement to improve achievement. *Principal,* 30-33. doi: 10.1002/pits.20113

Clay, M. M. (2002). *An observation survey of early literacy achievement*. (2nd ed.). Portsmouth, N.H.: Heinemann.

Clay, M. M. (2005). *An observation survey of early literacy achievement*. (3rd ed.). Portsmouth, N.H.: Heinemann.

Cook, R. G. (2003). The utility of DIBELS as a curriculum-based measurement in relation to reading proficiency on high-stakes tests (Unpublished master's thesis, Marshall University Graduate College). Retrieved from https://mds.marshall.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1350&context=etd

Cowen, J. (2003). *A balanced approach to beginning reading instruction: A synthesis of six major U.S. research studies.* Newark, NJ: International Reading Association.

Crowell, C. G. (2015). Miscue analysis v. DIBELS: A tale of resistance. *Talking Points*, *26*(2), 2-9.

Cunningham, J., Cunningham, P., Hoffman, J., & Yopp, H. (1998). *Phonemic awareness and the teaching of reading: A position statement from the board of directors of the International Reading Association.* Newark, DE: International Reading Association. Retrieved from https://www.literacyworldwide.org/docs/default-source/where-we-stand/phonemic-awareness-position-statement.pdf?sfvrsn=944ea18e_6

Cunningham, P. (2004) *Phonics they use: Words for reading and writing* (4th ed.). New York: Allyn & Bacon.

Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Oranje, A. (2005). Fourth-grade students reading aloud: NAEP 2002 special study of oral reading. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006469

Deno, S. L. (1985). Curriculum-based measurement: The emerging alterative. *Exceptional Children, 52,* 219-232. doi: 10.11770014402985052003O3

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*(3), 184-192. doi: 0.1177/00224669030370030801

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

DeVries, B. A. (2011). *Literacy assessment and intervention for classroom teachers* (3rd ed.). Scottsdale, AZ: Holcomb Hathaway Publishers.

Dworkin, A. G. (2005). The No Child Left Behind Act: Accountability, high-stakes testing, and roles for sociologists. *Sociology of Education, 78*(2), 170-174. doi: 10.1177/003804070507800205

Duffy, M., Giordano, V. A., Farrell, J. B., Paneque, O. M., & Crump, G. B. (2008). No Child Left Behind: Values and research issues in high-stakes assessments. *Counseling and Values, 53,* 53-66. doi: 10.1002/j.2161-007X.2009.tb00113.x

Duke, N. & Carlisle, J. (2011). The development of comprehension. *Handbook of Reading Research, 4,* 199-228.

Dynamic Measurement Group, Inc. (2011a). *DIBELS Next assessment manual*. Retrieved from http://dibels.org

Dynamic Measurement Group, Inc. (2011b). Why is retell a required part of DORF in DIBELS Next? Retrieved from https://acadiencelearning.org/papers/WhyRetell_DIBELSNext.pdf

Dynamic Measurement Group, Inc. (2016, September 12). DIBELS Next benchmark goals and composite score. Retrieved from https://acadiencelearning.org/papers/DIBELSNext

BenchmarkGoals.pdf

Educational Testing Service (ETS). (2012). K-12 State assessment programs: Tennessee Comprehensive Assessment Program (TCAP)—TNReady and TCAP alternative assessments. Retrieved from http://www.ets.org/k12/programs/custom_assessments

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*(2), 250-287.

Farstrup, A. (2002). There is more to effective reading instruction than research. In A. Farstru & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed.) (pp. 1-7). Newark, DE: International Reading Association.

Foorman, B. R., Fletcher, J. M., Francis, D. J., Carlson, C. D, Chen, D., Mouzaki, A., & Taylor, R. H. (1998). *Texas Primary Reading Inventory* (1998 Edition). Houston, TX: Center for Academic and Reading Skills and University of Houston.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Winikates, D., Mehta, P., Schatschneider, C. (1997). Early interventions for children with reading problems. *Scientific Studies of Reading, 1*, 255-275. doi: 10.1207/s1532799xssr0103_5

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (1999). *Matching books to readers: Using leveled books in guided reading, K-3*. Portsmouth, NH: Heinemann.

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

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2012). Smart RtI: A next-generation approach to multilevel prevention. *Exceptional Children, 78*(3), 263-279. doi:

10.1177/001440291207800301

Fuchs, L. S. (2016). Curriculum-Based Measurement as the emerging alternative: Three decades later. *Learning Disabilities Research & Practice*, *32*(1), 5-7. doi: 10.1111/ldrp.12127

Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness of intervention. *Exceptional Children, 39*(5), 14-20. doi: 10.1177%2F004005990703900503

Fuchs, L. S., Fuchs, D., & Deno, S. (1982). Reliability and validity of curriculum-based informal reading inventories. *Reading Research Quarterly, 18*, 6-26. doi: 10.2307/747536

Good, R. H., Simmons, D. C., & Kame'enui, E. J., (2001). The importance and 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

Goodman, K. S. (Ed.). (2006). *The truth about DIBELS: What it is what it does.* Portsmouth, NH: Heinemann.

Griffith, L. W., & Rasinski, T. V. (2004). A focus on fluency: How one teacher incorporated fluency with her reading curriculum. *The Reading Teacher, 58*(2), 126-137. doi: 10.1598/RT.58.2.1

Gushta, M., Parisi, D., Richards, K., & York, A. (n.d.). mCLASS Reading 3D: Amplify Atlas book set development (White Paper). Retrieved from https://www.mclasshome.com/support_center/mclass_reading3d_book_set_development.pdf

Hagan-Burke, K., Burke, M., & Crowder, C. (2006). The convergent validity of the Dynamic Indicators of Basic Early Literacy Skills and the test of word reading efficiency for the

beginning of first grade. *Assessment for Effective Intervention, 31*(4), 1-15. doi: 10.1177/073724770603100401

Haladyna, T., Haas, N., & Allison, J. (1998). Continuing tensions in standardized testing. *Childhood Education, 74*(5), 262-273. doi: 10.1080/00094056.1998.10521950

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review, 32*, 228-240.

Harn, B. A., Stoolmiller, M., & Chard, D. J. (2008). Measuring the dimensions of alphabetic principle on the reading development of first graders. *Journal of Learning Disabilities, 41*(2), 143-157. doi: 10.1177/0022219407313585

Harris, T., & Hodges, R. (1995). *The literacy dictionary*. Newark, DE: International Reading Association.

Holcomb, Z. C. (2017). SPSS basics: Techniques for a first course in statistics. Glendale, CA: Pyrczak Publishing.

Huddleston, A. P., & Rockwell, E. C. (2015). Assessment for the masses: A historical critique of high-stakes testing in reading. *Texas Journal of Literacy Education, 3*(1), 38-49.

International Reading Association, & The National Association for the Education of Young Children. (1998). Learning to read: Developmentally appropriate practices for young children. *Young Children, 53*(4), 30-46.

Invernizzi, M., Justice, L., Landrum, T. J., & Booker, K. (2004). Early literacy screening in kindergarten: Widespread implementation in Virginia. *Journal of Literacy Research, 36*(4), 479-500. doi: 10.1207/s15548430jlr3604_3

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2007). *Handbook of Response to Intervention: The science and practice of assessment and intervention*. New

York, NY: Springer.

Johnson, P. (2006). *One child at a time: Making the most of your time with struggling readers, K-6.* Portland, ME: Stenhouse.

Johnson, P., & Keier, K. (2010). *Catching readers before they fall: Supporting readers who struggle, K-4.* Portland, ME: Stenhouse.

Kaminski, R., & Cummings, K. D. (2007). DIBELS: Myths and facts. Retrieved from https://acadiencelearning.org/papers/Myths_0208.pdf

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215-227.

Kimmel, M. K. (2008). *The successes and challenges of response to intervention: A case study of the impact of RTI implementation*. (Doctoral dissertation, University of Southern California, Los Angeles). Retrieved from http://digitallibrary.usc.edu/cdm/ref/collection/p15799coll127/id/46254.

Klein, A. (2015, Apr. 10). No Child Left Behind: An overview. *Education Week, (34)* 27. Retrieved from https://www.edweek.org/ew/section/multimedia/no-child-left-behind-overview-definition-summary.html

Kuhn, M. R., & Rasinski, T. R. (2011). Best practices in fluency instruction. In L .M. Morrow and L. B. Gambrell (Eds.), *Best Practices in Literacy Instruction* (pp. 276-294). New York, NY: The Guilford Press.

Kuhn, M. R., & Stahl, S. A. (2000). Fluency: A review of developmental remedial practices. (Ciera Report #2-008). Retrieved from http://www.ciera.org/library/reports/inquiry-2/2-008/2-008.pdf

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in

reading. *Cognitive Psychology, 6*, 293-323.

Lane, H. S., & Pullen, P. (2004). *Phonological awareness assessment and instruction: A sound beginning*. Boston: Allyn & Bacon.

Madaus, G.F. (1988). The influence of testing on the curriculum. In L.N. Tanner (Eds.), *Critical issues in curriculum: 87th yearbook of the National Society for the Study of Education* (pp. 83-121). Chicago, IL: University of Chicago Press.

Mansouri, D. (2018, June 4). High expectations, rigorous assessment, and closing the honesty gap in Tennessee. Retrieved from https://tnscore.org/high-expectations-rigorous-assessment-and-closing-the-honesty-gap-in-tennessee/

Manzo, K. K. (2005). National clout of DIBELS test draw scrutiny: Critics say reading tool's scope fails to justify its broad use. *Education Week, 25*(5), 1-4.

Marcotte, A. M., & Hintze, J. M. (2009). Incremental and predictive utility of formative assessment methods of reading comprehension. *Journal of School Psychology, 47*(5), 315-335. doi: 10.1016/j.jsp.2009.04.003

Marion, S. (2016, Feb. 19). Considerations for state leaders in the design of school accountability systems under the Every Student Succeeds Act. National Center for the Improvement of Educational Assessment. Retrieved from https://eric.ed.gov/?q=no+child+left+behind+and+essa&id=ED570461

Marks, D. (1989). Statewide achievement testing: A brief history. *Educational Research Quarterly, 13*(3), 36-43.

Martin, M. (2016). *School accountability systems and the Every Student Succeeds Act.* The Hunt Institute's re:VISION. Retrieved from https://files.eric.ed.gov/fulltext/ED569952.pdf

McGlinchey, M.T., & Hixson, M.D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*(2), 293-203.

McKenna, M. K., & Good, R. H. (2003). Assessing reading comprehension: The relation between DIBELS oral reading fluency, DIBELS retell fluency, and Oregon State Assessment Scores (Technical Report) Eugene, OR: University of Oregon.

McLaughlin, M. (2008). Reading comprehension: An evolution of theory, research, and practice. In M. J. Fresch (Eds.), *An essential history of current reading practices* (pp. 82-105). Newark, DE: International Reading Association.

Medina, J., & Riconscente, M. M. (2006). Accounting for quality. *The Journal of Education, 186*(3), 3-10.

Meisels, S. J. (1989). High-stakes testing in kindergarten. *Educational Leadership, 46*(7), 16-22.

Meisinger, E. B., Bradley, B. A., Schwanenflugel, P. J., Kuhn, M. R., & Morris, R. D. (2009). Myth and reality of the word caller: The relation between teacher nominations and prevalence among elementary school children. *School Psychology Quarterly, 24*(3), 147-159.

Mellard, D. F., & Johnson, E. (2008). *RTI: A practitioner's guide to implementing Response to Intervention*. Thousand Oaks, CA: Corwin Press. doi: 10.4135/9781483329772

Munger, K. A., LoFaro, S. A., Kawryga, E. A., Sovocool, E. A., & Medina, S. Y. (2014). Does the Dynamic Indicators of Basic Early Literacy Skills Next assessment take a "simple view" of reading? *Educational Assessment, 19*, 204-228.

National Reading Panel & National Institute of Child Health and Human Development (U.S.). (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, D.C.:

National Institute of Child Health and Human Development, National Institutes of Health. Retrieved from https://www1.nichd.nih.gov/publications/pubs/nrp/Pages/smallbook.aspx

Nelson, H. (2013). Testing more, teaching less: What America's obsession with student testing costs in money and lost instructional time. New York: American Federation of Teachers. Retrieved from https://www.aft.org/sites/default/files/news/testingmore2013.pdf

O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3,* 159-197. doi: 10.1207/s1532799xssr0302_4

Pearson, P. D. (2000). Reading in the twentieth century. Washington, DC: Office of Educational Research and Improvement (CIERA). Retrieved from https://files.eric.ed.gov/fulltext/ED479530.pdf

Pearson, P. D. (2010). American reading instruction since 1967. In Bean, R. M., Heisey, N., & Roller, C. M. (Eds.), *Preparing reading professionals* (pp. 7-38). Newark, DE: International Reading Association.

Pearson, P. D., and Stephens, D. (1994). Learning about literacy: A 30-year journey. In Ruddell, R. E., Ruddell, M. R., & Singer, H. (Eds.), *Theoretical models and processes of reading*. (4th ed.) (pp. 22-43). Newark, DE: International Reading Association.

Popham, W. J. (2001). Uses and misuses of standardized tests. *NASSP Bulletin, 85*(24), 24-31. doi: 10.1177/019263650108562204.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In Kamil, M. L., Mosenthal, P. B., Pearson, P. D., Moje, E. B., & Barr, R. (Eds.), Handbook of reading research (3rd ed.) (pp. 545-561). Mahway, NJ: Lawrence Erlbaum Associates.

Pressley, M., Hilden, K., & Shankland, R. (2005). An evaluation of end-grade-3 Dynamic

Indicators of Basic Early Literacy Skills (DIBELS): Speed reading without comprehension, predicting little (Technical Report). East Lansing, MI: Michigan State University, Literacy Achievement Research Center.

Rasinski, T. (1985). A study of factors involved in reader-text interactions that contribute to fluency in learning-disabled children. (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.

Rasinski, T. (2003). *The fluent reader*. New York, NY: Scholastic Professional Books.

Rasinski, T. V., Reutzel, D. R., Chard, D., & Linan-Thompson, S. (1996). Reading fluency In Kamil, M. L., Pearson, P. D., Moje, E. B., & Afflerbach, P. P. (Eds.), *Handbook of reading research* (2nd ed.). (pp. 287-319). New York, NY: Routledge.

Rathvon, N. (2004). *Early reading assessment: A practitioner's handbook.* New York, NY: The Guildford Press.

Readence, J. E., & Moore, D. W. (1983). Why questions? A historical perspective on standardized reading comprehension tests. *Journal of Reading, 26*(4), 306-313.

Riddle Buly, M. R., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis, 24*, 219-239. doi: 10.3102/01623737024003219

Riedel, B. W., & Samuels, S. J. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students (with commentary). *Reading Research Quarterly, 42*(4), 546-567. doi:10.1598/RRQ.42.4.5

Roberts, G., Good, R., & Corcoran, S. (2005). Story retell: A fluency-based indicator of reading comprehension. *School Psychology Quarterly, 20*, 304-317. doi: 10.1521/scpq.2005.20.3.304

Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS Oral Reading Fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology, 46*(3), 343-366. doi:

10.1016/j.jsp.2007.06.006

Ross, J. A. (2004). Effects of running records assessments on early literacy achievement: Results of a controlled experiment. *Journal of Educational Research, 97*(4), 186-194. doi: 10.3200/JOER.97.4.186-195

Rosenblatt, L. M. (1994). The transactional theory of reading and writing. In Ruddell, R. E., Ruddell, M. R., & Singer, H. (Eds.), *Theoretical models and processes of reading*. (4th ed.) (pp. 1057-1092). Newark, DE: International Reading Association.

Routman, R. (2003). *Reading essentials: The specifics you need to teach reading well.* Portsmouth, NH: Heinemann.

Rothman, R., & Marion, S. F. (2016). The next generation of state assessment and accountability. *Phi Delta Kappan, 97*(9), 34-37. doi: 10.1177%2F0031721716647016

Samuels, S. J. (2006, May). Introduction to fluency. Paper presented at the annual meeting of the International Reading Association, Chicago, IL.

Scanlon, D. M. (2011). Response to intervention as an assessment approach. In R. L. Allington and A. McGill-Franzen (Eds.), *Handbook of Reading Disability Research* (pp. 139-148). New York, NY: Taylor and Francis.

Shankweiler, D., Lundquist, E., Katz, L, Stuebing, K. K., Fletcher, J. M., Brady, S.,…Shaywitz, B. A. (1999). Comprehension and decoding: Patterns of associations in children with reading difficulties. *Scientific Studies of Reading, 3*, 69-94. doi: 10.1207/s1532799xssr0301_4

Shapiro, E. S. (2008). Best practices in setting progress monitoring goals for academic skill improvement. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 114-158). Bethesda, MD: National Association of School Psychologists.

Shaw, R., & Shaw, D. (2002). DIBELS Oral Reading Fluency-based indicators of third grade reading skills for Colorado State Assessment Program (CSAP). (Technical Report). Eugene, OR: University of Oregon.

Sheninger, E. C., & Murray, T. C. (2017*). Learning transformed: 8 keys to designing tomorrow's schools, today.* Alexandria, VA: ASCD.

Shepard, L. A. (2016). Testing and assessment for the good of education: Contributions for AERA Presidents, 1915-2015. *Educational Researcher, 45*(2), 112-121. doi: 10.3102%2F0013189X16639599

Shinn, M. R. (2010). Building a scientifically based data system for progress monitoring and universal screening across three tiers, including RtI using curriculum-based measurement. *Interventions for Achievement and Behavior Problems in a Three-Tier Model including RTI,* 259-292.

Smith, F. (1999). Why systematic phonics and phonemic awareness instruction constitute an educational hazard. *Language Arts, 77*(2), 150-155.

Snow, A. B. (2014). A comparative evaluation of instructional levels determined by the Text Reading and Comprehension (TRC) assessment and an informal reading inventory. (unpublished doctoral dissertation). Appalachian State University, Boone, NC.

Snow, A. B., Morris, D., & Perney, J. (2018). Evaluating the effectiveness of a state-mandated benchmark reading assessment: mCLASS Reading 3D (Text Reading and Comprehension). *Reading Psychology*, *39(4)*, 303-334. doi: 10.1080/02702711.2017.1422302

Snow, C., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, D.C.: National Academy Press.

Spear-Swerling, L. (2011). Patterns of reading disabilities across development. In R. L. Allington and A. McGill-Franzen (Eds.). *Handbook of Reading Disability Research* (pp. 149-161). New York: Taylor and Francis.

Squire, J. R. (1994). Research in reader response, naturally interdisciplinary. In Ruddell, R. E., Ruddell, M. R., & Singer, H. (Eds.), *Theoretical models and processes of reading*. (4th ed.) (pp. 637-667). Newark, DE: International Reading Association.

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.

Stahl, K. A. D. (2004). Proof, practice, and promise: Comprehension strategy instruction in primary grades. *The Reading Teacher*, *57,* 598–609.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 306-406. doi: 10.1177%2F0022057409189001-204

State Collaborative on Reforming Education (SCORE). (2015, September). Teaching, testing, and time: Educator voices on improving assessment in Tennessee. Retrieved from https://tnscore.org/wp-content/uploads/2018/09/Teaching-Testing-and- Time_ PolicyReport-2015.pdf

Strauss, V. (2014, June 25). Is this really how we should test reading development in kids? *The Washington Post.* Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2014/06/25/is-this-really-how-we-should-test-reading-development-in-kids/?noredirect=on&utm_term=.bdfa4b92e8fd

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

Tennessee Department of Education (2015). *Implementation guide: Response to instruction and*

intervention framework. Nashville, TN: Education, Tennessee Department of Education. Retrieved from https://tn.gov/content/dam/tn/education/special- education/rti/rti2_ implementation_guide.pdf

Tennessee Department of Education. (2017). Every Student Succeeds Act: Building on success in Tennessee, ESSA state plan. Retrieved from https://tsba.net/wp-content/uploads/2017/ 11/ESSA_state_plan.pdf

Tennessee Department of Education. (2018, April). First steps: A report on elementary grades reading in Tennessee. Retrieved from https://www.tn.gov/content/dam/tn/education/ reports/rpt_first_steps_reading_report.pdf

Tennessee Department of Education. (n.d.a). How TNReady is created for Tennessee students. Retrieved from https://www.tn.gov/content/dam/tn/education/documents/ How_TNReady_is_Created_FINAL.pdf

Tennessee Department of Education (n.d.b). Overview of testing in Tennessee. Retrieved from https://tn.gov/education/assessment/testing-overview.html

Tennessee Department of Education. (n.d.c.). Student assessment in Tennessee. Retrieved from https://www.tn.gov/education/assessment.html

Tennessee Department of Education (n.d.d). Tennessee's education transformation. Nashville, TN; Tennessee Department of Education. Retrieved from https://tn.gov/governor/priorities/tennessee-education-transformation.html

Tennessee Department of Education. (n.d.e). 2018-19 TNReady English language arts grades 3-4 assessment fact sheet. Retrieved from https://www.tn.gov/content/dam/tn/education/ testing/Grades_3-4_ELA_Fact_Sheet.pdf

Tierney, R. J., & Thome, C. (2006). Is DIBELS leading us down the wrong path? In Goodman,

K. S. (Eds.), *The truth about DIBELS: What it is what it does.* (pp. 50-59). Portsmouth, N.H.: Heinemann.

Torgesen, J.K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice, 15*, 55-64. doi: 10.1207/SLDRP1501_6

United States Department of Education. (2001). *No Child Left Behind (NCLB).* Retrieved from https://www2.ed.gov/policy/elsec/leg/esea02/index.html

United States Department of Education. (2015). *Every Student Succeeds Act (ESSA).* Retrieved from https://www.ed.gov/essa

VanDerHeyden, A. (2011). Technical adequacy of response to intervention decision. *Council for Exceptional Children, 77*(3), 335-350. doi: 10.1177/001440291107700305

Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly*, *35*(4), 554-566. doi: 10.1598/RRQ.35.4.7

Wilson, J. (2005). *The relationship of Dynamic Indicators of Basic Early Reading Skills (DIBELS) Oral Reading Fluency to performance on Arizona Instrument to Measure Standards (AIMS).* (Research Brief). Tempe, AZ: Assessment and Evaluation Department Tempe School District No. 3.

W. K. Kellogg Foundation. (n.d.). *Wireless Generation ushers in new wave of educational technology.* Battle Creek, MI: Author. Retrieved from https://www.wkkf.org/what-we-do/featured-work/wireless-generation

Yopp, H. K., & Yopp, R. H. (2000). Supporting phonemic awareness development in the classroom. *The Reading Teacher, 54*(2), 130-143. doi: 10.1598/RT.54.2.2

Zhao, H., & Von Secker, C. (2008, August). Evaluation of the criterion-related validity of

Montgomery County Public Schools assessment program in primary reading. (Research Brief). Retrieved from: https://montgomeryschoolsmd.org/departments/shared accountability/reports/2008/CRValidity_2008-9-10.pdf

APPENDICES

APPENDIX A

Letter to School District
Permission to Conduct Study

To:

FROM: JoDee W. Dotson

DATE: July 3, 2018

SUBJECT: District Permission to Conduct Study

I would like your permission to conduct a research study in _____Schools as part of my doctoral dissertation at Milligan College. The purpose of this correlational quantitative study is to explore the relationship between kindergarten and first grade students' DIBELS Next and TRC scores and third grade students' TCAP ELA scores. It is also to validate the integrity of DIBELS Next and TRC by establishing their predictive validity. I will be using student scores from 2014-2015, 2015-2016, and 2017-2018. Finding a relationship between these assessments could assist district personnel in making better decisions about universal screeners, placements of students, and instructional decisions.

There will not be any identifiable data used in the process of gathering student scores. Schools will also not be referenced in this study. _____ has graciously agreed to serve as liaison in the data collection process to assure the confidentiality of students and the school district.

I will be conducting this study under the supervision of Milligan College. I believe the results of the study will enable our school district to better serve our students, I appreciate your consideration of this important matter. Thank you!

Respectfully,

JoDee W. Dotson

APPENDIX B

IRB APPROVAL

Date: July 23, 2018

From: The Institutional Review Board (IRB) at Milligan College

RE: IRB Proposal for JoDee Dotson

Submission type: Initial Submission

Dear JoDee,

On behalf of the Milligan College Institutional Review Board (IRB), we are writing to inform you that your study *'The Relationship between Kindergarten and First Grade DIBELS Next and TRC Scores and Third Grade TCAP ELA Scores at a Selected School District in Upper East Tennessee'* has been approved as expedited. This approval also indicates that you have fulfilled the IRB requirements for Milligan College.

All research must be conducted in accordance with this approved submission, meaning that you will follow the research plan you have outlined here, use approved materials, and follow college policies.

Take special note of the following important aspects of your approval:

- Any changes made to your study require approval from the IRB Committee before they can be implemented as part of your study. Contact the IRB Committee at **IRB@milligan.edu** with your questions and/or proposed modifications.
- If there are any unanticipated problems or complaints from participants during your data collection, you must notify the Milligan College IRB Office within 24 hours of the data collection problem or complaint.

The Milligan College IRB Committee is pleased to congratulate you on the approval of your research proposal. Best wishes as you conduct your research! If you have any questions about your IRB Approval, please contact the IRB Office and copy your faculty advisor if appropriate on the communication.

Regards,

The IRB Committee