2019

An Evaluation of the Relationship Among North Carolina Reading Assessments

Sophia Latrell Crawford-Mapp

AN EVALUATION OF THE RELATIONSHIP AMONG NORTH CAROLINA READING ASSESSMENTS

By
Sophia Latrell Crawford-Mapp

A Dissertation Submitted to the
Gardner-Webb University School of Education
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

Gardner-Webb University
2019

Approval Page

This dissertation was submitted by Sophia Latrell Crawford-Mapp under the direction of the persons listed below. It was submitted to the Gardner-Webb University School of Education and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Gardner-Webb University.

Morgan Blanton, Ed.D. Committee Chair	Date
Sydney Brown, Ph.D. Committee Member	Date
Lane Wesson, Ed.D. Committee Member	Date
Prince Bull, Ph.D. Dean of the School of Education	Date

Acknowledgements

I would like to thank my family, friends, and colleagues who provided me with feedback and motivation to help me complete my dissertation. All kindness shown to me in this process is greatly appreciated. I also want to thank my kids, Carter and Ava, for being understanding, smiling, and fussing when mommy was unavailable. Finally, I want to thank God for allowing me to persevere and endure! Hebrews 10:36: "You need to persevere so that when you have done the will of God, you will receive what he has promised." Luke 6:40: "The student is not above the teacher, but everyone who is fully trained will be like their teacher." 1 John 5: 14-15: "This is the confidence we have in approaching God; that if we ask anything according to his will, he hears us. And if we know that he hears us- whatever we ask- we know that we have what we asked of him."

Abstract

AN EVALUATION OF THE RELATIONSHIP AMONG NORTH CAROLINA READING ASSESSMENTS. Crawford-Mapp, Sophia L Latrell, 2019: Dissertation, Gardner-Webb University.

The purpose of this study was to extend the Bowles (2014) study in North Carolina to (a) determine the relationship between the scores of the North Carolina beginning-of-grade (BOG) and end-of-grade (EOG) reading assessments, the scores of the mClass Text Reading and Comprehension (TRC) assessment, and the scores of the CogAT assessment; and (b) determine the degree the TRC, CogAT, and NC BOG predict scores on the NC EOG reading assessment in third-grade classrooms of nine elementary schools. This study was conducted in two parts to best address the research questions. The first part consisted of descriptive, variance, and inferential statistics calculated by demographic variables. This part described the strength of the relationship between the predictor variable (BOG, TRC, and CogAT) and the EOG score on the reading assessment. The second part consisted of calculating multiple regression analyses using the assessment scores. This part described the predictability of BOG, TRC, and CogAT to student scale scores on the reading comprehension portion of the EOG.

This study found that there was a positive correlation and a strong relationship between the NC BOG and NC EOG. The NC BOG scores were statistically significant when predicting the NC reading EOG. Additionally, the second and third grade EOG TRC and the NC EOG had a strong positive correlation and relationship. This study will be important for educators to accurately base instructional decisions on existing and newly collected data.

Keywords: elementary schools, mClass, CogAT, curriculum-based assessment

Table of Contents

Page

Figures

Chapter 1: Introduction

Introduction

Richard Whitmore argued that there are multiple reasons for the reading achievement gap, from academic difficulties to cultural issues (Loveless, 2015). Whitmore postulated that the reading gap transcends country boundaries and is deeper than subpar reading instruction and text interest (Loveless, 2015). Loveless (2015) stated that three most prominent justifications for the reading achievement gap are "biological/ developmental, school practices and cultural influences" (p. 10).

This study examined the current body of knowledge concerning assessments, student achievement, and reading. This research built upon current understanding of these variables within the context of elementary school classrooms in a public school. Chapter 1 provides the context for this quantitative study. Student achievement was measured by mClass text reading and comprehension (TRC), the North Carolina (NC) Beginning of Grade (BOG), and the NC End of Grade (EOG) reading assessments. Student reasoning and problem-solving skills were measured by using the verbal score of the Cognitive Ability Test (CogAT) assessment. This chapter discusses background, rationale for conducting this study, research questions, and key terms and definitions.

Statement of the Problem

The notion of an achievement gap is not new; John Dewey's research introduced this concept. Dewey (1916) wrote, "it is the aim of progressive education to take part in correcting unfair privilege and unfair deprivation, not to perpetuate them" (p. 82). Due to state-wide deficits in reading achievement, lawmakers in NC initiated legislation entitled Excellent Public Schools Act HB 950 (NC Department of Public Instruction [NCDPI], 2015). This law pinpoints strategies and various methods of instruction for assisting student reading proficiency in early elementary grades. NCDPI (2015) Read to Achieve

(RtA) legislation stated that early elementary level students will be assessed with "valid, reliable, formative, and diagnostic reading assessments" (p. 5). The premise behind this law is to increase the likelihood that students will read at or above grade level by the end of third grade to have future educational success. This premise stems from the No Child Left Behind Legislation (NCLB) which postulates that students learn and achieve at an equivalent rate, yet the achievement gap exists and persists. Students are assessed many times during their educational career; therefore, studying the relationship among these data points should allow educators to gain a better understanding of assessment predictors to student success.

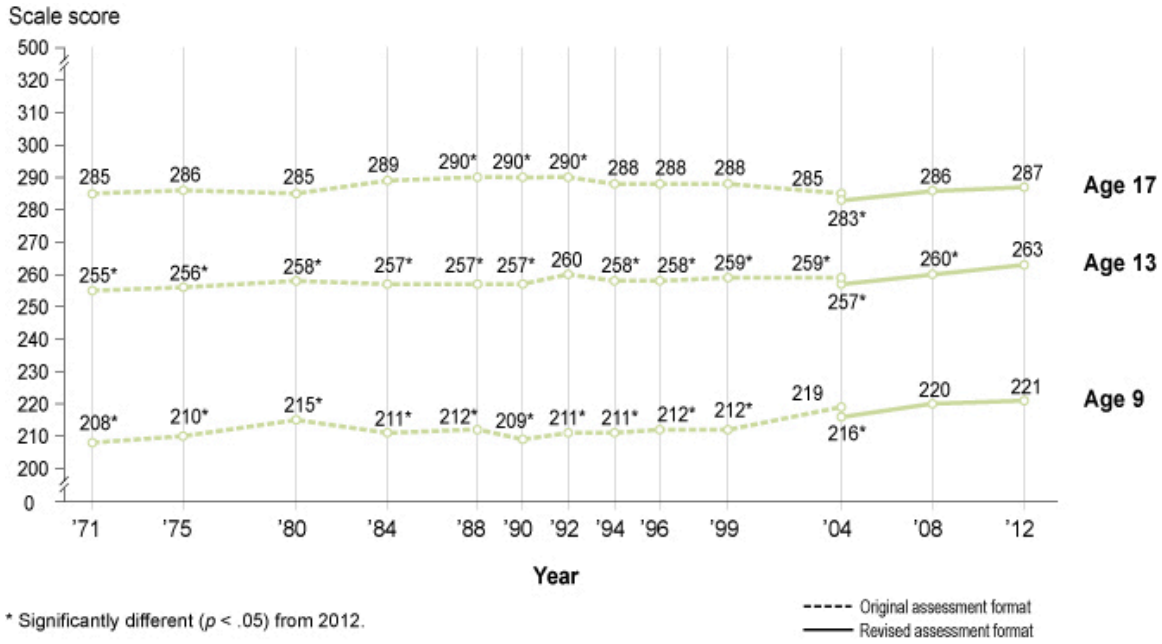


Figure 1. Long-Term National Reading Trends.

Figure 1 shows reading trends from a report by the National Association of Educational Progress (NAEP) that is based on a national representation of 9-, 13- and 17-year-old students attending schools in the United States (National Center for Education Statistics [NCES], 2015). Students in the fourth grade are typically 9 years old, students

in the eighth grade are typically 13 years old, and students in the 12th grade are typically 17 years old. Reading achievement trends from 1971 to 2012 show that there has been a minimal amount of growth over this time frame: 17-year-olds only had a 2-point scale score gain, 13-year-olds had an 8-point scale score gain, and 9-year-olds had a 13-point scale score gain. The line break in 2004 illustrates the change in the reading assessment format in 2004 which re-normed the assessment and altered proficiency levels (NCES, 2015).

Table 1

Poverty Percentage and Report Card Ranking in NC

	A (NG)	A	B	C	D	F
50% or More Poverty	9.5	17.9	27.2	68.7	91.9	98.0
Less than 50% Poverty	90.5	82.1	72.8	31.3	8.1	2.0

Table 1 consists of the letter grade distribution for schools. The table is broken into 2 sections to show scores of schools that have 50% or more of the students who live in poverty and those schools that have less than 50% of the students who live in poverty. All NC schools receive a letter grade based on their EOG scores in reading and math. Each school's score is comprised of 80% achievement and 20% growth. Research shows a correlation between the percentage of poverty and test scores; the higher the poverty percentage, the lower the test score. Typically, schools that have lower grades have 50% or more students living in poverty.

This study took place in Carter School District. The number of students reading on or above grade level has remained stagnant for some subgroups, despite the increased focus on reading instruction in Carter School District and nationally (National Early Literacy Panel, 2008). Reading is categorized as a basic foundational skill that is paramount for academic success which typically develops in early elementary school

(International Literacy Association [ILA], 2007). For many students who fail to develop basic reading comprehension on both fiction and nonfiction texts in early elementary grades, their reading achievement levels are negatively impacted later in school (Routman, 2007).

This research study aimed to investigate the predictability of mClass TRC, the NC BOG and Cognitive Abilities Test (CogAT) scores on the NC EOG reading assessments. NC students must be assessed with BOG, EOG, TRC and CogAT, but no direct or indirect link to the success, or lack thereof, from one NC reading achievement assessment to the other is noted. This study aimed to extend portions of the Bowles (2014) study to determine if the results are transferable to different schools with both similar and different demographics.

Extension Study

The research design for this study was an extension of Bowles's (2014) dissertation, *The Relationship between mClass Reading 3D Assessment and the North Carolina End of Grade Assessment of Reading Comprehension in an Elementary School*. Bowles researched the relationship between scores of the NC EOG Reading Comprehension assessment; the scores from the mClass Reading 3D assessment in third, fourth, and fifth grades; and the degree to which mClass Reading 3D predicted the reading NC EOG scores. This study aimed to extend portions of the Bowles study to determine if the results are transferable to a different school district with both similar and different demographics. Conceptual replication studies differ from the original study and could have one or more different features (Schmidt, 2009). Conceptual replication identifies the generalizability of the original study (Schmidt, 2009).

Bowles (2014) researched the relationship between NC EOG reading and mClass Reading 3D assessment results and the predictability of mClass Reading 3D assessments

to NC EOG scores in one elementary school. Assessment data of students in Grades 3, 4, and 5 were analyzed.

Bowles (2004) conducted a quantitative, correlation study that was conducted in two parts. The first section analyzed inferential, variance, and descriptive statistics calculated by various demographic components. The second section calculated regression analyses and identified the predictability of mClass Reading 3D.

Table 2

Comparison of the Bowles (2014) Study

	Bowles Study	This Study
Participants	One school	Nine schools
Independent Variables	mClass Reading 3D	mClass TRC, NC BOG, CogAT
Subgroups	Sex, race	Exceptional Children (EC), race

This study is an extension of the Bowles (2014) study on the evaluation of NC reading assessments. Table 2 illustrates the conceptual replication changes between Bowles and this study with a difference in population and assessments. This study explored a different population by analyzing data from nine different elementary schools and only analyzing data for third-grade students. Bowles analyzed a sampling from third-, fourth-, and fifth-grade students at one elementary school. Bowles analyzed the mClass Reading 3D assessment, and this study analyzed the TRC section of the mClass Reading 3D assessment, NC BOG reading assessment and the CogAT. Similar to the Bowles study, this study was strictly quantitative.

There are three justifications for extending the Bowles (2014) study. The first justification to include additional elementary schools is to test the generalization of data findings. The second justification for narrowing the grade level to only third grade is to narrow the chance of change and variation with CogAT data, to use the same test and test specification for one grade level and to have three data points because TRC is not

assessed on all fourth and fifth graders. The third reason is to add CogAT assessment data because this assessment is revered higher than mClass Reading 3D components due to the accuracy of data across the school district.

Purpose of the Study

The purpose of reading research was to initiate, create, expand, or validate a reading theory; collect additional reading data; and identify or create improvements with instructional models (ILA, 2014). The purpose of this study was to identify the correlation between reading scores in order to determine predictive validity. By extending the Bowles (2014) study in NC, the researcher aimed to (a) determine the relationship between the NC BOG, TRC, and CogAT reading assessment scores on the NC EOG reading assessment and (b) determine the degree the TRC, BOG, and CogAT predict scores on the NC EOG reading assessment.

Bowles (2014) made nine recommendations for future studies, and this study aimed to focus on three of those recommendations. The first recommendation is to include different schools within this study. The second recommendation is to follow a cohort to determine the impact over time. The third recommendation is to analyze the relationship between mClass and NC EOG based on variable factors such as Academically or Intellectually Gifted (AIG) and Exceptional Children (EC). More details about the Bowles study will be provided in Chapters 2 and 3.

Background

The National Institute of Child Health and Human Development (NICHHD, 2000) noted a massive shift about teaching reading in K-12 public education and the approach to use within each grade level. Teachers of early reading needed to shift instruction as students move from nonreaders to readers (Kennedy et al., 2012). Teachers of proficient readers will have to shift instruction from teaching students how to read to

teaching students to read to learn. Public education school systems shifted the instructional delivery from teaching reading using phonics or using whole language to a combination of the two (Duibhir & Cummins, 2012; Kennedy et al., 2012). Thirty-eight percent of NC fourth graders and 30% of eighth graders are reading at or above grade-level proficiency. During the 2015-2016 school year, there were less than 60.5% of students who performed at or above grade level in reading on each grade level from third to eighth grade. Table 3 shows the individual breakdown of students at or above reading expectation for the state of NC and the school district represented in this study. Proficiency for this data is based on the NC EOG reading assessment for the 2015-2016 school year.

Table 3

Percentage of Students Performing at or Above Reading Proficiency

	NC	Carter School District
Third grade	57.7	58.5
Fourth grade	58.0	57.9
Fifth grade	55.4	55.5
Sixth grade	58.7	59.7
Seventh grade	58.5	60.4
Eighth grade	53.4	56.2

Table 3 shows that Carter School District's scores are very similar to overall results of NC. The state of NC only outscored Carter School District for students performing at or above reading proficiency in fourth grade.

The term assessment is described as a process of strategically gathering and analyzing data about student achievement (Clay, 2001). Based on NC mClass assessment data, students continue to struggle with various areas of reading including but not limited to comprehension, fluency, and phonetics. According to National Reading Panel students who are not proficient readers in early elementary school will continue to struggle with

reading tasks for the remainder of their educational career (NICHHD, 2000). Based on NC data for Carter School District, during the 2011-2012 school year, 68.8% of third-grade students were reading at or above grade-level expectancy compared to the 2016-2017 school year where 58.4% of third graders were reading at or above grade-level expectancy based on NCDPI EOG assessments. Based on these statistics, NC teachers have a long arduous task to shift the future.

Snow, Burns, and Griffin (1998) wrote a report titled Preventing Reading Difficulties in Young Students. This study noted that the most important factors to preventing reading difficulties are a well-prepared and knowledgeable teacher, utilization of language skills and strategies, and a clear understanding of alphabetic principles in reading. This report also noted that reading instruction should include but not be limited to the following components: phonemic awareness, comprehension, letters knowledge, and opportunities to read and write often. Being proficient allows readers to understand and comprehend a variety of texts independently. Students will encounter text levels of varying complexity through content area reading and reading in English language arts. Student success in school after third grade is dependent upon their ability to read and gain meaning and information through text (Bond & Dykstra, 1997). The research from Bond and Dykstra (1997) is still applicable today because students continue to struggle with the basics of reading, which in turn directly affects their comprehension.

Achievement testing of students began in 1845 and was administered to students in a uniform manner (Bond & Dykstra, 1997). Bond and Dykstra (1997) noted that achievement tests are used in education for two purposes: political and professional reform. Political reform refers to the notion that schools are equipping students and individuals with the information and knowledge to take part in the election process. Professional reform refers to a model that educators are free to make the professional

judgments needed for school improvement. Bond and Dykstra conducted research in reading and found that reading is not a singular act but is comprised of numerous integrated pieces.

NCLB mandated that all children without significant cognitive disabilities must work toward the same standards, Common Core State Standards (CCSS; NCDPI, 2015). With the adoption of the CCSS, the state of NC chose to use mClass to measure academic progress of each student. The mClass assessment with subtests for phonemic awareness, phonics, fluency, and comprehension provides schools with a tool that meets the requirements for individual frequent monitoring (mClass: Reading 3D, 2010). More information about this instrument is provided in Chapter 2.

Definition of Terms

The following significant terms are used throughout this research. The definitions are included below.

BOG. Reading comprehension baseline for third-grade students that is administered at the beginning of the school year assessment (NCDPI, 2015).

CogAT. An intelligence assessment that measures quantitative, nonverbal, and verbal ability (Cognitive Abilities Test Form 7 Research and Development Guide, 2012, p. 20).

EOG. Reading assessment that is given to all NC students in Grades 3-8 that assesses reading comprehension skills on grade-level text based on the CCSS (NCDPI, 2015).

Literacy. The ability to receptively and expressively understand and disseminate information (Fountas & Pinnell, 2006).

mClass Reading. mClass Reading is a digit assessment that incorporates TRC and early reading behaviors (mClass Reading 3D, 2010).

Reading comprehension. The ability to actively make meaning, nonvisible processing in the brain, which allows the reader to gather and collect multiple sources of information from the text and construct meaning (Fountas & Pinnell, 2006).

Reading level. The level that a reader can read a text. There are three levels: independent, instructional, and frustration (Clay, 2001). Independent is when a student can read a text with 95% accuracy or higher. Instructional is when a student reads a text with 90-94% accuracy. The frustration level is when a student reads a text with 89% or less accuracy.

Running records. Running records are used to capture reading behaviors of readers while reading a text (Clay, 2001). They are used to guide instruction based on what the reader did and did not do while reading.

Significance of the Study

This research added to the existing knowledge related to the predictability of mClass TRC and CogAT on NC BOG and EOG. There is a wide array of literature about reading, reading practices, and assessments. There are minimal studies that investigate the interconnectedness and predictability of these reading assessments to each other.

This research may benefit teachers, school administrators, and supervisors. It can enhance teacher awareness in problem-solving student needs based on student reading testing data and enhance teacher perceptions of assessment tools and their applications. When a teacher meets a student's specific needs, this type of instruction will improve student achievement. The findings of this study can support leaders to gain a better understanding of assessment results and help teachers enhance teaching practices to improve literacy instruction.

Research Questions

By extending the Bowles (2014) study, this study was designed to determine the predictive validity of the NC reading assessments by answering the following questions:

1. What is the relationship between third grade BOG reading comprehension and third grade EOG reading comprehension?
2. To what extent does the third grade BOG reading comprehension accurately predict student scores on the third grade EOG reading comprehension?
3. What is the relationship between CogAT verbal score and third grade EOG reading comprehension?
4. To what extent does the CogAT verbal score accurately predict student scores on the third grade EOG reading comprehension?
5. What is the relationship between second grade EOY TRC and third grade EOG reading comprehension?
6. To what extent does the second grade EOY TRC accurately predict student scores on the third grade EOG reading comprehension?

Summary

This study posed a quantitative inquiry into the relationship between mClass text levels to BOG and EOG scores and to CogAT scale scores in order to determine the predictive validity of these assessments. There is a limited amount of relevant research available on the combination of the topics in this study; additional research will be needed to strengthen the relationship between the CogAT, mClass Reading 3D TRC, NC BOG assessment, and EOG reading comprehension assessments. This study investigated whether these assessments are strong predictors of the scores on the NC EOG reading comprehension assessment. The results of this study will be valuable for educators because the usage of these assessments is mandated in NC. The identification of these

relationships can help teachers improve effective reading instruction. The results from this study can be used to develop professional development activities for reading teachers. The next chapters consist of information that will enhance the understanding principles that undergird this study. Chapter 2 provides a review of pertinent literature, and methodology is described in Chapter 3.

Chapter 2: Literature Review

Introduction

The purpose of this chapter is to provide pertinent research and background information through the fleshing out of the conceptual framework and a review of literacy literature. Additional components and elements will be discussed that add value to the understanding of student achievement. This chapter contains a review of the literature that deals with reading instruction with an emphasis on reading assessments, testing, and student achievement.

The literature review will discuss the relevant literacy components that contribute to student reading achievement. A brief history of reading research; components of classroom literacy instruction; and the four assessments, mClass TRC, NC BOG, NC EOG, and CogAT emphasized will provide a context for this study. This chapter will conclude with a summary of reading literature.

Overview

Reading is a skill that students need to develop because they will use these skills throughout life to develop an understanding of various concepts in the workplace. Reading requires a tremendous amount of practice and consideration of several complex reading and comprehension processes (Routman, 2003). To be successful, students need to exceed the status quo level of reading comprehension and accuracy proficiency. Literacy skills require certain levels of socialization between individuals and a text. Reading skills occur in a social context (Routman, 2005).

Becoming academically literate at all grade levels is an important skill to master in order for a student to have academic success throughout school. Students often struggle in the area of literacy. The state of NC required public schools to use mClass Reading 3D data three times a year on student reading achievement levels in grades

kindergarten to third grade. Data gathered from these assessments were used to develop and alter whole group and small group reading instruction and literacy interventions for students in need of remediation and enrichment. This study explored the TRC part of the mClass Reading 3D's capacity to aid with the prediction of student scale scores on the reading comprehension assessment of the NC BOG and EOG assessment to figure out if mClass TRC is a useful assessment to use and alter instruction for the nine schools in this study.

Replication

Replication is the intentional duplication of prior research to refute or substantiate preceding data results (Makel & Plucker, 2014). Replication studies first appeared in educational journals in 1938. C. Peters wrote a paper titled, "An Example of Replication of the Experiment for Increased Reliability," that was published in the Journal of Education Research (Makel & Plucker, 2014). There is not an agreed upon list of essential and satisfactory features that comprise a replication study (Makel & Plucker, 2014).

The entire premise of teaching and education developed based on study replication and building upon the knowledge and ability of others. Science inquiry evolves around the notion of replicating studies (Schmidt, 2009). Study replication is important to research because it controls components that affect the validity of studies while measuring the impact on studies (Schmidt, 2009). There are two main types of replication studies, direct and conceptual. Direct replication studies mirror the original study and use the same design, method, and sampling (Schmidt, 2009). These studies assess the accuracy of the original study. Conceptual replication studies may use a different analysis, design, methods, and/or sampling (Schmidt, 2009). Using this replication type allows opportunities to test the construct versus the data or method of the

original study.

Schmidt (2009) noted five roles of study replication: regulate errors with sampling, regulate research artifacts, regulate fraud, generalization of outcomes across populations, and authenticate undergirding hypotheses. The first three roles of replication closely align to direct study replication, while the last two roles are to extend upon the information of the original study. The types of replication formulated by Schmidt can be categorized into two groups, direct and conceptual. The importance of each replication is established by determining the goal of the study. Direct studies seek to confirm the original findings and conceptual studies assess models.

Lykken (1968) postulated three replication types: literal, operational, and constructive. A literal replication is an exact copy of the original researcher's conditions, technique, and methods and procedures. Researchers feel that this type of research is impossible and could have similar bias (Makel & Plucker, 2014). Operational replication exists when researchers attempt to duplicate the sampling and procedures. Constructive replication is when the original methods are avoided and follows claims of the first researcher (Makel & Plucker, 2014).

Toncar and Munch (2008) found three reasons to replicate a study. The first reason is to measure the reliability and validity of the original study. The second reason is to show how variables change based on time and location. The third reason is to assess the external validity of the original study.

Replication is known as the Supreme Court of research (Makel & Plucker, 2014). Makel and Plucker (2014) conducted research to determine how frequent educational journals published relocation studies. They reviewed hundreds of top education journals: 461 of 164,589 articles contained the word replicate; 221 articles were replications; and 18 journals never used the term. The replication publication rate for educational journals

was .13%, whereas the replication rate for psychology journals was eight times higher. Makel and Plucker noted that the reduced publication and lack of publication is decreasing the significance of replication studies. Additionally, 67% of replication studies were conducted by the original researcher.

Education should be built upon sound policy and practice, and replication research is the key to identifying best practices and educational concept correlations. In order to move to a reliable education system, research finding must be deemed valid, reliable, and transferable. Reliance on single study findings will weaken the field of education (Makel & Plucker, 2014).

Bowles (2014) Findings

This section will only discuss the TRC assessment findings from the Bowles (2014) study because this variable is shared with the current study. Bowles found that there was a positive correlation between NC EOG and mClass Reading 3D assessments. Additionally, Bowles found that there was a significant statistical relationship between NC EOG and mClass Reading 3D assessments between students in Grades 3-5. There was not a positive correlation or significant statistical relationship within ethnicity and gender between these assessments. Bowles also found that mClass Reading 3D gave a significant statistical prediction of NC EOG scores.

Bowles (2014) found that in Grades 4 and 5, females outperformed males; and in Grade 3, males outperformed females on all assessments. Hispanic students in Grade 5 outperformed other ethnicities on all assessments. Multi-racial students in Grades 3 and 4 outperformed other ethnicities on the NC EOG. African-America students were the largest ethnicity subgroup in all three grades, and they never outperformed any other group in this study.

Bowles (2014) found that TRC proficiency scores were on average below

proficiency on the NC EOG in Grades 3 and 4. Further, girls outperformed boys in Grades 4 and 5, and boys outperformed girls in Grade 3. Based on ethnicity, Black students never outperformed any other ethnicity on any assessment, and Hispanic students outperformed all other ethnicities in fourth and fifth grade on TRC. Multi-racial students outperformed other participants on third and fourth grade NC EOG assessments.

This study extended Bowles (2014) by investigating the relationship between the NC EOG and mClass TRC. Additionally, this study also included an examination of the relationship between NC BOG scores. Furthermore, Carter School District also used the CogAT to find student academic potential, so this assessment was also included in this study.

mClass DIBELS components are assessed in a pseudo-standardized manner. The teacher of record generally administers the DIBELS assessment and a different teacher administers the TRC; therefore, the creditability of the validity and reliability are impacted. In Carter School District, data are directly linked to the teacher education appraisal system. This study shifted from the analysis of the DIBELS components to the analysis of the CogAT to add an assessment that is administered by a third party, measures ability, and is nationally normed.

Reading

"Readers must read with divided attention to solve words without losing meaning or fluency" (Fountas & Pinnell, 1996, p. 13). Proficient reading is known as the entranceway to societal accomplishment. Becoming a proficient and college and career ready (CCR) reader in early elementary grade levels is directly linked to later educational achievement (Coyne, Kame'enui, & Carnine, 2011). Reading success or failure is determined in early elementary based on a student's performance on standardized assessments (Coyne et al., 2011). Educational disparities exist between unprivileged and

privileged students in kindergarten. This concept is known as the Matthew effect (Coyne et al., 2011).

Reading development. The main purpose for reading shifts around third grade or when a student reaches a third-grade reading level, when reading shifts from learning to read to reading to learn (Spandel, 2008). During the beginning years, teachers implement a variety of interventions during reading instruction for students that range from word work, writing, and reading comprehension to help insure reading success (Spandel, 2008).

Stage	Name	The Learner
Stage 0: Birth to Grade 1	Emergent Literacy	Gains control of oral language; relies heavily on pictures in text; pretends to read; recognizes rhyme
Stage 1: Beginning Grade 1	Decoding	Grows aware of sound/symbol relationships; focuses on printed symbols; attempts to break code of print; uses decoding to figure out words
Stage 2: End of Grade 1 to End of Grade 3	Confirmation and Fluency	Develops fluency in reading; recognizes patterns in words; checks for meaning and sense; knows a stock of sight words
Stage 3: Grade 4 to Grade 8	Learning the New (Single Viewpoint)	Uses reading as a tool for learning; applies reading strategies; expands reading vocabulary; comprehends from a singular point of view
Stage 4: Secondary and Early Higher Education	Multiple Viewpoints	Analyzes what is read; reacts critically to texts; deals with layers of facts and concepts; comprehends from multiple points of view
Stage 5: Late Higher Education and Graduate School	A Worldview	Develops a well-rounded view of the world through reading

Figure 2. Stages of Reading Development. (Roskos, Strickland, Haase, & Malik, 2009, p. 5).

Roskos et al. (2009) depicted reading development through five stages. Each stage has a grade equivalency; the students in this study would be categorized in Stage 2, confirmation and fluency. These readers have reading fluency and can problem solve a variety of words. Stage 0 is comprised of emergent literacy skills such as speaking, book orientation, and beginning phonetics and phoneme skills. Stage 1 students are generally in first grade and is comprised of students learning to read and decode unknown works.

Stage 3 students begin to learn new concepts based on what was read and is identified in students in fourth through eighth grade. Stage 4 readers are in secondary and higher education and begin to analyze texts. Finally, in Stage 5, readers develop a view of the world through reading and this happens in higher education and later (Roskos et al., 2009).

Learning to read. Primary reading instruction places and emphasis on phonological awareness, alphabetic understanding, and automaticity with sounds. Phonological awareness instruction insinuates the understanding and knowledge that words are composed of sounds (Calkins, 1994). Typically, students just learning to read will hear individual and combined letter sounds before they understand that those sounds have letter combinations that match (Coyne et al., 2011). Complete alphabetic understanding is necessary to assist a student to read and write. This developmental understanding may be confusing to some students because of English language is comprised of countless irregularities and variations. A student who possesses automaticity of sounds is about to read and write words fluently and efficiently (Clay, 2001).

Reading to learn. The shift to reading to learn stems from the premise that once a student starts to take content classes and delve in deep into reading comprehension, much of the information they learn will come from texts versus visible concrete teaching (Kerr & Frese, 2017). Dooley and Matthews (2009) stated that in order for students to have reading and learning success, they must learn to read and read to learn simultaneously and continue that into a student's middle school career. The concept is derived from the premise that students construct meaning in diverse ways and the teaching of deeper comprehension skills cannot wait until a certain grade or reading level (Dooley & Matthews, 2009).

Reading difficulties. Kerr and Frese (2017) noted four reasons students have difficulties with instruction: "unpreparedness, lack of motivation, time constraints and underestimation of reading importance" (p. 28). First, students primarily struggle with instruction because they do not have the skills to read the information presented to them. Kerr and Frese found that students in high school read at a level that is three to four levels below their actual grade level. Second, low intrinsic motivation was noted as one of the reasons students struggle with higher level reading and a lack of academic success. Third, students are overwhelmed by the course load or uncertain about instructional expectations (Kerr & Frese, 2017). Finally, students struggle with the notion of reading to learn because they view reading as optional but not required to learn new information.

Reading assessments to drive instruction. The notion of designing instruction based on assessments has been criticized because the perception is that students are not using higher level thinking skills (Badger & Christmann, 2009). Lower level comprehension such as this is the building block for higher order thinking, teaching, and assessment. The utilization of Bloom's Taxonomy within instruction and assessment will shift the level of student thinking and application of concepts from the knowledge level to the evaluation and application level. Using assessments to design instruction will allow content to be organized in a sequential order (Badger & Christmann, 2009).

Legislation

In 2001, NCLB placed an emphasis on reading and made a goal to have all students reading on grade level and have access to national standards. By the end of 2014, NCLB required all students to read on grade level by the end of their third grade school year. Initiatives were implemented in various states to help make this goal a reality. As part of these initiatives, the National Reading Panel (NRP) was created. One of the roles and responsibilities of NRP was to determine the effectiveness of various

teaching practices in reading. NRP created the document "The Report of the National Reading Panel: Teaching Children to Read" (NICHHD, 2000). This document identified five essential components of teaching reading: (a) phonics (b) phonemic awareness, (c) fluency, (d) comprehension, and (e) vocabulary. Additionally, this document stated that early intervention of students is vital to help curb and decrease students experiencing struggles with reading.

In 2003, NC received federal funding from the Reading First grant. NC used this grant funding to give training to teachers on research-based instruction and best practices. The overall goal was to improve the reading skills of students who are not performing on grade level. NC's goal is to have students reading proficiently before or during their third grade school year (NCDPI, 2011).

In 2011, educational waiver implementation allowed states an alternative to the NCLB mandates. With waivers, states agreed to implement a uniform set of standards and assessments that propelled students forward to meet the needs of the work force and higher education. Additionally, states had to add a component to the teacher evaluation system that accounted for student progress.

NC adopted a comprehensive reading plan in 2012 for students in grades kindergarten to 12th grade. The Excellent Public Schools Act has one portion that focused on the development of a comprehensive reading plan to improve reading achievement in public schools at all grade levels. This act is based on the premise that reading instruction is research based and meets rising demands and challenges within reading expectations.

The NC comprehensive reading plan was divided into several sections that would address specific needs pivotal to educational levels from kindergarten to high school. The framework for the Comprehensive Reading Plan stressed six key areas: leadership,

instruction, standards-based curriculum, assessment, professional development and partnership, and communication.

The comprehensive version of this plan outlined expectations for school districts, individual schools, and teachers. Each group has a particular list of points within each of the six areas. The six areas are composed of actions and key indicators to provide guidance to help understand each action area. The primary area of focus stresses the importance of having a print-rich environment within all classrooms.

Another key point within the NC Comprehensive Plan is a section titled Read to Achieve (RtA), written in 2012 (NCDPI, 2016). The goal of this section was to emphasize that all students are reading at or above grade level by the end of third grade. RtA has numerous indicators to be completed at each grade level. A few of these components are providing intervention instruction for struggling students, interval assessments related to key standards, summer camp for students who fail to meet grade level, and transition classes for students who continue to need support after summer instruction and interventions. RtA has a component that requires schools to monitor student growth through assessments. Additionally, RtA noted that student assessment should be frequent and have a comprehensive component. NC placed mandates on students to complete standardize assessments to measure growth and proficiency.

Assessment

Tests are considered to be high stakes if they are used for comparison (Decker & Bolt, 2008). High stakes testing is one of the highest contentious aspects of testing students in elementary school (Decker & Bolt, 2008). High stakes tests impact all individuals in a school differently. For students, this test can determine if a student is promoted or retained, and they also determine the academic path students take in school (Decker & Bolt, 2008). For teachers and teacher assistants, this test determines if one has

job security or not, the level of rating you receive on your evaluation, and also if you receive additional funds or not (Decker & Bolt, 2008). The results of high stakes assessments determine the overall rating and report card grade a school receives (Decker & Bolt, 2008). These high stakes tests are potentially used to compare various aspects in education such as students, teachers, schools, and even states.

In Carter School District, the NC EOG is considered a high stakes test. The students who do not pass the NC EOG, scoring a level three, four, or five, are required to retake the assessment. The students who do not pass the retest are required to attend summer school and retake the NC EOG up to three more times. The growth or lack thereof over the retakes will determine is a student will be placed in a third-grade class, a third/fourth combination class or moved on to the fourth grade (NCDPI, 2016).

ILA (2014) noted that the success or lack thereof can significantly alter the path a student takes in education. ILA found that high stakes tests are used as a key indicator to school districts to help find alterations that need to be made to current curriculum pacing. Decker and Bolt (2008) noted that schools and school districts need to proceed with caution when using high stakes testing for any reason other than the authors' intended purposes (ILA, 2014). High stakes assessments are aligned to the CCSS in NC (ILA, 2014). NC administers high stakes tests to students in Grades 3-12.

The Department of Education uses the EOG tests in NC in a variety of methods. These assessments are considered high stakes tests. The EOG assessments are used as a means of comparing the success or failure of students and public schools (Worthen & Spandel, 1991). In some districts, EOGs are used to make end-of-year promotion decisions for students and to determine whether teachers and building-level administrators are deserving of bonuses or sanctions for low-student performance (NCDPI, 2015).

The NC EOG assessments in science, reading, and math were developed to ensure accountability within schools and districts because the North Carolina General Assembly wanted tests that would measure student knowledge and give them an avenue to judge the educational performance and progress of students, schools, districts, and the state school system (NCDPI, 2015). The North Carolina General Assembly required an instrument to measure student academic progress from 1 year to the next year. The creators of this assessment chose to use developmental scale scores. The scale scores are acquired by computing the number of items answered correctly and then using a formula to translate this raw score into a developmental score (NCDPI, 2000).

Third graders are administered a BOG assessment within the first 3 weeks of the beginning of school. The BOG and EOG have the same developmental scale score scale and achievement levels. The purpose of administering the BOG proves a point for schools and the state to measure growth in reading, because these students have not been assessed in this manner in previous years. The results from the third-grade BOG is compared to the third-grade EOG results to measure student growth for the year (NCDPI, 2015). This assessment was first eliminated in the 2009-2010 school year but was implemented again in the 2013-2014 school year (NCDPI, 2015). The BOG assessment serves as a pretest for third-grade students.

The EOG in reading typically consists of four literary passages and four nonfiction passages. The length of these reading passages is indicative of the diverse styles of reading students encounter while reading throughout school. The questions on the reading assessments were devised to measure student abilities to comprehend, analyze, interpret, and apply the content they read in these passages (NCDPI, 2000).

The EOG yields a developmental scale score, which translates to a particular achievement level. The scale score range varies by grade level and it is also dependent

on the year the EOG is administered. This scale score denotes individual growth in reading and allows schools to measure a student's growth. Additionally, the EOG provides an achievement level for students. The EOG achievement levels are divided into five specific scale score ranges: 1, 2, 3, 4, and 5. The first two levels denote non-mastery of skills, and the last three denote levels mastery of skills and concepts (NCDPI, 2015).

NC adopted CCR academic achievement standards for the EOG assessment during the 2014-2015 school year. NC EOG assessment and the third-grade BOG reading assessment have five achievement levels to determine student proficiency. Each achievement level represents a varying degree of proficiency and level of understanding in regard to the content material that is taught. The higher the level of proficiency a student has achieved, the higher the level of material was learned in reading. Level five represents the highest level a student can achieve, and a level one represents the lowest level that a student can achieve. Table 4 displays the new five achievement levels for NC EOG assessment and third-grade BOG assessment (NCDPI, 2015).

Table 4

Proficiency and CCR Explanation

Achievement Level	Meets On-Grade-Level Proficiency Standard	Meets CCR Standard
Level 5 denotes Superior Command of knowledge and skills	Yes	Yes
Level 4 denotes Solid Command of knowledge and skills	Yes	Yes
Level 3 denotes Sufficient Command of knowledge and skills	Yes	No
Level 2 denotes Partial Command of knowledge and skills	No	No
Level 1 denotes Limited Command of knowledge and skills	No	No

The NC State Board of Education (NC SBE) created descriptors for the five achievement levels.

- Level 1 – Student has limited command of the knowledge and skills contained in the Common Core State Standards for literature, limited command of informational text, and limited command of language when determining the meaning of a word (NCDPI, 2015, p. 4).
- Level 2 – Student has partial command of the knowledge and skills contained in the Common Core State Standards for literature, partial command of informational text, and partial command of language when determining the meaning of a word (NCDPI, 2015, p. 4).
- Level 3 – Student has sufficient command of grade-level knowledge and skills contained in the Common Core State Standards for literature; student is ready for the next grade level (NCDPI, 2015, p. 4).

- Level 4 – Student has solid command of the knowledge and skills contained in the Common Core State Standards in literature, has solid command of informational text, and has solid command of language when determining the meaning of a word (NCDPI, 2015, p. 4).
- Level 5 – Student has superior command of the knowledge and skills contained in the Common Core State Standards in literature, has superior command of informational text (NCDPI, 2015, p. 4).

This external summative assessment provides teachers with information about student learning and teacher effectiveness and is used to evaluate the overall performance of the school. It is imperative that teachers understand the purpose for summative assessment and identify what information can be gleaned from this assessment (NCDPI, 2016).

mClass

Clay (2015) concluded that notable reading assessments must be tailored to the individual. Clay (2015) also noted that these assessments should document student reading responses on various texts. Each assessment should give understanding into the student's skill strengths and weaknesses as they read and problem solve a text. The results of each assessment should be compared with standardized behaviors used by children who were successful reading. The instruction that ensues the assessment should be grounded on the results found in the individualized reading assessment and not generalized grade-level instruction.

mClass Reading 3D (Reading 3D) is an assessment that encompasses phonics, phonemic awareness, fluency assessments, and comprehension components (mClass: Reading 3D, 2010). Reading 3D is a formative assessment measure that incorporates DIBELS assessment and TRC assessment. The TRC portion of the assessment involves a

running record which allows teachers to identify a student's accuracy rate, fluency rate, and oral and/or written comprehension level. mClass assesses student reading using both fiction and nonfiction passages and texts (Reading 3D Brochure, 2009). NC adopted this reading assessment titled mClass Reading 3D in the 2013-2014 school year (NCDPI, 2015).

There are seven different components that DIBELS measures: oral reading fluency (ORF), retell fluency, letter naming fluency, initial sound fluency, phoneme segmentation fluency, nonsense word fluency, and DIBELS maze comprehension. Each component provides a teacher with an array of data of each student who is being assessed that paints a picture of the student's strengths and weaknesses in the area of reading (mClass Reading 3D, 2010). This assessment is formally administered to students three times a year: beginning of the year, middle of the year, and end of the year. mClass also has an ongoing assessment measure that is used to periodically monitor students at risk for reading failure (Reading 3D Brochure, 2009). Each student's proficiency level determines the frequency at which the teacher progress monitors each student. There are three proficiency levels in each component: above level, on level, and below level. The cut scores for each level is dependent upon the student's grade level.

mClass Reading 3D components are administered with a computerized device. With the TRC components, the person administering the assessment digitally records each student's performance while they read a leveled text or book. Notations of errors, corrections, omissions, verbal processing, and time are recorded to analyze and address student needs. By using a digital system to administer the assessments, teachers have instant access to data and can address a student's strengths and concerns without delay.

The National Center on Response to Intervention determined the reliability and validity of the TRC. Concurrent validity was identified as 0.72, predictive validity was

identified as 0.76, marginal reliability was identified as 0.86, and the inter-rater reliability was identified as 0.73 (mClass Reading 3D, 2010).

The TRC assessment component is based on running records. Marie Clay identified running records as a formative reading measure in the 1960s. "If RRs [running records] are taken in a systematic way they provide evidence of how well children are learning to direct their knowledge of letters, sounds, and words to understanding the messages in the text" (Clay, 2015, p. 49). Typically, running records are used to guide reading instruction, identify and monitor reading levels, and identify reading progress at certain times (Clay, 2015).

After the running record section of the assessment is completed, the child retells various portions of the text read and their interpretation of what they read. This is the basic level of the comprehension assessment portion. Other assessments require students to answer additional comprehension questions and write out their responses (Fountas & Pinnell, 2001). Upon completion of the comprehension section, the teacher analyzes the information gathered from the running record and oral and/or written comprehension to paint a picture of each individual reading need (Clay, 2015). From this data, a student's reading and comprehension level is identified for work in a small group or in a one-on-one teaching scenario.

Clay (2015) identified the original reliability of running records. Within this research, Clay (2015) identified the error rates to be correlated at r=.98. She also identified the self-correction rates as r=.68. Based on a chi-square test, there was no significant differences at the .01 level based on the scoring and recording behaviors based on self-correction and errors (Clay, 2013).

In 2011, Coyne et al. researched the effects of additional early reading instruction and the effects on achievement using mClass Reading DIBELS and TRC assessments.

The results of this research noted that when students are given appropriate and accurate reading instruction, their reading achievement increased.

The mClass written component of their assessment can be defined as a constructivist aspect of assessment. A constructivist perspective utilizes a performance assessment which allows students to demonstrate their level of understanding through a written response. The written portion of the assessment is graded based on a rubric (Anderman & Sinatra, 2009).

CogAT

The Cognitive Abilities Test (CogAT) is a standardized assessment that evaluates and assesses student intelligences. This assessment assesses verbal, quantitative, and nonverbal areas. The results of these areas are scored in isolation and collaboration, and a score is calculated to obtain a composite score for each student assessed (Riverside Publishing, 2012). The authors of the CogAT measured the internal consistency of the whole assessment and its subtests as being in the .90 range or higher with this type of measurement (Riverside Publishing, 2012).

The Cognitive Abilities Test Form 7, CogAT 7, is a widely used test for students from kindergarten through high school. The test measures a student's reasoning abilities that are considered a crucial factor to distinguish gifted learners (Lohman, 2012). The CogAT scores are used to discover the gained reasoning skills through educational experience (Lohman, 2011). The CogAT contains two major parts: the full battery test and the screening test. The full battery test is used to measure children's cognitive abilities; the screening test is used to offer fast and reliable signs of children who need gifted education services. The screening test is a shorter form of the full battery and includes all subtests located within the full battery test (Lohman, 2012). The CogAT 7 has a quantitative, nonverbal, and verbal battery.

The seventh edition of CogAT has three avenues of interpreting the test results. "Score Levels" use a median age stanine (1 to 9) score scale, where 1 refers to the lowest score and 9 refers to the highest score. Figure 3 shows the CogAT median stanine scale. "Score Patterns" describes student results based on their Age Percentile Rank (APR). There is a system that supports individuals when deciphering CogAT data categorized as (A, B, C, or E) profiles: (A) profile means the student is at the same level in all batteries; (B) profile means the student is below or above in one or more of the batteries; (C) profile shows there is a contrast between two scores; and (E) profile means there is an extreme difference between the scores. "Ability Profile" uses the above two methods together along with + or - signs to refer to student strength or weakness (Riverside Publishing, 2012).

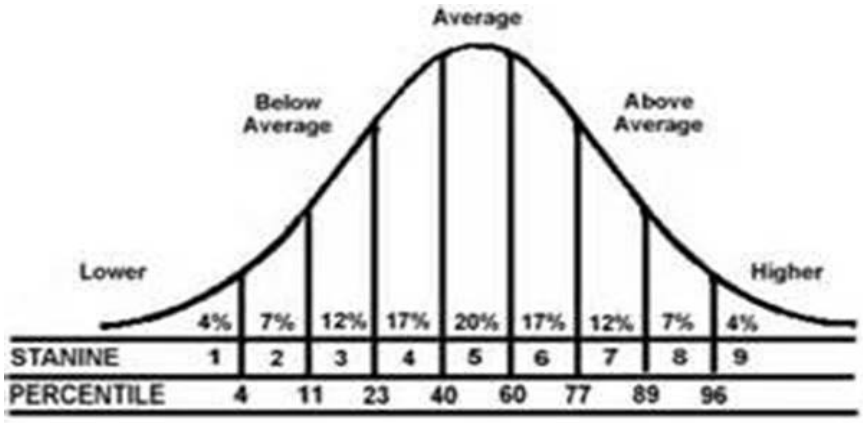


Figure 3. CogAT Stanine Bell Curve (Riverside Publishing, 2012).

The identification of students typically takes place during the early years in elementary school with the use of the CogAT scores. The use of CogAT scores as an educational prediction of success can provide countless benefits to a school (Gottfredson & Saklofske, 2009). Cognitive ability has been researched to be the leading factor to academic performance (Gottfredson & Saklofske, 2009).

Lohman, Gambrell, and Lakin (2008) noted that there should be multiple criteria to identify students as gifted because students typically have discrepancies within profile scores on CogAT. Lohman et al. also postulated that the use of the overall composite score could decrease the number of capable students from gifted and talented programs. Lohman et al. found that academic aptitude and reasoning abilities have the greatest impact on academic learning.

Ability and Achievement

Ability levels have been linked to student achievement. Research found that the relationship between ability and achievement was correlated; the correlation between ability assessments and achievement assessments was nearly .50 (Parker & Benedict, 2002). Parker and Benedict (2002) noted that school psychologists generally interpret the performance of one as an integral link to success in school achievement. Ability assessments given in early elementary school were also able to predict reading achievement years later. Parker and Benedict also noted that the correlations between these assessments were high because numerous tests ask similar questions.

The terms ability test and achievement test can best be identified as "jangle fallacy" (Lohman, 2006, p. 2). Jangle fallacy is defined as the tendency to treat words that sound different as if they signified different ideas and concepts (Lohman, 2006). Researchers believe that it is not possible to create an ability assessment that is unaffected by learning/achievement (Lohman, 2006). There is a strong correlation between student ability and achievement outcomes (Lohman, 2006). Lohman stated, "The measurement of achievement now emphasizes national comparisons while the measurement of ability increasingly emphasizes local comparisons" (Henshon, 2014, p. 8). Multiple points of view are essential for both achievement and ability assessments. Lohman noted that ability tests are more useful for students with limited background and educational

experiences (Henshon, 2014). Lohman also noted that ability assessments are more beneficial for students with talent who have not had ample opportunities to develop these talents (Henshon, 2014). This is why these types of assessments are typically given to younger students, because of the usefulness of the data gained. "Achievement tests sample a broad range of knowledge and skills acquired primarily in school … and ability tests emphasize reasoning abilities that ate required by and developed through experience in and outside of school" (Lohman, 2006, p. 4).

The idea of achievement and ability tests can be summed up through Robert Snow's theory which addresses the idea of a student's readiness to learn (Lohman, 2006). Aptitude to learn is directly related to what must be learned and to the learning context (Lohman, 2006).

To date, educators use standardized and performance assessments as well as anecdotal records to assess student needs. These assessments vary in format, administration, and objectivity. Standardized assessments are categorized as assessments that are administered within specific parameters. Performance assessments are used to gather a variety of information for a student to determine achievement levels. Anecdotal records are notes recorded by the teachers that speak specifically about a student's performance based on observations. Assessment is defined as, "the process of gathering information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences" (Huba & Freed, 2000, p. 8). Reading assessments are used to evaluate the strengths and needs of students' various reading areas to include but not be limited to comprehension, decoding, fluency, and vocabulary. Huba and Freed (2000) noted that it is vital to assess students using standardized, performance, and anecdotal assessments because the combination will depict an accurate picture of student

performance levels. The combination of assessments will increase the reliability and validity of the achievement results.

Within schools, grade levels, and classrooms, achievement levels of students can be distinctively different. Despite similarities in settings, instructional delivery, and instructional time, there is a difference in student achievement levels in reading. The achievement gap can be identified when a subgroup of students outperforms another group of students.

Achievement Gap

Regardless of the location, there continues to be an achievement gap between different ethnicities. Researchers argue that minority students may feel that academic success is synonymous to acting White (Wright, 2011). Students who feel that academics are linked to giving up or altering their ethnic roots will continue to struggle with the balance between self-identity and academic success. Wright (2011) noted that those students who outperformed their ethnic peers view academic success as cool and have a stronger home support than others. Additionally, those students who had academic successes defined their ethnicity as being more than the clothes they wear, manner in which they spoke, and music they listened to (Wright, 2011).

Graham and Erwin (2011) found similar commonalities as Wright (2011); they noted that African-American gifted students indicated that school was their responsibility and had a significant amount of support from home. The academically gifted students also reported that their identity was tied to their successes.

The Annie E. Casey Foundation (2013) stated that the achievement gap with between White and Black students should be considered a national crisis. The Foundation results found that Black students had the lowest reading scores all students in fourth and eighth grade. When they compared all states, Wisconsin had the lowest score,

which was 87% below the national average; and Wisconsin's score for White students' reading performance scores were 44% higher than the national average (Annie E. Casey Foundation, 2013).

Chin and Phillips (2004) found that the achievement gap between Caucasian and African-American students in reading is greater than the achievement gap between Caucasian and Latino students in reading. The reading achievement gap between Caucasian and African-American students in fourth grade was .83 standard deviations, and the reading achievement gap between Caucasian and Latino students in fourth grade was .70 standard deviations. Chin and Phillips noted that this gap continues to eighth grade reading scores.

Madrid (2011) noted that Latino students have made notable gains in their academics in the past 30 years; but in comparison to Caucasian students, the gains are insignificant. On a California assessment in 2010, only 48% of Latino students passed language arts assessments in comparison to 82% of Caucasian students (Madrid, 2011). Madrid noted that there are potentially five factors that contribute to Latino students' low scores: nature of the schools that serve Latino students, teacher quality and how teachers teach, programs and services for Latino students, teacher perception of students, and teacher perception of parents. Madrid noted that all stakeholders need to reevaluate the education system as it pertains to teaching Latino students because their current levels of performance and trends are dismal.

Research shows that students of color made significant gains in education in the 1960s to 1980s, but growth beyond this point seems to have hit a plateau. Hoffman, Llagas, and Snyder (2003) found that African-American students are more likely than Caucasian and Latino students to be enrolled in nontraditional primary education programs that focus more on the whole child than rigorous academics. Additionally,

minority students are more likely to attend a public school where they are the majority instead of the minority (Hoffman et al., 2003). The percentage of minority students continuing their education is on a rise, yet the achievement gap continues to increase.

Studies of students who come from low and middle socioeconomic statuses found that family and student values tend to be the leading contributor of student success or lack thereof in school (Sampson, 2002). Access to opportunities and books were the second leading contributor to lower achievement levels of students. Additionally, peer pressure and teacher expectations contributed to the success or detriment of student school achievement (Sampson, 2002).

U.S. Department of Education (2008) noted that the increase in the achievement gap is due to the lack of schools developing academic ability. "Academic ability is one expression of human intellective competence this is recognized as the universal currency of societies" (Bennett et al., 2004, p. 7). Academic ability is the cultivation and education of the whole child, not just the processing of mathematics and literacy. This ability includes but is not limited to

> critical literacy and numeracy; mathematical and verbal reasoning; skill in creating, recognizing and resolving relationships; classification of information and stimulus material; problem solving from both abstract and concrete situations, as in deductive and inductive reasoning; sensitivity to multiple contexts and perspectives; skill in accessing and managing disparate bodies and chunks of information; resource recognition and utilization; and self-regulation. (Bennet et al., 2004, p. 7)

Researchers noted that since academic ability is a developed ability, this ability needs to be cultivated in all environments in which students interact. Research has found that students of color are performing lower than other students. Recent research has been

focusing on African-American students in poor communities because they are on the lowest end of most assessments (Battle & Linville, 2006). Studies show that African-American, American India, and Hispanic students' school entry reading achievement levels are lower than Asian and White students (Battle & Linville, 2006). These achievement gaps continue into high school; African-American and Hispanic students enter high school an average of 3 years behind White and Asian students.

Based on NAEP reading assessments, the reading achievement gap among African-American and White students was first noted in the 1970s. African-American students performed on average one standard deviation lower than White students (Reardon, Valentino, & Shores, 2012). This gap has continued to fluctuate; and in 2008, the gap was 0.6 of a standard deviation apart (Reardon et al., 2012). The reading achievement gap among Hispanic and White students is similar to that of the African-American pattern. Reardon et al. (2012) noted that the performance of students of different races exists among all socioeconomic levels.

Gender

Similar to ethnicity, there is also an achievement gap between genders. As early as 1984, there was a noted difference between the performance of girls and boys. Eccles, Midgley, and Adler (1984) identified a subject matter gap between girls and boys; boys outperformed girls in math, and girls outperformed boys in reading. Additionally, girls outperformed boys in regard to their class grades.

Preckel, Götz, Pekrun, and Kleine (2008) conducted a study and noted that there was a difference between gender among typically developing students and students identified as gifted, but the difference is larger between gifted students. Although girls had higher aptitude scores than boys, boys generally outperformed girls in math. Preckel et al. found that the gender difference exists between students from the USA and

Germany.

Hyde (2005) noted teachers previously taught all genders the same even though girls and boys came in to education with different backgrounds and experiences with education. Expectations placed on students by caregivers, parents, and society are significantly diverse; but once students enter school, they are expected to be on the same academic level. Although students are expected to be on the same level, each level varies according to the school they attend. Based on the differences placed on boys and girls by parents and society, students come to school with drastic brain-based and biological differences. Even though the brains of boys are girls are the same, the regions and hemispheres of the brain are utilized differently (Hyde, 2005).

Wright (2011) noted that achievement scores vary school to school and state to state. The achievement gap spread varies among students, subjects, and classes with little to no difference in instructional strategies and implementation. The gap also exists between males and females of the same race (Wright, 2011). Research shows that an achievement gap exists among race, sex, and ethnicity; but the origin and cause of the various achievement gaps are unclear.

Reading achievement gaps are new to education. The academic gap among gender varies according to the subject that is being discussed. Achievement levels for male and female students have been approving over the years, but there is still a disparity between performance (Robinson & Lubienski, 2011). The trends in reading and math are opposite of each other. Male students outperform female students in math achievement.

Female students in Grades 2-8 have outperformed male students for nearly 50 years (Robinson & Lubienski, 2011). The reading achievement gap is not just a phenomenon in the United States. In fourth grade, female students outperform male students in 38 of 40 countries based on a literacy study in 2006 (Robinson & Lubienski,

2011). Robinson and Lubienski (2011) found that the reading achievement gap begins as early as first grade, and females are more proficient than male students. Female students are noted to have higher reading comprehension skills than male students in third grade.

Studies have been conducted on student attitude, self-confidence, parental involvement, teacher experience, and more; but the root cause for disparities in reading achievement are still unknown. Researchers are beginning to research the interplay of socialization, psychology, and biology to help determine the root cause of academic gender gaps (Robinson & Lubienski, 2011). The math achievement gap tends to fluctuate based on the age and grade of students; the reading gap tends to remain constant over time. Teachers tend to be the greatest lever in student performance regardless of community location (Robinson & Lubienski, 2011).

Summary

Based on the research, there are several factors that actively effect and impact the teaching and learning of reading. Although there are countless views and teaching methods for reading, several theorists believe that student reading ability is impacted the most by individualized lessons by an observant teacher (Calkins, 2011).

Comprehension strategies increase reader ability to understand word meaning. A teacher providing opportunities for students to activate background knowledge is one avenue that brought meaning to text, which increases student reading achievement. A student's increased reading achievement with various strategies can help build a reader's knowledge of the world and language skills. As students venture through school and life, he or she continuously builds upon reading levels and has to make meaning or comprehend through asking questions.

"Meaningful reading is defined as reading that is accountable, moderately expressive and highly leveraged" (Lemov, 2010, p. 254). Teachers have the power and

should have the resources to ensure that all literacy tasks are meaningful and positively impact student achievement in literacy.

Several research studies were conducted on the DIBELS component of the mClass Assessment using the ORF component which analyzed the predictability to the NC EOG reading assessment. Currently, there is little research that has been conducted on the TRC portion and its predictability to the NC EOG assessment. The research study aimed to fill the research void by analyzing the NC BOG, mClass Reading 3D TRC, CogAT assessment data, and the predictability to the reading comprehension portion of the NC EOG assessment.

Chapter 3: Methodology

Introduction

The purpose of this study was to extend the Bowles (20140 study in NC to determine (a) the relationship between the scores of the NC BOG and EOG reading assessment, the scores of the mClass TRC assessment, and the scores of the CogAT Assessment; and (b) the degree the TRC, CogAT, and NC BOG predict scores on the NC EOG reading assessment (Bowles, 2014).

This study analyzed archival data over a 2-year period during the 2015-2017 school years. The research questions were answered by gathering archived quantitative data from participant test results. The data collected included assessment scores from third-grade students for the NC reading BOG, EOG, those students' second grade CogAT scores, and their EOY second mClass TRC levels. Two years of data were collected and analyzed by statistical tests in order to draw possible conclusions. This chapter describes the study's methodology, including a reiteration of the research questions, the research type and design, a description of the participants, and the methods of data collection and analysis.

Research Questions

Based on the current body of literature, this study sought to investigate the predictability of mClass TRC, NC Reading BOG, and CogAT on NC Reading EOG assessment. The research questions for this study were

1. What is the relationship between third grade BOG reading comprehension and third grade EOG reading comprehension?
2. To what extent does the third grade BOG reading comprehension accurately predict student scores on the third grade EOG reading comprehension?
3. What is the relationship between CogAT verbal score and third grade EOG

reading comprehension?

4. To what extent does the CogAT verbal score accurately predict student scores on the third grade EOG reading comprehension?
5. What is the relationship between second grade EOY TRC and third grade EOG reading comprehension?
6. To what extent does the second grade EOY TRC accurately predict student scores on the third grade EOG reading comprehension?

Setting

This study examined the relationship among elementary reading assessments used in NC. The study was conducted in Carter School District (pseudonym) which is located the southwestern piedmont area of NC. There are 12 school districts in this region of NC.

Table 5

District Comparison of Third Grade EOG Data

	Students assessment	Percent CCR	Percent Grade Level Proficient
District 1	279	38.0	46.6
District 2	2493	48.0	60.0
District 3	429	31.2	43.6
District 4	1196	46.4	58.9
District 5	2366	39.5	51.0
District 6	1476	48.6	59.5
District 7	425	57.4	70.8
District 8	848	54.5	66.5
Carter School District	12257	46.5	58.4
District 10	1530	39.0	51.0
District 11	655	45.3	59.5
District 12	2984	58.8	70.4

Table 5 compares third grade reading EOG results of the 12 school districts in the region. Carter School District had the largest number of third-grade students assessed. Additionally, Carter School District had the sixth highest number of students performing at the CCR level and the eighth highest score of students performing at the grade-level

proficiency.

Based on the 2017-2018 school year, Carter School District had over 147,000 students in grades kindergarten through 12. There were 176 schools in all: 93 elementary schools (k-5 and PK-5), 44 middle schools (6-8), 36 high schools (9-12), eight prekindergarten through eighth-grade schools (PK-8), one kindergarten through 12th-grade school (K-12), one sixth- through 12th-grade school (6-12), and three alternative schools. Forty-seven of these schools were magnet schools. The students in this district represent over 160 different cultural and ethnic backgrounds. The student demographics were 3% American Indian/Multiracial, 6% Asian, 24% Latino, 28% Caucasian, and 38% African American. There are more than 18,000 employees in Carter School District. Approximately 9,100 of these employees are certified teachers. During the 2017-2018 school year, the graduation rate was over 89%.

The research took place in the elementary section of nine K-5 schools within Carter School District. There are eight K-5 schools and one PK-5 school. Table 6 shows the ethnicity ratios for all students who were enrolled in these schools during the 2016-2017 school year.

Table 6

Ethnicity Breakdown

	Meares	Gause	Rogers	Samuels	Crawford	Pickney	Cooper	Mapp	Davis
Caucasian	70.3	45	67	43.8	66.3	74.3	67	80	35
African American	21.8	14.3	9	40.5	9.5	8.3	8	9	39
Asian	2.6	19	16	0	3.9	3.6	14	9	5
Latino	0	17.5	7	6.8	18.2	10.1	8	0	16
Multi-racial	2.2	0	1	4.3	0	3.5	3.5	0	0
Other	3.1	4.2	0	2.1	2.1	0	2	2	5

Table 6 shows the subgroup number for the students who were enrolled in the school for the 2016-2017 school year. All schools have been given pseudonyms to ensure confidentiality. Caucasian students represent over 35% of the total population. The percent of Latino students represents the next largest ethnicity group at Crawford and Pickney Elementary. African-American students represent the next largest ethnicity group at Meares and Samuels Elementary. Asian students represent the next largest ethnicity group at Gause, Rogers, and Cooper Elementary. The schools have a wide range of ethnic representation.

For the 2016-2017 school year, Meares Elementary had a total of 735 students, Gause Elementary had a total of 737 students, Rogers Elementary had a total of 913 students, Samuels Elementary had a total of 634 students, Crawford Elementary had a total of 814 students, Pickney Elementary had a total of 1033 students, Cooper Elementary had a total of 515 students, Mapp Elementary had a total of 740, and Davis Elementary had a total of 746 students.

The schools in this study performed similarly on the NC EOG reading test as compared to the rest of the schools in the district. The data in Table 7 show the level of

proficiency on the NC EOG reading assessments in all test grades at each school.

Table 7

2016-2017 Reading EOG Proficiency Results

School	Level 1	Level 2	Level 3	Level 4	Level 5
Meares	6.8	15.2	9.8	37.9	30.3
Gause	10.8	12.3	13.1	44.6	19.2
Rogers	5.7	8.2	8.8	40.3	37.1
Samuels	5.3	19.5	15.0	41.6	18.6
Crawford	10.0	11.4	5.7	48.6	24.3
Pickney	5.5	18.8	17.1	34.3	24.3
Cooper	<5	12.2	11.0	41.5	31.7
Mapp	5.0	20.0	17.5	41.3	16.3
Davis	5.0	20.0	17.5	41.3	16.3

Table 7 represents the nine schools that were used in this research. This table shows the reading proficiency for the third-grade students who take the NC EOG test. Each of the schools have a noted percentage of students in all achievement levels except Cooper Elementary for level 1. On average, those students performing below grade-level proficiency range from 12% to 25%. Each school has 70% of their population who scored level 3 or higher. Mapp Elementary and Davis Elementary had the largest percent of students in Levels 1 and 2, which is 25%.

When analyzing data, identifying the number of students who perform at each achievement level is essential as well as identifying which subgroups are meeting expectations. The highest level of proficiency is considered to be CCR, which denotes those students who scored a level 4 and level 5. There is a noted disparity between Caucasian and African-American and Latino students. The star denotes subgroups that did not have enough students to statically analyze the data.

Table 8

2016-2017 Reading EOG Results for CCR (Levels 4 & 5)

	Students	Female	Male	Asian	African American	Latino	Caucasian
Meares	68.2	72.3	64.2	*	21.4	*	81.7
Gause	63.8	72.6	52.6	61.9	47.4	54.5	73.4
Rogers	77.4	79.1	75.3	77.4	55.6	61.5	84.0
Samuels	60.2	66.7	54.2	*	46.8	*	73.1
Crawford	72.9	78.7	68.4	*	50.0	34.6	84.2
Pickney	58.6	60.4	56.5	*	26.3	50.0	63.3
Cooper	73.2	75.0	71.1	80.0	*	*	66.7
Mapp	76.9	84.1	69.0	*	63.6	*	80.6
Davis	57.5	62.3	53.8	*	43.9	44.4	72.4

*p<10.

Table 8 illustrates the CCR levels represent the highest two levels a student can achieve on the reading EOG. There are a few schools that have a star listed instead of a percentage because these schools did not have a large enough population in the listed subgroup for the Department of Education to calculate the percentage. In all schools, the female students outperformed the male students. When listed, Asian students outperformed other ethnic groups.

Table 9 shows the second proficiency grouping, which is considered those who met grade-level proficiency. These students are considered to be on or above grade level in the subject area that is being assessed.

Table 9

2016-2017 Reading EOG Results for Grade Level Proficiency (Levels 3-5)

	Students	Female	Male	Asian	African American	Latino	Caucasian
Meares	78.0	80.0	76.1	*	32.1	*	90.3
Gause	76.9	84.9	66.7	81.0	63.2	72.7	79.7
Rogers	86.2	87.2	84.9	87.1	61.1	76.9	91.5
Samuels	75.2	79.6	71.2	*	66.0	*	82.7
Crawford	78.6	82.0	75.9	*	60.0	46.2	88.4
Pickney	75.7	77.1	74.1	*	52.6	75.0	78.4
Cooper	84.1	81.8	86.8	>95	*	*	76.5
Mapp	86.0	93.7	77.6	*	63.6	*	89.2
Davis	75.0	78.3	72.5	*	66.7	55.6	87.9

*p<10.

In Table 9, the grade-level proficiency indicates students who received a score of a level 3, 4, and 5. These levels indicate students who passed the reading EOG assessment. Each school has an increased proficiency. The star denotes there were not enough students in that subgroup to calculate the percentage. All of the schools had 75% of students score a level 3 or higher on the reading EOG during the 2016-2017 school year. Female students outperformed male students. White students outperformed other ethnic groups except at Gause Elementary and Cooper Elementary. Black students were the lowest performing group at all schools except Crawford Elementary and Davis Elementary, where Hispanic students performed the lowest.

Table 10 indicates the years of experience of all the teachers in the nine schools involved in this study. The majority of all the teachers in the schools have 10 or more years of experiences, ranging from 40% of teachers at Gause Elementary to 66.7% of teachers at Samuels Elementary.

Table 10

Percentages of Teacher Years of Experience Per School

	0-3	4-10	10+
Meares	32.5	27.5	40
Gause	7.0	27.9	65.1
Rogers	26.5	20.4	53.1
Samuels	17.9	15.4	66.7
Crawford	6.7	28.9	64.4
Pickney	15.4	34.6	50.0
Cooper	20.6	26.5	52.9
Mapp	15.0	27.5	57.5
Davis	31.0	21.4	47.6

*p<10.

One of the most important indicators of student success is teacher experience and teacher morale. Teachers are more likely to stay in schools where they find success. Student achievement is directly related to teacher knowledge and experience (Fountas & Pinnell, 2001). Table 10 indicates that these schools have experienced teachers.

In NC, all schools are evaluated and receive report cards made available to the general public. Each school receives a grade in tested subjects and an overall performance grade. The grade is comprised of 80% of the school achievement score in reading, and 20% is based on academic growth. These scores are published in various reports to be shared with stakeholders.

Table 11

School Report Card Rating in Reading

School	Rating
Meares	B
Gause	C
Rogers	B
Samuels	C
Crawford	B
Pickney	B
Cooper	B
Mapp	A
Davis	B

Each school's grade for the 2016-2017 school year is provided in Table 11.

Grades range from A to C with a mode grade of B. Mapp Elementary received a score of A, Gause and Samuels Elementary received a score of C, and all other schools received a score of B in reading.

Carter School District encompasses a wide array of students, teachers, and ability levels. Carter School District has areas of success and areas that need improving in reading. Reading is essential for student academic success, unified growth, and achievement. Within the researched school district, even schools that have a large percentage of students performing at or above grade level proficiency still have a large achievement gap between students of color.

The schools in this research study were selected to participate by using the following criteria: (a) the school was not identified as a Magnet school, (b) the school had a reportable subgroup of EC which is equal to or greater than 10 students, and (c) each school had a Talent Development population that was equal to or greater than 10 students. Each of the schools in this research study met all of the three data points being used in this study. The number of students who are categorized as Economically Disadvantaged (EDS) and Limited English Proficient (LEP) was not used to restrict the school sampling of this study. The number of students with a disability (SWD) and AIG students were also analyzed in this study. Based on NC accountability department, each subgroup listed above was disaggregated. Students in each subgroup who passed the reading EOG were divided into two groups: CCR and grade-level proficient (GLP). Level 1 and 2 scores are considered not proficient, level 3-5 scores are considered GLP, and level 4 and 5 scores are considered CCR.

Table 12

Number of Third-Grade Students by Subgroup at CCR and GLP Levels

School	EDS	LEP	SWD	AIG
Meares-CCR	28	<10	13	37
Meares- GLP	28	<10	13	37
Gause-CCR	37	15	10	26
Gause- GLP	37	15	10	26
Rogers-CCR	19	20	12	34
Rogers- GLP	19	20	12	34
Samuels-CCR	29	<10	14	23
Samuels- GLP	29	<10	14	23
Crawford-CCR	38	27	13	28
Crawford-GLP	38	27	13	28
Pickney-CCR	33	<10	14	25
Pickney- GLP	33	<10	14	25
Cooper-CCR	17	<10	12	22
Cooper- GLP	17	<10	12	22
Mapp-CCR	22	<10	14	32
Mapp- GLP	22	<10	14	32
Davis-CCR	72	12	11	24
Davis- GLP	72	12	11	24

Table 12 shows that the number of EDS ranges from 17-72, Cooper Elementary to Davis Elementary respectively. Four of the schools, Gause, Rogers, Crawford, and Davis Elementary, had an identifiable subgroup, more than 10, in third-grade students who are classified as LEP. Gause Elementary only had 10 third graders who are identified as SWD and Samuels, Pickney, and Mapp Elementary each had 14 SWDs. All schools had at least 22 students but not more than 37 students who were identified at AIG in third grade.

Participants

There were nine elementary schools in this study. The number of third-grade students assessed during the 2016-2017 school year on the third grade reading assessment ranged from 82-181 in each school. There were approximately 1,800 students for both years in this study; about 900 students in year one and 900 in year two.

All nine schools in this study had a notable achievement gap among students of different ethnicities. Each of the schools in this study had less than 10 students identified as American Indian and two or more races; therefore, this information was not included in the ethnicity breakdown.

Table 13

CCR and GLP Breakdown by Ethnicity

	Asian	African American	Latino	Caucasian
Meares-CCR	<10	28	<10	93
Meares- GLP	<10	28	<10	93
Gause-CCR	21	19	22	64
Gause- GLP	21	19	22	64
Rogers-CCR	31	18	13	94
Rogers- GLP	31	18	13	94
Samuels-CCR	<10	47	<10	52
Samuels- GLP	<10	47	<10	52
Crawford-CCR	<10	10	26	95
Crawford-GLP	<10	10	26	95
Pickney-CCR	<10	19	20	139
Pickney- GLP	<10	19	20	139
Cooper-CCR	15	<10	<10	51
Cooper- GLP	15	<10	<10	51
Mapp-CCR	<10	11	<10	93
Mapp- GLP	<10	11	<10	93
Davis-CCR	<10	57	27	58
Davis- GLP	<10	57	27	58

Only Gause Elementary and Rogers Elementary had more than 10 individuals in each of the subgroups identified in Table 13. Crawford Elementary, Pickney Elementary, and Davis Elementary had at least 10 students in three of the ethnicity subgroups. There are more Caucasian students in each of the nine schools than any other subgroup who are in the CCR and GLP levels on the reading EOG for third grade.

Table 14 illustrates the number of students in each school who were assessed in years 2016-2017 and 2015-2016. The table also shows the percentage of students who meet GLP which include Levels 3, 4 and 5. Additionally, the table shows the percentage

of students that meet CCR which include Levels 4 and 5.

Table 14

Comparison of Third-Grade Students Assessed and Scores 2016-2017 and 2015-2016

School	Number of Students 2016-17	Percentage of Students CCR	Percentage of Students GLP	Number of Students 2015-16	Percentage of Students GLP	Percentage of Students CCR
Meares	132	68.2	78.0	105	82.9	88.6
Gause	130	63.8	76.9	111	62.2	73.0
Rogers	159	77.4	86.2	164	74.4	83.5
Samuels	113	60.2	75.2	119	58.8	67.2
Crawford	140	72.9	78.6	125	69.6	83.2
Pickney	181	58.6	75.7	164	69.5	76.2
Cooper	82	73.2	84.1	80	68.8	75.0
Mapp	121	76.9	86.0	105	77.1	82.9
Davis	160	57.5	75.0	115	67.0	74.8

Table 14 illustrates a comparison of the number of students in each of the nine schools who were assessed in the 2015-2016 school district is less than the 2016-2017 school year except Rogers Elementary. Each year Carter School District projects to have an increase of students. Meares Elementary, Pickney Elementary, Mapp Elementary, and Davis Elementary level 4 and 5 percentages decreased from the 2015-2016 to 2016-2017 school year. Gause Elementary, Samuels Elementary, Cooper Elementary, and Mapp Elementary percentage of GLP students increased from the 2015-2016 to 2016-2017 school year.

Students were qualified for involvement in the study if they met all of the following criteria: (a) enrolled in third grade during the 2015-2016 or 2016-2017 school year, (b) had a TRC score from mClass Reading 3D from both the BOG third grade and EOG assessment from second grade, (c) acquired a score from the reading comprehension portion of the BOG and EOG NC reading assessment, (d) had a TRC score from mClass Reading 3D from the EOY assessment for second grade and (e) had

CogAT scores from second grade.

From the participating schools, 2015-2017 third grade reading scores were used as part of this study. Students for the study were chosen if they had all of the following data points: third grade BOG and EOG scores on the reading assessment and EOY TRC scores from second and third grade, BOG scores from third grade, and CogAT scale scores from second grade. If a student did not have all of the data points, they were omitted from this study. The sample was obtained from records of the accountability departments for the schools involved, and all identifiable indicators were removed from the scores to preserve confidentiality.

Cohort 1. Included in the study were third graders who attended the schools of the participating district during the first year of the study, 2015-2016. The average class size for third grade for the 2015-2016 school year was 23 students. There were approximately 900 students in the first year.

Cohort 2. Included in the study were third graders who attended the schools of the participating districts during the second year of the study, 2016-2017. The average class size for third grade for the 2016-2017 school year was 25 students for the district in this study. There were approximately 900 students analyzed in the second year.

All of the data were retrieved from NCDPI based on each of the represented years above.

Instrumentation

The instruments that were used for this study were the second grade EOY mClass TRC tests, second grade CogAT scores, and the third grade NC BOG and EOG reading comprehension tests. Scores on these tests were gathered from the school district of this study. The collected data were quantitative in nature and based on predetermined questions from the tests as well as data on student performance. Proficiency for BOG,

EOG, and TRC assessment was determined by the state of NC. CogAT scale scores and stanine levels were set by national norms.

TRC assessment. The TRC is a digital assessment that uses running records (RR) to determine a student's reading level. A running record is an instrument that supports teachers to recognize patterns and behaviors in student readings of a text (Fountas & Pinnell, 2009). In a running record, students are directed to read a text and answer oral or written comprehension questions based on the level of the text that is being read. While a student reads a text, the teacher records observations and reading behaviors of the student. Upon successful completion of the running record, a student is administered the comprehension portion of the assessment. The comprehension sections assist teachers with understanding the degree to which the student constructs meaning of the text. TRC has three comprehension components: retelling, oral comprehension, and written comprehension (Amplify, 2013). After reading, students are instructed to retell the text read from text levels E and below. Students have to retell the major components from the beginning, middle, and end of the text. Students have to demonstrate oral comprehension for text levels D and higher. Students have to successfully answer five questions about the text. A student has to demonstrate written comprehension at text level F and higher. To successfully complete this portion, a student has to answer two comprehension questions by writing their response.

The student's accuracy percentage of the running records and the comprehension component(s) are combined to assign the reading level of the student (Amplify, 2013). The instructional reading level is determined by the following criteria: the accuracy percentage is 90-94%, a score of a two or higher on retell, a four on oral comprehension, and a two or higher on written comprehension (Amplify, 2013). Each student's reading score is identified as a letter (A-Z) based on the system created by Fountas and Pinnell

(2010). Grade-level criteria is identified and described as at or above grade-level proficiency, just below grade-level proficiency, or far below grade-level proficiency (mClass Reading 3D, 2010). Based on NC standards, the goal students are expected to achieve at the end of each grade level is as follows: a level M in second grade, a level P in third grade, a level S in fourth grade, and a level U in fifth grade (Amplify, 2013).

> The three main forms of validity to look for are
>
> (a) content validity (do the items measure the content they were intended to measure?), (b) predictive or concurrent validity (do scores predict a criterion measure? Do results correlate with other results?), and (c) construct validity (do items measure hypothetical constructs or concepts?). (Creswell, 2008, p. 190)

The National Center on Response to Intervention determined the reliability and validity of the TRC. Concurrent validity was identified as 0.72, and predictive validity was identified as 0.76. "Reliability refers to the accuracy or precision of a measurement procedure. Internal reliability deals with the consistency of collecting, analyzing and interpreting the data" (Creswell, 2008, p. 192). The marginal reliability was identified as 0.86, and the inter-rater reliability was identified as 0.73 (mClass Reading 3D, 2010). This study seeks to determine the predictive validity between TRC scores and NC EOG scores.

Table 15 represents the text level gradient for elementary school. Text levels are represented by a letter (A-Z) system. Levels A-U correspond with elementary grades, and U-Z represent middle school levels. There is some overlap from end of year testing and beginning of year testing scores (Fountas & Pinnell, 2010).

Table 15

Text Level Gradient

Reading Level	Grade Level Equivalency
A-D	Kindergarten
D-K	Grade 1
K-M	Grade 2
M-P	Grade 3
P-S	Grade 4
S-U	Grade 5

The end of 1 year and the beginning of the next are the same because students are expected to maintain their level over the summer. Second and fifth grade have the fewest levels to grow (two levels). First grade has the greatest number of levels to grow (seven levels).

For the purpose of this study, the letter score on the mClass TRC was converted to a number score for statistical analysis as shown in Table 16.

Table 16

TRC Conversion

TRC Letter	Number Conversion
A	1
B	2
C	3
D	4
E	5
F	6
G	7
H	8
I	9
J	10
K	11
L	12
M	13
N	14
O	15
P	16
Q	17
R	18
S	19
T	20
U	21
V	22
W	23
X	24
Y	25

NC (BOG and EOG) reading comprehension tests. The NC reading comprehension test is administered to students in third grade in August (BOG) and May (EOG). This test is mandated by the state and includes all students unless they are in their first year in U.S. schools (NCDPI, 2014a). The following is a description of the NC EOG reading comprehension test according to the Public Schools of NC website: "The North Carolina End-of-Grade Reading Comprehension Tests measure the goals and objectives as specified in the North Carolina English Language Arts Standards" (NCDPI, 2014a, p. 1). On this assessment, a student reads nonfiction and fiction texts and then answers questions based on their comprehension of the text read.

The NC EOG assessment is deemed reliable and valid as documented by NCDPI. The reliability for the third-grade assessment is 0.925, fourth grade is 0.912, and fifth grade is 0.900 (NCDPI, 2011). The criterion-related validity for third grade is 0.66, fourth grade is 0.63, and fifth grade is 0.61 (NCDPI, 2011).

The tests were measured by five achievement levels: Achievement Level 1, Achievement Level 2, Achievement Level 3, Achievement Level 4, and Achievement Level 5 (NCDPI, 2015). Students scoring at Achievement Levels 1 and 2 are considered performing below grade level, while students scoring at Achievement Levels 3, 4, and 5 are considered to be performing at or above grade-level expectations (NCDPI, 2015).

Table 17 represents the five different achievement levels for third-grade students in reading. Students are considered proficient once they master grade-level reading goals based on the CCSS for NC; and this is denoted at Levels 3, 4, and 5. Mastery is calculated based on responses on informational and literary text. This study used the scaled score data from the NC BOG and EOG reading assessments administered during a student's third grade school year.

Table 17

Third Grade Reading BOG and EOG Levels

Level	Scale Scores
1	≤ 431
2	432-438
3	439-441
4	442-451
5	≥ 452

In Carter School District, the NC EOG is considered a high stakes test. The students who do not pass the NC EOG with a scale score of 439 or higher (Table 17), are required to retake the assessment. The students who do not pass the retest are required to

attend summer school and retake the NC EOG for a second time. Those students who elect to not attend summer school will have to repeat the third grade. The students who attend summer school but do not successfully pass the second retake are placed in a third/fourth grade combination class. The students in the combo class receive intensive reading instruction and will retake the assessment for a fourth time. If a student scale score increases over the retakes or if they pass the assessment determines if a student will remain in the combination class or be moved on to the fourth grade. The acceptable level of growth is determined by each individual school (NCDPI, 2014b).

CogAT. The CogAT is administered to all second-grade students in Carter School District at the beginning of the second semester. CogAT consists of three batteries: verbal, nonverbal, and quantitative. The assessment is comprised of a total of 154 questions (Lohman, 2011). It may be administered in a large or small group or in a one-on-one setting (Lohman, 2011). The test utilized in this study was administered as a whole group unless a student was absent, then it was administered in a small group or one on one.

CogAT scores are reported in the form of Universal Scale Scores (USS; the score used for grade and age norms on a continuous growth scale) and Standard Age Scores (SAS; the score that is used to compare level and rate of student cognitive ability to others students the same age; Lohman, 2012). The USS is a normalized standard score and is the fundamental CogAT scale. USS are considered developmental scores. The SAS scale for each separate battery were developed using smoothed cumulative frequency distributions of USS scores of students at common age levels. The SAS scale can typically range from 50-150. In all age groups, the mean is 100, and the standard deviation is 16 (Riverside Publishing, 2012).

Kuder-Richardson Formula 20, or coefficient alpha reliabilities, average .95 for

the Verbal Battery, .94 for the Quantitative Battery, and .95 for the Nonverbal Battery for both fall and spring administrations of CogAT (Lohman, 2012). Based on Creswell (2008), a coefficient of .9 and greater has excellent reliability and a coefficient of .8 to .9 has a good reliability. The composite scores on CogAT are highly reliable. The three-battery composite reliabilities average .98 for both the fall and spring administrations (Lohman, 2011).

In Table 18, stanine scores range from 1-9. Stanines are groupings of percentile ranks. A higher stanine equates with a higher level of cognitive abilities development. A comparison of stanines and percentile ranks are summarized in the Table 18.

Table 18

CogAT Stanine Categories

Stanine	Percentile Rank	Description
9	96-99	Very High
8	89-95	Above Average
7	77-88	Above Average
6	60-76	Average
5	40-59	Average
4	23-39	Average
3	11-22	Below Average
2	4-10	Below Average
1	1-3	Very Low

Groupings of above average, average, and below average consist of three different stanines. Below average consists of a stanine of 3 or lower and a percentile rank of 22 or lower. The average group encompasses stanines 4-6 and percentiles 23-76. The top grouping is divided into two sections, above average and very high. Above average includes stanines 7 and 8; and the top grouping, very high, is stanine 9 and percentiles 96-99.

This study used the verbal scale score data from the CogAT of students that is administered at the beginning of their second grade school year. The verbal scale score

was used in this study because it corresponds most closely to the other reading assessments (BOG, EOG, and TRC) being analyzed in this study.

Extension Study Details

Bowles (2014) had 143 student participants in Grades 3, 4, and 5. These students were from one public elementary school. The study included all students in these three grades, whereas both test scores were available during the 2010-2011 school year (Bowles, 2014).

Bowles (2014) found a statistically significant and high positive correlation relationship in third, fourth, and fifth grade between the EOG and TRC. Bowles also found that TRC significantly predicted third, fourth, and fifth grade scale scores.

Instead of replicating the study, the researcher decided to extend Bowles (2014) by making revisions to some of the variables and the setting. The differences between this study and Bowles are

1. Bowles was conducted when there were only four NC reading assessment levels, and now there are five levels.
2. Bowles took place at one elementary school, and this study took place at nine elementary schools.
3. Bowles analyzed data from 1 school year, and this study analyzed data over 2 years.
4. Bowles looked at the mClass ORF scores, and this study looked at the CogAT scores.

The first difference between these studies was due to changes the state of NC made with the EOG reading assessment. The second difference was made because the researcher wanted to have a larger sample size than the original study to strengthen the possible data results. The third difference was made because the researcher wanted to

link student potential as measured by the CogAT to achievement scores. The similarities of these studies include elementary school data, analyzing TRC data and EOG data of third-grade students, and using archival data. The studies' research questions are the same, and the analysis is strictly quantitative.

Reliability of data across researchers is essential when developing and measuring an intervention or program efficacy (Kratochwill, Levin, & Horner, 2018). Findings that are positive and negative play an important role in educational system growth. Positive findings provide individuals with sound practices and negative or contradictory results provide individuals with various viewpoints about sampling, methodology, expectations in research, and platforms for advocates to stand on (Kratochwill et al., 2018).

Because many reading research studies utilize small sample sizes, replication studies are important because less than half of replication research finds the same results as the original study (Shanahan, 2016). Shanahan (2016) postulated several reasons why replication research in education had differing findings than the original: pedagogy varies, differences in human behavior, and variable differences. Shanahan noted that when several educational researchers find successful results, the likelihood of result generalization will be strengthened (Shanahan, 2016).

The extension of the Bowles (2014) study potentially provided opportunities for meta-analysis of reading assessments in NC (Shanahan, 2016). This research sought to determine if consistency of findings could be established which would increase the credibility of the Bowles study results that found a strong correlation between the TRC and EOG data for third-grade students.

Research Design

The methodology used for this study was a nonexperimental quantitative approach. This approach is characterized as one that the researcher uses "post positivist

claims for developing knowledge, employs strategies of inquiry such as experiments and surveys, and collects data on predetermined instruments that yield statistical data" (Creswell, 2008, p. 148). Nonexperimental research is illustrated by considering occurrences without providing interventions (Gall, Gall, & Borg, 2007).

The research design was descriptive as well as correlational. Gall et al. (2007) noted that a correlational research design is beneficial in education when studying problems that may exist. Correlational research design provides information about the relationship that exists between variables. Figures, tables, and the Pearson correlation test were used to examine the data of the study and answer the research questions. The Pearson correlation was used in this study to show the strength of the relationship between the student's scores on the mClass assessment, CogAT, and BOG assessments with their scores on the EOG reading comprehension test. This test was used because it took each pair of scores and generated a coefficient that showed the relationship between the data points (Lund Research Ltd., 2013). Predictive validity was later used to determine the strength of the correlations. The Statistical Package for the Social Sciences (SPSS) software program was used to run the Pearson correlation test. The variables that were analyzed were the overall composite scores on the mClass TRC assessment, CogAT verbal battery, and the scale scores on the BOG and EOG reading comprehension test.

Procedures

Carter School District required permission from its local Research Review Panel. The school district used in this study also required the support of a district-level employee who is linked to the primary area of research study. Carter School District did not require data owner consent to access anonymous archival data. Carter School District's panel also required follow-up questioning once the application was submitted and approved to provide clarity to panel members about the research that was being

conducted (see Appendix).

Table 19

Data Reporting Method

CogAT	TRC second	BOG	EOG
Verbal Score	Reading Level Letter	Scale Score	Scale Score

In Table 19, the CogAT scores were reported using a verbal scale score, and the BOG and EOG were reported using a scale score. The TRC assessed in second grade was reported using the reading level letter grade.

Once the study was approved, archival data for this study were collected from test results from the 2015-2017 school years. The data collected were scores from assessments. To collect the data, the researcher contacted the testing and accountability department of the district. The researcher requested the following information:

1. Students who were in second grade in the 2014-2015 school year EOY mClass TRC scores and CogAT scale score and percentile and their 2015-2016 BOG and EOG reading scale scores and levels.
2. Students who were in second grade in the 2015-2016 school year EOY mClass TRC scores and CogAT scale score and percentile and their 2016-2017 BOG and EOG reading scale scores and levels.

The researcher created a database of scores that included the three indexes of the CogAT: verbal, quantitative, and nonverbal reasoning. The verbal composite score was used as an overall measure of student performance, and scores from second grade were collected. The mClass TRC scores from second grade were also collected. The BOG and EOG reading assessment scores from third grade were also collected. The data that were received were coded so there would be no identifiable information that would allow others to identify students.

Once the requested information was retrieved, it was coded to meet the study's specifications. The student information was received from the school district's office of accountability. Data were composed into a document, with one row assigned to each student that encompasses all of the data points. Data were entered from all sources. Student identities were kept confidential by using student identification numbers to represent each student's name.

The data were analyzed to measure the relationship between the TRC, NC reading BOG, and EOG achievement scores. The purpose was to determine the relationship between the three assessments and the degree to which mClass Reading 3D and the NC reading assessment predicted student achievement levels on the reading comprehension part of the NC BOG and EOG.

For each question, Pearson correlations were computed between the variables. A Pearson correlation coefficient was computed to assess the relationship between the TRC and NC reading assessment. Pearson r was used to define the strength of the linear relationship between the two variables through scatterplots and descriptive statistics. The results are illustrated in the form of figures and tables. Predictive validity was discussed based on the results of the Pearson correlation.

Data Analysis

This study sought to determine (a) the relationship between third-grade performance on the BOG reading comprehension test and third-grade performance on the EOG reading comprehension test and (b) the relationship between the end of year second grade TRC and the third-grade performance on the EOG reading comprehension test.

Three variables of interest were studied: third grade BOG scores, third grade EOG scores, and second grade EOY TRC scores. The data analyzed consisted of test scores from each assessment that were obtained from archived files in the research school

district. The predictor variables were mClass TRC and NC BOG scores, and the outcome measure was the NC EOG. Pearson correlations (gender, race, and subgroup), descriptive statistics (frequency and central tendency), and measures of variability (standard deviation) were calculated to determine the strength of the relationship and the degree of predictive validity (multiple regression).

Through the use of Pearson correlations, positive, negative, or no correlation was identified. The strength level of each correlation was identified for the absolute value of R. A very weak correlation was identified as .00-.19, and a weak correlation was identified as .20-.39. A moderate correlation was identified as .40-.59. A strong correlation was .60-.79, and a very strong correlation was identified as .80- 1.0 (Creswell, 2012).

Limitations

"Limitations are potential weaknesses or problems with the study identified by the researcher" (Creswell, 2012, p 199). The test administrators for mClass vary from class to class. Varying administrators is a limitation because there is a time restriction to assess all students, so many teachers are rushed. Data from nine of the 93 schools in this district were analyzed in this study. The sample size of archival data that were collected potentially limited the statistical variability of the study because only 2 years' worth of data were being analyzed, so the outcomes may not be generalizable to other settings or years. The study did not explore if any student within the sample size attended tutoring that could impact their reading comprehension growth. This study did not take into consideration the following factors: teacher attendance, student attendance, instructional delivery, or reading instructional time.

Delimitations

"The delimitations of a study are those characteristics that arise from limitations

in the scope of the study (defining the boundaries) and by the conscious exclusionary and inclusionary decisions made during the development of the study plan" (Simon & Goes, 2013, p. 4). This research study was limited by the use of data that were collected from nine elementary schools in the southwest region of NC, which was a relatively small sample size from the larger district size; the researched school district had a size restriction on research studies. The study did not research academic areas outside of education. This study only researched reading and three assessments, mClass TRC, and the NC EOG reading assessment. Potentially, there were instrumentation issues within the realm of the study. TRC administrators and test training varied from school and district. There is human error that could transpire when assessing students, and this could not be corroborated due to using archival data. The study only compared second and third grade student data. The third grade BOG and EOG are geared toward only one grade level.

Summary

The purpose of this study was to (a) determine the relationship between the scores of the NC BOG and EOG reading assessments and the mClass TRC assessment and (b) determine the relationship and predictability of the TRC and NC BOG scores on the EOG reading assessment by extending the Bowles (2014) study in NC. The participants in this study were third-grade students and archived data were used from two years. The methodology described in this chapter planned for a quantitative design to measure the relationship between two reading assessment. Once data were collected, the results of the research were described and are illustrated in Chapter 4 of this dissertation.

Chapter 4: Results

Introduction

The purpose of this study was to examine the degree to which mClass TRC, NC BOG, and CogAT scores correlate with NC EOG reading comprehension test proficiency levels. This study analyzed archival data over a 2-year period during the 2015-2017 school years. The research questions were answered by gathering archived quantitative data from participant test results. The data collected included assessment scores from third-grade students for the NC reading BOG, EOG, those students' second grade CogAT scores and their EOY second mClass TRC levels. Two years of data were collected and analyzed by statistical tests in order to draw possible conclusions. This chapter details how the data were disaggregated and evaluates the statistics to address the research questions. This chapter concludes with a summary of the research results.

After requesting data from Carter School District, the researcher learned that the CogAT data point was not accessible (and may never be accessible) at the time of the research study request; therefore, Research Questions 3 and 4 were removed and not addressed in this research study. Based on this change, there were four remaining research questions. Due to this shift, this study aligned more closely with the Bowles (2014) study because the CogAt assessment was not used in the original study. The Bowles study and this study analyzed mClass assessments with NC EOG.

Research Questions

This study investigated the predictability of mClass TRC and NC Reading BOG on NC Reading EOG assessment. The research questions for this study were

1. What is the relationship between third grade BOG reading comprehension and third grade EOG reading comprehension?
2. To what extent does the third grade BOG reading comprehension accurately

predict student scores on the third grade EOG reading comprehension?

3. What is the relationship between second grade EOY TRC and third grade EOG reading comprehension?
4. To what extent does the second grade EOY TRC accurately predict student scores on the third grade EOG reading comprehension?

Disaggregation Procedures

Before the data analysis, data were coded and placed into an Excel sheet from the accountability department in Carter School District. The data for each year were entered into the SPSS program.

There were nine elementary schools in each of the data collection years, 2014-2015, 2015-2016, and 2016-2017. Student data were analyzed for only students who had assessment data for each of the assessments: TRC, BOG reading assessment, and EOG reading assessment. Data were analyzed in two sections, Cohort 1 and Cohort 2. Cohort 1 represents a group of students who had second grade TRC data in the 2014-2015 school year and third grade TRC, BOG, and EOG data in the 2015-2016 school year. Cohort 2 represents a group of students who had second grade TRC data in the 2015-2016 school year and third grade TRC, BOG, and EOG data in the 2016-2017 school year. In this study, archival data were analyzed for two different sets of students to increase the sample size and to determine if the same trends and predictability occurred 2 years in a row.

Descriptive and Inferential Statistics

Descriptive statistics were used to summarize data to provide an accurate depiction of results. Descriptive analysis was calculated for Cohort 1 and Cohort 2 for the third-grade students who participated in this study to determine the relationship between second grade TRC, third grade TRC and BOG, and the reading EOG.

Frequency distributions were calculated to determine the frequency of data to determine an accurate depiction of the demographics of participants in this study. Central tendency and variability were calculated to determine scale score range, mean, and standard deviation of the BOG and EOC and text level range of the second and third grade TRC scores.

Inferential statistics were used on the TRC, EOG, and BOG to make generalizations about the data points. The inferential test of significance used in this study equaled .05. Pearson correlation was used to determine the relationship among data points. Through the use of Pearson correlations, a positive, negative, or no correlation was identified. The strength level of each correlation was identified for the absolute value of R. A very weak correlation was identified as .00-.19, and a weak correlation was identified as .20-.39. A moderate correlation was identified as .40-.59. A strong correlation was .60-.79, and a very strong correlation was identified as .80-1.0 (Creswell, 2012). Multiple regression analysis was calculated to determine the extent of predictability of data points. This study used r to represent the multiple correlation coefficient in the Pearson correlation. The notation r^2 is used to notate the proportion of variance in the dependent variable (EOG). R is used to determine the dependent variable quality of the prediction. The Analysis of Variance (ANOVA) was used to determine if the difference among variables is significant.

Cohort 1

Cohort 1 consists of data from students who were from the 2014-2015 and 2015-2016 school years. Within the nine schools in this study, data were received for 1,226 students. After elimination of students who were missing data points, there were 777 students in the Cohort 1 sample. One key factor was that Mapp Elementary did not have any students in this sample size because none of the third-grade students were

administered the TRC End of Year (EOY) assessment. The reason for not administering the TRC at the end of the school year is unknown because it is a district and state expectation that this assessment is administered.

Table 20 illustrates the number of students who were in the third grade for each academic year of the study, 2015-2016. Year 2015-2016 represented the first cohort that was analyzed. This table also illustrates the number of students in each school who meet the analysis criteria.

Table 20

Research Sample Size by School for Cohort 1

School	Total 2015-2016	Sample size 2015-2016
Meares	127	96
Gause	135	86
Rogers	176	133
Samuels	228	87
Crawford	128	92
Pickney	192	135
Cooper	76	53
Mapp	113	0
Davis	130	95

Analysis of data included descriptive and inferential statistics for the nine schools in this study to answer the research questions. The analysis of data in this study was calculated by school and subgroups. The subgroups analyzed were race, EC, LEP, and AIG. Attendance was also analyzed to determine if there was a link between days enrolled and achievement levels.

This study had a representation of seven races. White students made up the majority of the participants, with 64.5%. Pacific Islander had the smallest percentage at 0.3%, which was slightly lower than American Indian (see Table 21).

Table 21

Race Disaggregation for Cohort 1

Race	Frequency	Percent
African American	130	16.7
Hispanic	75	9.7
Asian	46	5.9
White	501	64.5
American Indian	3	.3
Pacific Islander	2	.3
2 or More	20	2.6
Total	777	

Descriptive analysis that includes frequency and percentage of participants in comparison to the Carter School District is in Table 22. The largest sample size in this study was Pickney Elementary, and the smallest participating school sample size was Cooper Elementary.

Table 22

School Disaggregation for Cohort 1

School	Frequency	Percent
Meares	96	12.4
Gause	86	11.1
Rogers	133	17.1
Samuels	87	11.2
Crawford	92	11.8
Pickney	135	17.4
Cooper	53	6.8
Mapp	0	0
Davis	95	12.2

Table 23 consists the gender breakdown of participants in Cohort 1. There was a relatively even distribution of gender representation in this study. Females had a six tenths larger representation in this study.

Table 23

Gender Disaggregation for Cohort 1

Sex	Frequency	Percent
Male	386	49.7
Female	391	50.3

LEP students were also identified in this study. The students identified in this study have taken all assessments and have been in the program for a minimum of 6 months. Table 24 represents the overall number of LEP and non-LEP students in Cohort 1 of this study.

Table 24

LEP Disaggregation for Cohort 1

Category	Frequency	Percent
LEP	53	6.8
Non-LEP	724	93.2

Table 25 illustrates the EC category breakdown which includes non-EC students, SWD, and AIG students. Non-EC students made up the largest sample, and SWD had the smallest number of students identified.

Table 25

EC Category Disaggregation for Cohort 1

Category	Frequency	Percent
Non-EC	519	66.8
SWD	77	9.9
AIG	181	23.3

The TRC second grade EOY assessment range of levels spanned from kindergarten to fifth grade. Table 26 reflects the largest percent of the students' text level was M and N which is on and slightly above grade-level expectation. There was one student who ended the year on a kindergarten level, which was a level C.

Table 26

Second Grade EOY Disaggregation for Cohort 1

Level	Frequency	Percent
C	1	.1
E	15	1.9
F	11	1.4
G	8	1.0
H	17	2.2
I	13	1.7
J	36	4.6
K	40	5.1
L	85	10.9
M	178	22.9
N	103	13.3
O	55	7.1
P	55	7.1
Q	36	4.6
R	58	7.5
S	49	6.3
T	6	.8
U	11	1.4

Students are expected to take the TRC at the end of the year in third grade. This assessment is administered around the same time as the EOG. The text level ranged from first grade to fifth grade reading levels. The second largest frequency was a level Q, which is the beginning of the fourth-grade level. A level U which is fifth grade had the largest number of students, with this score that represented 23.4%.

Table 27

Third Grade EOY Disaggregation for Cohort 1

Level	Frequency	Percent
E	1	0.1
F	2	0.3
G	2	0.3
H	6	0.8
I	3	0.4
J	7	0.9
K	1	0.1
L	10	1.3
M	16	2.1
N	21	2.7
O	48	6.2
P	86	11.1
Q	123	15.8
R	114	14.7
S	113	14.5
T	42	5.4
U	182	23.4

At the beginning of third grade, students in Carter School District take a reading BOG which serves as a baseline. This assessment is administered in a standardized manner similar to the EOG. The data from this study had the majority of the students as proficient, scoring a level 3, 4, and/or 5. The largest frequency was a level 4 that represented 31.5%.

Table 28

BOG Achievement Levels Disaggregation for Cohort 1

Level	Frequency	Percent
1	184	23.7
2	186	23.9
3	114	14.7
4	245	31.5
5	48	6.2

The number of proficient students increased from the BOG to the EOG. Students scoring a level 4 remained the largest achievement level, similar to the BOG. Level 1

had the lowest frequency; and from the BOG to EOG, 134 students moved up to the next achievement level. All proficient levels increased, and level 5 had the largest amount of growth, which was 167 students added to this level.

Table 29

EOG Achievement Levels Disaggregation for Cohort 1

Level	Frequency	Percent	BOG to EOG change
1	50	6.4	-134
2	109	14.0	-77
3	61	7.9	+53
4	342	44.0	+97
5	215	27.7	+167

Table 30 contains descriptive statistics of the EOG and BOG scale scores. There was a 4-point difference between the BOG and EOG minimum scale, score but there was only a 1-point difference between the maximum scale score. The mean score for the BOG is 438, and this score equates to a level 2. The mean score for the EOG is 445, and this score equates to a level 4.

Table 30

BOG and EOG Scale Score Statistics for Cohort 1

Assessment	Minimum	Maximum	Mean	Std Deviation
BOG Scale score	412	461	438.41	9.409
EOG Scale score	416	462	445.53	8.876

Independent Samples *t* Test

An independent *t* test was used to see if there is a difference between gender and EOG scale scores and LEP category and EOG scale scores. The analysis of the variables in this study are identified in Table 31.

Table 31

Gender and LEP Category Disaggregation for Cohort 1

		F	Sig.	t	df	Sign (2-tailed)	Mean Difference	Std Error Difference
EOG Scale score & gender	Equal variance assumed	.010	.920	-1.660	775	.097	-1.056	.636
EOG Scale score & LEP		.093	.761	-4.084	775	.000	-5.106	1.250

The first analysis was conducted on gender. The significance was not smaller than .05; therefore, we accept the hypothesis that these two groups do not have a statistical difference between gender and EOG scores. The significance for gender was .097. The second analysis was conducted on LEP category. The significance was smaller than .05; therefore, we reject the hypothesis that these two groups do not have a statistical difference between LEP category and EOG scores. The significance for LEP was .000.

One-Way ANOVA

One variable with a least two different independent levels are needed to conduct a one-way ANOVA. Table 32 depicts the descriptive difference among race using the EOG scale score data. The between groups analysis has a .000 significance.

Table 32

Gender ANOVA Cohort 1

Race	N	Mean	Std Deviation	Std Error
African American	130	441.93	9.445	.828
Hispanic	75	441.00	9.119	1.053
Asian	46	449.09	7.650	1.128
White	501	446.85	8.182	.366
American Indian	3	440.00	14.526	8.386
Pacific Islander	2	441.00	18.385	13.00
2 or More	20	445.90	9.679	2.164

The mean EOG scale score for students identified as African Americans, Hispanics, American Indians, and Pacific Islanders was a level 3. The mean EOG scale score for students identified as Asians, Whites, and two or more Races was a level 4. American Indian representation had the gender subgroup with the lowest EOG scale score of 440, and Asian students represented the gender subgroup with the highest scale score of 446.

In regard to EC category, the mean score for all categories was 445, which is a level 4 with a standard deviation of 8.876. Table 33 illustrates the mean score for each category and includes standard (Std) deviation and Std error.

Table 33

EC Category ANOVA Cohort 1

Category	N	Mean	Std Deviation	Std Error
Non-EC	519	444.82	7.530	.331
SWD	77	433.79	9.386	1.070
AIG	181	452.55	5.411	.402
Total	777	445.53	8.876	.318

SWD had the largest standard error of 1.070 and the largest standard deviation of 9.386. Academic and Intellectually Gifted students had the smallest standard deviation of 5.411 and non-EC students had the smallest standard error.

The frequency of achievement level scored based on the second grade TRC level is depicted in Table 34.

Table 34

Second Grade Book Level to Achievement Level Frequency Cohort 1

Level	1	2	3	4	5	Total
C	1	0	0	0	0	1
E	7	5	0	3	0	15
F	5	4	1	1	0	11
G	4	3	1	0	0	8
H	6	5	1	4	1	17
I	5	5	3	0	0	13
J	6	10	3	16	1	36
K	5	11	7	12	5	40
L	6	21	7	39	12	85
M	4	32	21	94	27	178
N	4	32	21	94	27	103
O	0	2	4	33	16	55
P	0	5	3	20	27	55
Q	0	0	0	19	17	36
R	1	0	1	26	30	58
S	0	1	0	12	36	49
T	0	0	0	0	6	6
U	0	0	0	1	10	11
	50	109	61	342	215	777

The higher the text level, the more likely a student would achieve a proficient EOG achievement level. There were 81 students who were proficient by text level standards but did not pass the EOG achievement level. Additionally, there were 15 students who were reading on a first-grade level or below who scored a proficient level on the EOG.

The frequency of achievement level scored based on the third grade TRC level is depicted in the Table 35.

Table 35

Third Grade Book Level to Achievement Level Frequency Cohort 1

Level	1	2	3	4	5	Total
E	1	0	0	0	0	1
F	2	0	0	0	0	2
G	2	0	0	0	0	2
H	4	1	1	0	0	6
I	1	1	0	1	0	3
J	6	0	0	1	0	7
K	0	1	0	0	0	1
L	7	3	0	0	0	10
M	8	6	1	1	0	16
N	4	10	2	5	0	21
O	6	13	8	14	7	48
P	3	20	10	48	5	86
Q	4	28	15	64	12	123
R	1	18	14	60	21	114
S	1	6	6	60	40	113
T	0	1	3	23	15	42
U	0	1	1	65	115	182
	50	109	61	342	215	777

The higher the text level, the more likely a student would achieve a proficient EOG achievement level. There were students who were proficient by text level standards but did not pass the EOG achievement level. Additionally, there were students who were at least 2 levels below text level expectations who scored a proficient level on the EOG. Text levels E to I are considered first grade reading levels, and two students in Cohort 1 who had these scores passed the EOG. Text levels J to M are considered second grade reading levels, and three students who had these scores passed the EOG. Students who scored a level Q or higher are reading beyond third grade level, and 60 students with this score did not pass the EOG.

Table 36 illustrates that female students in this study outperformed male students in achievement levels 3 and 5.

Table 36

Gender and EOG Cross Tabulation Cohort 1

Sex	1	2	3	4	5	Total
Male	29	55	27	177	98	386
Female	21	54	34	165	117	391
Total	50	109	61	342	215	777

Males outperformed females in achieving a level 4 on the EOG. There were 48 male students and 75 female students who scored a not proficient level on the EOG. A total of 302 males and 316 females were proficient, scoring a level 3, 4, or 5 on the EOG.

Noted in Table 37, LEP students scored substantially lower than non-LEP students on the reading EOG.

Table 37

LEP Frequency of Achievement Levels Cohort 1

Category	1	2	3	4	5	Total
LEP	6	16	5	19	7	53
Non LEP	44	93	56	323	208	724
Total	50	109	61	342	215	777

There were only 31 LEP students who passed the EOG for reading. Both sets of students had the largest sampling of students score a level 4 on the EOG. The smallest sampling of LEP students scored a level 3, and the smallest sampling of non-LEP students scored a level 1.

Table 38 illustrates the cross tabulation of EC categories and EOG achievement levels.

Table 38

EC Category Frequency of Achievement Levels Cohort 1

	1	2	3	4	5	Total
Non-EC	18	82	54	265	100	519
SWD	32	25	5	11	4	77
AIG	0	2	2	66	111	181
Total	50	109	61	342	215	777

The majority of non-EC and AIG students scored a level 4, and the majority of SWD scored a level 1. Students categorized as AIG did not have a representation in achievement level 1. Non-EC students' smallest sampling was a level 1, and SWD's smallest sampling was a level 5.

Correlation Analysis (Cohort 1)

To answer Research Questions 1 and 3 based on Cohort 1's data, the researcher used Pearson r to determine the strength of the relationship between second grade EOY TRC scores and the third grade reading EOG scores as well as third grade BOG scores and the third grade reading EOG scores.

Pearson r measures the correlation between two variables. Table 39 represents the correlation between assessments in this study.

Table 39

TRC and BOG Pearson Correlation Cohort 1

		EOG Scale Score
Second grade EOY TRC Level	Pearson Correlation	.638
	Significance	.000
	Number	777
Third grade EOY TRC Level	Pearson Correlation	.680
	Significance	.000
	Number	777
BOG	Pearson Correlation	.819
	Significance	.000
	Number	777

To determine the strength of the correlation, the researcher used the following scale (Creswell, 2012):

- .00-.19 very weak
- .20-.39 weak
- .40-.59 moderate
- .60-.79 strong
- .80-1.0 very strong

There is a strong correlation between the second grade EOY TRC and the reading EOG; the correlation is .638, and the significant 2-tailed is at the .000 level. There is a strong correlation between the third grade EOY TRC and the reading EOG; the correlation is .680, and the significant 2-tailed is at the .000 level. There is a very strong correlation between the reading BOG and the reading EOG; the correlation is .819, and the significant 2-tailed is at the .000 level. The TRC assessments both have a strong correlation to the EOG, and the BOG has a very strong correlation to the EOG.

To answer Research Questions 2 and 4 based on Cohort 1's data, the researcher ran a linear regression to determine the extent to which the second grade EOY TRC and third grade BOG scores accurately predict the third grade reading EOG scores.

A prediction equation is based on the correlation of two variables. In order to create this prediction equation, one must first generate the coefficient. A linear regression analysis was conducted in order to generate the coefficients for the variables.

The first prediction equation is for the independent variable of second grade EOY TRC and the dependent variable of the reading EOG. The correlation study generates the number necessary to create this equation depicted in Table 32. The first equation was y=

1.733x + 421.730. The Y, dependent variable, was EOG scale score and the X, independent variable, was the second grade EOY TRC data. If the second grade EOY TRC score is M, which correlates to number 13, the EOG scale score should be 443.829, which was a level 4.

The second equation was y= 2.183x + 406.603. The Y, dependent variable, was EOG scale score and the X, independent variable, was the third grade EOY TRC data. If the second grade EOY TRC score is P, which correlates to number 16, the EOG scale score should be 441.531, which was a level 3. The sample score used in each equation for the TRC was the EOG expectation set in Carter School District.

The third equation was y= .773x + 106.765. The Y, dependent variable, was EOG scale score and the X, independent variable, was the third grade BOG data. If the BOG scale score is 439, which is the lowest scale score for a level 3, the EOG scale score should be 446.112, which was a mid-level 4.

Table 40

Second Grade EOY TRC Linear Regression Coefficients Cohort 1

	Unstandardized Coefficients	
	B	Std. Error
(Constant)	421.730	1.062
2- EOY TRC Level	1.733	.075
(Constant)	406.603	1.526
3- EOY TRC Level	2.183	.085
(Constant)	106.765	8.523
BOG	.773	.019

The first equation was y= 1.733x + 421.730. The Y, dependent variable, was EOG scale score and the X, independent variable, was the second grade EOY TRC data. If the second grade EOY TRC score is M, which correlates to number 13, the EOG scale score should be 443.829, which was a level 4.

The second equation was y= 2.183x + 406.603. The Y, dependent variable, was

EOG scale score and the X, independent variable, was the third grade EOY TRC data. If the second grade EOY TRC score is P, which correlates to number 16, the EOG scale score should be 441.531, which was a level 3. The sample score used in each equation for the TRC was the EOG expectation set in Carter School District.

The third equation was y= .773x + 106.765. The Y, dependent variable, was EOG scale score and the X, independent variable, was the third grade BOG data. If the BOG scale score is 439, which is the lowest scale score for a level 3, the EOG scale score should be 446.112, which was a mid-level 4.

Cohort 2

Cohort 2 consists of data from students who were from the 2015-2016 and 2016-2017 school year. Within the nine schools in this study, data were received for 1,386 students. After elimination of students who were missing data points, there were 861 students in the Cohort 1 sample size. One key factor was that Mapp Elementary only had all three data points for two third-grade students.

Table 41 illustrates the number of students who were in the third grade for each academic year of the study, 2016-2017. Year 2016-2017 represents the second cohort that was analyzed. This table also illustrates the number of students in each school who meet the analysis criteria.

Table 41

Research Sample Size by School for Cohort 2

School	Total 2016-2017	Sample Size 2016-2017
Meares	142	108
Gause	155	106
Rogers	179	122
Samuels	137	82
Crawford	160	107
Pickney	212	149
Cooper	94	61
Mapp	136	2
Davis	171	124

Each school has a smaller sample size than the actual number because each school had students who were missing at least one assessment. Pickney Elementary had the largest total and sample size, and Cooper Elementary has the smallest number of students; but Mapp Elementary had the smallest sample size. Each school had at least 33 fewer students and as large as 134 fewer students in the sample size.

Table 42 illustrates descriptive data for absences, second grade TRC, third grade TRC, and BOG and EOG scale scores.

Table 42

Descriptive Statistics Cohort 2

	Minimum	Maximum	Mean	Std. Deviation
Absences	0	65	6.76	5.74
TRC second Grade	2	21	14.32	3.19
TRC third Grade	5	21	17.71	2.90
BOG Scale Score	411	461	461	9.79
EOG Scale Score	411	461	461	8.98

The maximum number of student absences for Cohort 2 was 65 days, but the mean number of days was 6.76 days. The minimum TRC level for students at the end of second grade was a level B, which equates to the middle of kindergarten level; and the maximum level achieved was a U, which equates to the end of fifth grade in Carter

School District. The second grade end of year (EOY) mean was a level N. In Carter School District, a level N is the reading level of a student at the beginning of third grade. The minimum TRC score attained by a student for the end of the year of third grade was a level E, which is the beginning of first grade. The mean score attained on the EOY third grade TRC was a level Q. Level Q aligns with the beginning of fourth grade. The lowest and highest scale scores attained on the BOG and EOG were the same, 411 and 461.

Frequency descriptive data for 11 of the data points are listed in Tables 43-47. Each data point provides information to gain a better understanding of the participants in this study. In Table 43, Mapp Elementary has the smallest group of participants, and Pickney Elementary has the largest sample group.

Table 43

Schools in Cohort 2

School	Frequency	Percent
Meares	108	12.5
Gause	106	12.3
Rogers	122	14.2
Samuels	82	9.5
Crawford	107	12.4
Pickney	149	17.3
Cooper	61	7.1
Mapp	2	.2
Davis	124	14.4

In Table 43, Mapp Elementary has the smallest group of participants, and Pickney Elementary has the largest sample group.

Table 44

Gender Frequency for Cohort 2

Gender	Frequency	Percent
Male	430	49.9
Female	431	50.1

Table 44 show the gender breakdown for Cohort 2. There was almost the same number of male and female students. There was one more female student in this cohort than male students.

Table 45

LEP Category Cohort 2

Category	Frequency	Percent
LEP	69	8
Non LEP	792	92

Students who are identified as LEP have not attained a passing score on the ACCESS test components. The number of non-LEP students was drastically lower than LEP students. Only 8% of the students in this study in Cohort 2 are identified as LEP.

Table 46

EC Category Disaggregation Cohort 2

Category	Frequency	Percent
Non-EC	580	67.4
SWD	85	9.9
AIG	196	22.8

Cohort 2 had a larger population of students identified as AIG than the first cohort. There were 22.8% of students identified as AIG, which was approximately 13% higher than SWD.

In Table 47, the race of the study participants was broken down into seven subgroups.

Table 47

Race Disaggregation Cohort 2

Race	Frequency	Percent
African American	147	17.1
Hispanic	85	9.9
Asian	65	7.5
White	544	63.2
American Indian	1	.1
Pacific Islander	2	.5
2 or More	15	1.7

The largest group was White students, with 544 students; and the smallest group was American Indian, with only one student. There were only two students identified as Pacific Islander.

End of second grade text levels were also analyzed. Text levels can range from Level A to Level U based on TRC categories. The researcher used a numeric code to calculate means, frequencies, and correlations that aligned with the alphabetic level system used to describe TRC levels.

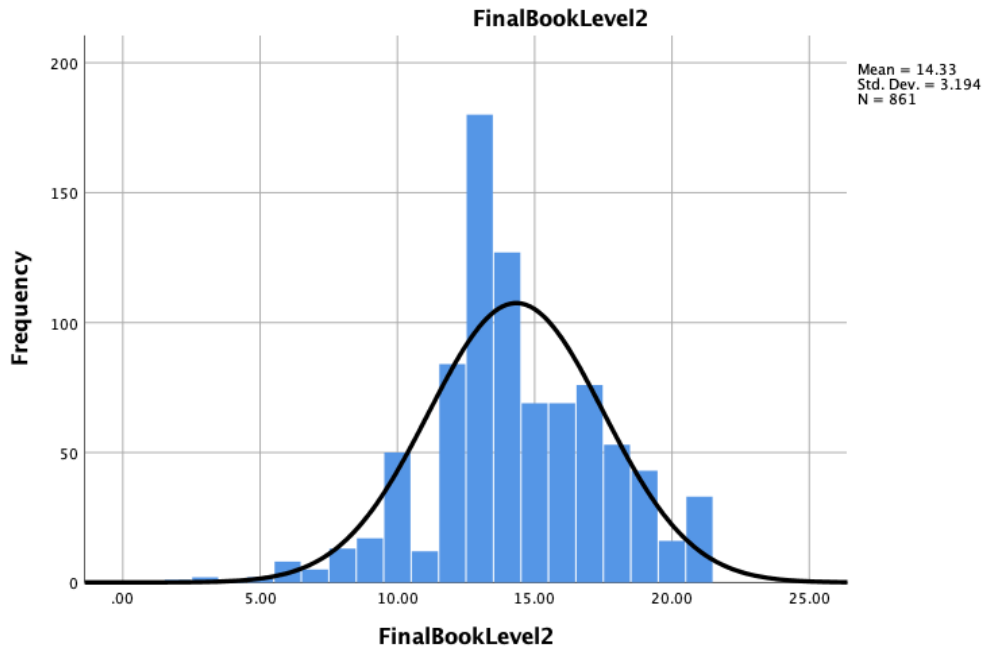


Figure 4. TRC Level for EOY Second Grade Frequency Cohort 2.

The majority of participants scored a Level M (13) which is the EOG proficiency expectation for second-grade students. The mean score is 14.33, which is above grade level expectation. Cohort 2 text level ranges from Level C (3) to Level U (21).

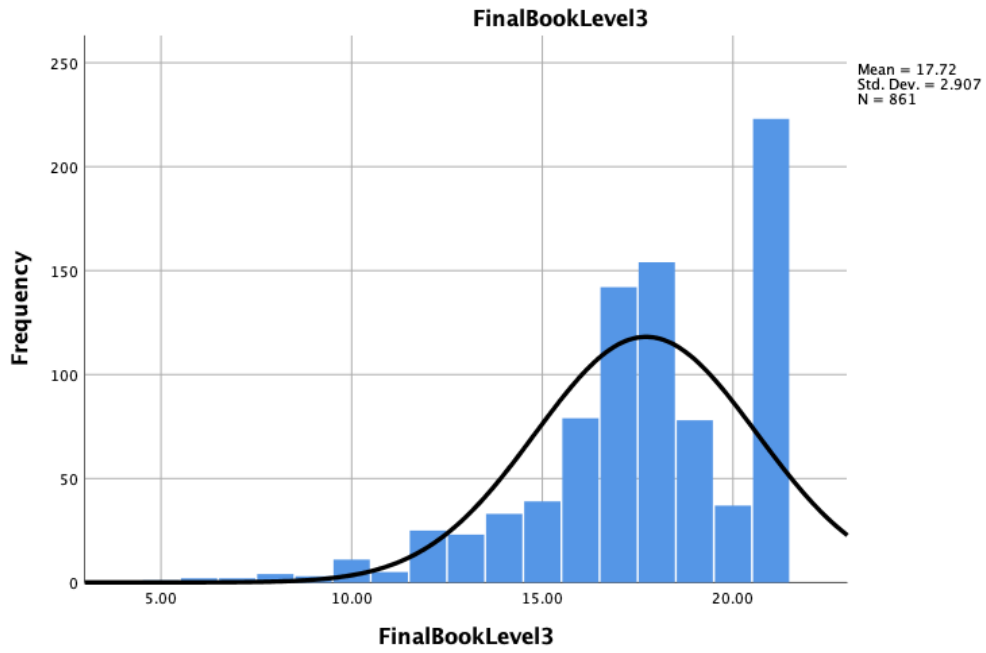


Figure 5. TRC Level for EOY Third Grade Frequency Cohort 2.

The EOG level TRC frequency distribution is in Figure 5. The mean score was 17.72, which is slightly above a level Q. This mean score is one level higher than EOG expectations for third grade students.

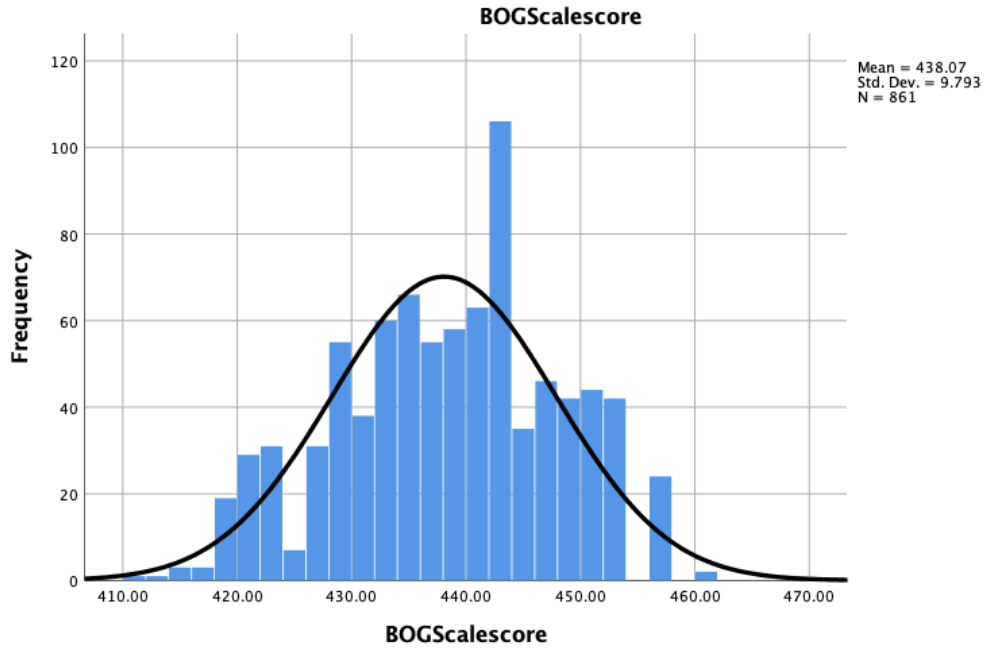


Figure 6. BOG Scale Score Frequency Distribution.

Figure 6 depicts the BOG frequency distribution for Cohort 2 students. Student BOG scale scores range from level 1 (411) to level 5 (461). The maximum level 1 score is 431, and the minimum level 5 score is 452. The mean score is 438.07, which is the highest level 2 score.

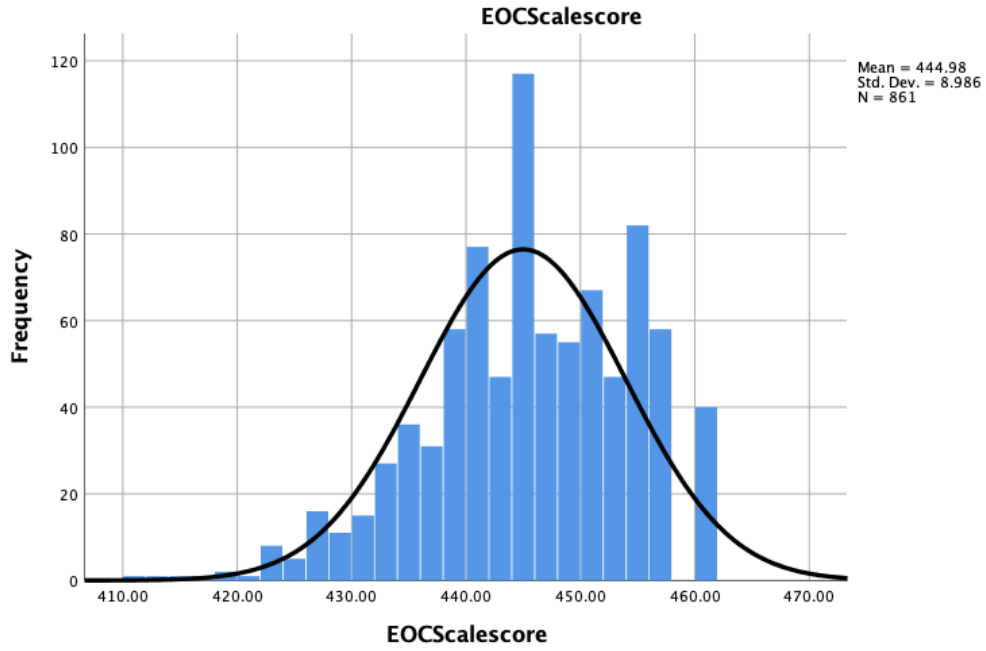


Figure 7. EOG Scale Score Frequency Distribution.

Figure 7 depicts the EOG frequency distribution for Cohort 2 students. Student BOG scale scores range from level 1 (411) to level 5 (461). The maximum level 1 score is 431, and the minimum level 5 score is 452. The mean score is 444.98, which is the highest level 4 score.

Correlation Analysis (Cohort 2)

To answer Research Questions 1 and 3 based on Cohort 2's data, the researcher used Pearson r to determine the strength of the relationship between second grade EOY TRC scores and the third grade reading EOG scores as well as third grade BOG scores and the third grade reading EOG scores

Table 48 indicates the Pearson correlation for all assessments from Cohort 2. To determine the strength of the correlation, the researcher used the following scale (Creswell, 2012):

- .00-.19 very weak

- .20-.39 weak
- .40-.59 moderate
- .60-.79 strong
- .80-1.0 very strong

The statistically significant relationship used is .05, and all assessments had .000 significance.

Table 48

Correlation of All Assessments Cohort 2

		EOG Scale score	TRC Second	TRC Third	BOG Scale Score
EOG Scale score	Pearson Correlation	1			
	Sig. (2-tailed)				
	N	861			
TRC Second	Pearson Correlation	.678**	1		
	Sig. (2-tailed)	.000			
	N	861	861		
TRC Third	Pearson Correlation	.698**	.739**	1	
	Sig. (2-tailed)	000	.000		
	N	861	861	861	
BOG Scale Score	Pearson Correlation	.822**	.699**	.709**	1
	Sig. (2-tailed)	000	.000	.000	
	N	861	861	861	861

** Correlation is significant at the 0.01 level (2-tailed).

All assessments were determined to have a positive correlation; second grade TRC had a .678 strong correlation, TRC third grade had a .698 strong correlation, and BOG scale score had a .822 very strong correlation to EOG scale score.

Table 49

Correlations of Gender (males) and EOG Scale score Cohort 2

		Gender	EOG Scale score
Gender	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	861	
EOG Scale score	Pearson Correlation	.085*	1
	Sig. (2-tailed)	.012	
	N	861	861

*. Correlation is significant at the 0.05 level (2-tailed).

There was a statistically significant and positive correlational relationship between gender (males) and the EOG scale score. This was identified by the very strong positive correlation result of .085, with the correlation coefficient being .012. Very strong correlation scores range from .80 to 1.0.

Table 50

Correlations of Gender (females) and EOG Scale Score Cohort 2

		LEP	EOG Scale score
LEP	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	861	
EOG Scale score	Pearson Correlation	.205*	1
	Sig. (2-tailed)	.000	
	N	861	861

*. Correlation is significant at the 0.015 level (2-tailed).

There was a statistically significant and positive correlational relationship between females and the EOG scale score. This was identified by the weak positive correlation result of .205, with the correlation coefficient being .000.

Table 51

Correlations of EC Category and EOG Scale Score Cohort 2

		Gender	EOG Scale score
EC Category	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	861	
EOG Scale score	Pearson Correlation	.384*	1
	Sig. (2-tailed)	.000	
	N	861	861

*. Correlation is significant at the 0.05 level (2-tailed).

There was a statistically significant and positive correlational relationship between EC category and the EOG scale score. This was identified by the weak positive correlation result of .384 with the correlation coefficient being .000.

Table 52

Correlations of Race and EOG Scale Score Cohort 2

		Gender	EOG Scale score
Race	Pearson Correlation	1	
	Sig. (2-tailed)		
	N	861	
EOG Scale score	Pearson Correlation	.287*	1
	Sig. (2-tailed)	.000	
	N	861	861

*. Correlation is significant at the 0.01 level (2-tailed).

There was a statistically significant and positive correlational relationship between race and the EOG scale score. This was identified by the weak positive correlation result of .287 with the correlation coefficient being .000.

A Pearson correlation by gender was completed to determine if there was a difference in the correlation by gender for each assessment. Table 53 depicts data for males, and Table 54 depicts data for females.

Table 53

Pearson Correlation for Males Cohort 2

		BOG Ach Level	EOG Ach. Level	TRC Second Grade	TRC Third Grade
BOG Ach. Level	Pearson Correlation	1			
	Sig. (2-tailed)				
	Covariance	1.858			
EOG Ach. Level	Pearson Correlation	.766**	1		
	Sig. (2-tailed)	.000			
	Covariance	1.293	1.532		
TRC Second Grade	Pearson Correlation	.643**	.635**	1	
	Sig. (2-tailed)	.000	.000		
	Covariance	2.816	2.526	10.315	
TRC Third Grade	Pearson Correlation	.650**	.676**	.732**	1
	Sig. (2-tailed)	.000	.000	.000	
	Covariance	2.623	2.477	6.957	8.763

**. Correlation is significant at the 0.01 level (2-tailed).

Covariance is the measure of how changes in a variable are linked with changes in another variable. Covariance measures the degree to which variables are associated with each other. Covariance shows a decreasing (negative numbers) or increasing (increasing) linear relationship between two different variables. The sample for gender was 431 students.

Table 54

Pearson Correlation for Females Cohort 2

		BOG Ach. Level	EOG Ach. Level	TRC Second Grade	TRC Third Grade
BOG Ach. Level	Pearson Correlation	1			
	Sig. (2-tailed)				
	Covariance	1.728			
EOG Ach. Level	Pearson Correlation	.756**	1		
	Sig. (2-tailed)	.000			
	Covariance	1.173	1.394		
TRC Second Grade	Pearson Correlation	.678**	.638**	1	
	Sig. (2-tailed)	.000	.000		
	Covariance	2.804	2.369	9.889	
TRC Third Grade	Pearson Correlation	.662**	.676**	.742**	1
	Sig. (2-tailed)	.000	.000	.000	
	Covariance	2.467	2.265	6.615	8.043

**. Correlation is significant at the 0.01 level (2-tailed).

To answer Research Questions 2 and 4 on Cohort 2's data, the researcher ran a linear regression to determine the extent to which the second grade EOY TRC and third grade BOG scores accurately predict the third grade reading EOG scores.

A linear regression was conducted to determine the linear relationship between gender and the independent assessment (TRC second, TRC third, and BOG scale score; [see Table 55]). R is one measure of the predictability of the dependent variable with a range from 0 to 1. The closer the score is to 1, the better the independent variable is to predicting the dependent variable. The r^2 accounted for .708 (70.8%), which denotes a high-quality prediction. The significance was .000. The r^2 denotes the proportion of variance within the dependent variable among independent variables. The adjusted r^2 helps to account for the bias among the independent variables.

Table 55

Regression[b] Model Summary Cohort 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.841[a]	.708	.707	4.86782

a Predictors: (Constant), Gender, TRC third grade, BOG Scale score, TRC second Grade
b Dependent Variable: EOG Scale score

The adjusted r^2 is .707 (70.7%). The closer the value is to 1, the better the fit; this study indicates a fit to the regression model.

Table 56

Multiple Regression for Gender, TRC 2 and 3, BOG Cohort 2

a.	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	421.5793	461.2247	444.9791	7.56077	861
Residual	-19.45177	14.90160	.00000	4.85649	861

a. Dependent Variable: EOG Scale score

In this analysis, the degree to which the dependent variable varies with an independent variable is notated by the unstandardized B. Table 57 illustrates the unstandardized coefficients.

Table 57

Coefficients Cohort 2

Model[a]	Unstandardized B	Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
(Constant)	183.385	10.280		17.839	.000
TRC Third Grade	.534	.092	.173	5.783	.000
BOG Scale score	.565	.026	.615	21.836	.000
TRC Second Grade	.340	.083	.121	4.094	.000
Gender	-.089	.334	-.005	-.268	.789

a Dependent Variable: EOG Scale score

The TRC third grade (.534), BOG scale score (.565), and the TRC second grade were all positive correlations; for each point increase of the independent variable, the dependent variable will increase. Gender in this study was not statistically significant.

An ANOVA test determines the ratio of the variance among groups to the variance within groups. This determines if the difference among variables is statistically significant. The independent variables in this analysis are gender, TRC second and third, and BOG scale score.

Table 58

ANOVA[a] Cohort 2

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	49162.091	4	12290.523	518.681	.000[b]
Residual	20283.533	856	23.696		
Total	69445.624	860			

a Dependent Variable: EOG Scale score

b Predictors: (Constant), Gender, TRC third grade, BOG Scale score, TRC second grade

The analysis denoted there is a statistically significance, F is 518.681 and the

significance is .000.

Research Question 1

Research Question 1 asked, "What is the relationship between third grade BOG reading comprehension and third grade EOG reading comprehension?" Both cohorts and all independent assessment variables based on the inferential and descriptive statistics have consistent positive findings. There was a positive correlation and a significant relationship between NC BOG and NC EOG. Within both cohorts, the BOG had the strongest correlations to the EOG. There is a very strong correlation between the reading BOG and the reading EOG; the correlation is .819 and the significant 2-tailed is at the .000 level. There was also a .638 positive correlation and a .000 level significant relationship between TRC and NC EOG.

Table 59

Pearson Correlation (BOG and EOG) All Schools

School	Correlation	Number
Meares Elementary	.812	204
Gause Elementary	.793	192
Rogers Elementary	.764	255
Samuels Elementary	.818	169
Crawford Elementary	.806	199
Pickney Elementary	.766	284
Cooper Elementary	.788	335
Davis Elementary	.790	219

In addition, each school's data were analyzed to see if each school had the same significance. The table above represents the correlation of all schools with data represented for both cohorts. Table 59 illustrates the results of a Pearson correlation analysis between EOG and BOG scale scores at all schools. Samuels Elementary results revealed the highest significant and positive relationship (.818) among all schools in this study. Rogers Elementary has the lowest correlation, .766. The data for Mapp

Elementary could not be analyzed. At least one of the variables is constant. This data consisted of only two students during the 2-year span.

Research Question 2

Research Question 2 asked, "What extent does the third grade BOG reading comprehension accurately predict student scores on the third grade EOG reading comprehension?"

Cohort 1 and 2 analysis illustrated that the independent assessments showed that they were statistically significant when predicting the EOG reading scale score. In addition, each school's data were analyzed to see if each school had the same significance. The strength of each school ranged from strong to very strong.

Table 60

ANOVA and Standardized Coefficient Summary

School	F	Beta
Meares Elementary	391.720	.812
Gause Elementary	439.535	.838
Rogers Elementary	354.879	.764
Samuels Elementary	337.512	.818
Crawford Elementary	365.465	.806
Pickney Elementary	401.294	.766
Cooper Elementary	345.788	.785
Davis Elementary	360.944	.790

Table 60 does not illustrate information for Mapp Elementary because the data statistics could not be computed because the sample size was not large enough to run an accurate analysis. Gause Elementary illustrates that the overall model justifies a significant proportion of variance or that the overall model is statistically significant. The independent variable had a significant effect on the EOG scale score, $F(1, 190) = 439.535$, $p < .001$. Gause Elementary coefficients show that BOG scale scores are linked to higher EOG scale scores (Beta = .838, $p < .001$) and were positively correlated,

signifying that higher levels on the BOG are linked with higher levels of EOG performance.

Research Question 3

Research Question 3 asked, "What is the relationship between second grade EOY TRC and third grade EOG reading comprehension?"

Both cohorts were analyzed together to determine the correlation for all participants in this study. A Pearson correlation analysis was conducted to examine the relationship between second and third grade end of year (EOY) TRC and EOG scale scores of all schools. There is a strong correlation between the second grade EOY TRC and the reading EOG; the correlation is .638, and the significant 2-tailed is at the .000 level. Additional information, there is a strong correlation between the third EOY TRC; the correlation is .680, and the significant 2-tailed is at the .000 level. Table 61 illustrates these data.

Table 61

Nine School Pearson Correlation

School	Second TRC	Third TRC	Number
Meares Elementary	.693	.771	204
Gause Elementary	.633	.667	192
Rogers Elementary	.586	.710	255
Samuels Elementary	.723	.708	169
Crawford Elementary	.633	.667	199
Pickney Elementary	.646	.677	284
Cooper Elementary	.623	.657	135
Davis Elementary	.715	.780	219

**. Correlation is significant at the 0.01 level (2-tailed).

Research Question 4

Research Question 4 asked, "What extent does the second grade EOY TRC accurately predict student scores on the third grade EOG reading comprehension?"

Table 62

Regression Summary Both Cohorts

	R	R Square	Adjusted R Square	Std. Error of the Estimate
Meares	.780[a]	.609	.605	5.694
Gause	.727[a]	.528	.523	6.298
Rogers	.716[a]	.512	.508	5.776
Samuels	.769[a]	.592	.587	5.550
Crawford	.727[a]	.528	.523	6.298
Pickney	.729[a]	.522	.517	5.330
Cooper	.720[a]	.538	.533	6.338
Davis	.790[a]	.624	.620	5.503

a. Predictors: (Constant), TRC third and TRC second

There was a positive correlation and relationship between TRC and NC EOG scores. The R-square is the proportion of variation in the dependent variable that is described by the independent variable. It is conveyed as a percentage. Rogers Elementary's r squared model illustrated 51.2% of the variation in overall performance can be explained by the independent variable (EOG achievement level) in the model. Rogers Elementary had the lowest relationship between assessment scores. Davis Elementary's data illustrated 62.4% of the variation in overall performance can be explained by the independent variable (EOG achievement level) in the model which was the highest relationship among all nine schools.

Summary

Participants in both cohorts had consistent descriptive and inferential findings. There was a positive correlation between TRC and NC BOG to NC EOG scores at the cohort and individual school levels. The NC BOG had a stronger relationship to the NC EOG than the TRC assessment. This provides the data to answer two research questions:

"What is the relationship between the TRC assessment and the NC EOG," and "What is the relationship between the NC BOG assessment and the NC EOG?"

Cohort 1 and 2 analysis illustrated that the independent assessments showed that they were statistically significant when predicting the EOG reading scale score. In addition, each school's data were analyzed to see if the same significance existed among schools and cohorts. The BOG assessment had a stronger predictability strength than the TRC to the NC EOG. The strength of each school ranged from strong to very strong. This provides the information to answer two research questions: "To what extent does the TRC predict student scores on the NC EOG," and "To what extent does the NC BOG predict student scores on the NC EOG? The TRC assessments both have a strong correlation to the EOG, and the BOG has a very strong correlation to the EOG. Chapter 5 elaborates on the analysis of data, research findings, and implications. Additionally, the connection to literature and future research recommendations are discussed.

Chapter 5: Discussion and Conclusions

Introduction

One key way to impact student trajectory is make data-driven decisions (Nonte, Hartwich, & Williems, 2018). Assessments used in this research can provide educators with data to alter instruction. Meeting the needs of students through classroom instruction or school intervention will increase student reading achievement, which will in turn increase school and student proficiency in reading. Nonte et al. (2018) noted that educators are striving to improve reading competencies. In order to successfully navigate school, students must develop their reading comprehension skills (Nonte et al., 2018). Chapter 5 elaborates on the analysis of data and research findings. Additionally, the connection to literature and future research are discussed.

Results and Connection to the Bowles (2014) Study

The research design for this study was an extension of Bowles's (2014) dissertation, *The Relationship between mClass Reading 3D Assessment and the North Carolina End of Grade Assessment of Reading Comprehension in an Elementary School*. Bowles researched the relationship between scores of the NC EOG reading comprehension assessment; the scores from the mClass Reading 3D assessment in third, fourth, and fifth grades; and the degree to which mClass Reading 3D predicted the reading NC EOG scores. This study replicated portions of the Bowles study to determine if the results are transferable to a different school district with both similar and different demographics.

Bowles (2014) found a statistically significant relationship and positive correlation between NC EOG and TRC which is identical to the findings of this study. Bowles found that TRC significantly predicted student achievement on the NC EOG assessment which is also supported by the findings of this study.

Study Summary

The purpose of this study was to examine the degree to which mClass TRC and NC BOG scores correlate with NC EOG reading comprehension test proficiency levels. This study analyzed archival data over a 2-year period during the 2015-2017 school years. This study measured the predicative implications of mClass TRC and NC Reading BOG on NC Reading EOG assessment.

The research questions for this study were

1. What is the relationship between third grade BOG reading comprehension and third grade EOG reading comprehension?
2. To what extent does the third grade BOG reading comprehension accurately predict student scores on the third grade EOG reading comprehension?
3. What is the relationship between second grade EOY TRC and third grade EOG reading comprehension?
4. To what extent does the second grade EOY TRC accurately predict student scores on the third grade EOG reading comprehension?

Quantitative data were collected and analyzed for two cohorts. Cohort 1 consisted of students who had second grade TRC data in 2014-2015 and third grade TRC, BOG, and EOG data in 2015-1016. Cohort 2 consisted of students who had second grade TRC data in 2015-2016 and third grade TRC, BOG, and EOG data in 2016-2017. Descriptive statistics were used to gain an accurate depiction of demographics and assessment data of the participants in this study; descriptive statistics were conducted. These data points provided a background and basis for the determination that there is a relationship between the independent variables (TRC and BOG) and the dependent variable (EOG). After the data relationship was determined, multiple regression analyses were calculated to determine if the independent variables significantly predicted the

dependent variable. The data determined that there was a statistical significance among NC reading assessments.

Pearson correlations were analyzed to determine the relationship among TRC and BOG reading and EOG reading assessments. There was a significant relationship and positive correlation between TRC, BOG, and the NC EOG assessment. A regression analysis was analyzed to determine if TRC and BOG predicts student achievement on the reading EOG. The findings of Cohort 1 and Cohort 2 both found that there is a significant predictability between the three NC assessments (TRC, BOG, and EOG) analyzed in this study.

Discussion of Findings

Race was analyzed as a variable in this research to determine the relationship between NC assessments in regard to this subgroup. There was a weak relationship between race and scores on the EOG; however, there was a difference in the overall scores among race. Although a specific race did not significantly predict TRC and NC EOG scores, the race to EOG relationship was a weak correlation of .287. Walkington, Clinton and Shivrai (2018) found that there is a difference between race on assessment.

Gender was analyzed as a variable in this research to determine the relationship between NC assessments in regard to this subgroup. Females outperformed males on TRC, BOG, and NC EOG assessments; and there was a strong correlation among these variables when controlling for gender. The BOG to EOG correlation for males was .766, and TRC to EOG correlation for males was .635. The BOG to EOG correlation for females was .756 and TRC to EOG correlation for females was .638. Similar to these findings, Zehner, Goldhammer, and Salzer (2018) found that the gender gap is prevalent in large scale reading assessments. Girls typically outperformed male students. Girls responded better on higher level questions, and they could divide their attention to a task

better than boys. Nonte et al. (2018) found that there was a difference in gender scores in four countries. Girls typically outperform boys, and their attitude towards reading was more positive as well.

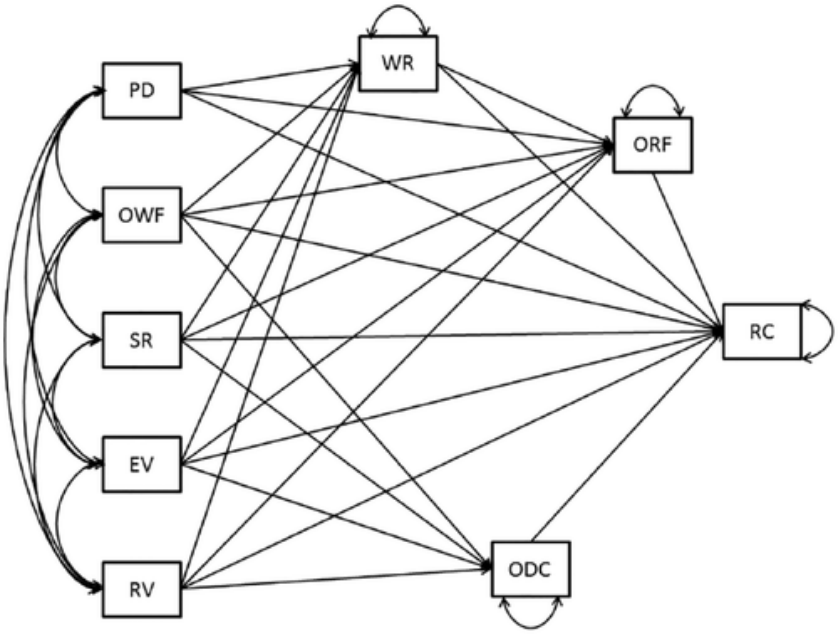
EC and LEP categories were analyzed in this research to determine the relationship between NC assessments in regard to these subgroups. EC and LEP student data points were lower than other students. The EC to EOG had a weak correlation and was noted at .384. The LEP to EOG had a weak correlation of .205. Cortiella and Horowitz (2014) noted that it is hard to identify students with reading difficulties in one subgroup to analyze data because the degree of reading difficulty will impact performance differently. In special education, there are 14 areas of disabilities, and each would present a different profile on reading assessments. Cortiella and Horowitz also noted that analyzing data of typically developing peers is difficult to compare and teach as a whole because of the vast range of factors impacting reading comprehension. Cortiella and Horowitz noted that in order to accurately identify trends and patterns in data, assessments should be administered at the same time to rule out biases, test administration discrepancies, and assessment administration.

Implications for Educators

The educational bar of excellence in education is being raised daily. Through the use of data analysis, teacher and student accountability will continue to be at the focal point of education and classroom instruction. Data in this study described the correlation between TRC and the EOG, high stakes assessments in reading. These data illustrated the importance of building student reading abilities as measured by TRC reading levels.

Figure 8 illustrates the combination of deficits that could potential impact reading comprehension. Each web line represents an area and/or skill that teachers need to teach to improve student reading comprehension (Parkin, 2018). The web represents various

pathways to reading comprehension. In order to improve reading comprehension (RC on the web), a teacher will need to teach to areas of weakness in all the lines and pathways that are directly and indirectly linked. One area of weakness (area on the web) may lead to a trickle-down effect and a teacher will need to have teaching points to address these areas. In order to improve EOG scores as linked to this research a student's TRC reading comprehension level will need to improve as well.



Note. PD = Pseudo word Decoding; OWF = Oral Word Fluency; SR = Sentence Repetition; EV = Expressive Vocabulary; RV = Receptive Vocabulary; WR = Word Reading; ORF = Oral Reading Fluency; RC = Reading Comprehension; ODC = Oral Discourse Comprehension

Figure 8. Initial Path Model.

Implications for classroom teachers. Based on these findings, the better a student's level of reading comprehension, the better their chances are to succeed on the EOG. These data also supported the use of student intervention to meet the needs of individual students because one may have a level E and M in the same class; and in order to improve the student's odds of "reading success," their individual needs must be met.

Weikert (2018) noted that students who were provided reading intervention made 2-year reading level growth. Weikert noted that human intervention had the strongest impact on student intervention achievement. Although the majority of the students who passed the NC EOG had text levels that were indicative of those who have mastered third grade content, there were a few exceptions to the rule. Based on these findings, text Levels E to I are considered first grade reading levels, and two students who had these scores passed the EOG. Text Levels J to M are considered second grade reading levels, and three students who had these scores passed the EOG.

NC's RtA legislation aims to have students reading proficiently by the end of third grade. This study adds to this premise because it determined that there is a link to high stakes assessment. The RtA legislation stands on the notion of intervention and meeting student needs; and this research found a positive, strong correlation between TRC assessment data and EOG achievement. A balanced literacy approach should be utilized by classroom teachers to help address student deficits through the five components of balanced literacy: shared reading, shared writing, reading groups, writing groups, and word work. Perkins Greene (2014) found that students at schools that implemented balanced literacy outperformed schools that were not implementing balanced literacy instruction or only those schools partially implementing balanced literacy in classrooms. Perkins Green (2014) found that there are phases to balanced literacy implementation, and schools that were in phase 3 outperformed those in phase 2 and phase 1 as measured by TRC reading scores. A classroom that utilizes balanced literacy can provide students with opportunities to meet students' needs through each of the five components. Gradual release of responsibility (GRR) provides specific instruction to a student moving from heavy teacher support to student independence (Fountas & Pinnell, 2009). Each component has a GRR so students receive varying

levels of instructional support based on their needs (Fountas & Pinnell, 2009). Often, the GRR model includes teacher modeling (I Do It), guided practice (We Do It), and collaborative practice (You Do It Together) and ends with independent practice (You Do It). In the teacher modeling portion, a teacher models a teaching strategy for students. The guided practice component allows students to work through the same strategy but with teacher scaffolding. Before moving to independent practice, students have added support of their classmates to practice what they have learned through collaborative practice. Finally, the independent practice component provides teachers with opportunities to see if a student has mastered the strategy or if they need additional support. After the three portions are completed, the students have time to work on a task independently and then the teacher can provide small group support. Through the utilization of balanced literacy, classroom teachers can address deficits in phonics, vocabulary, fluency, and comprehension. These additional supports will help students improve their reading skills and, based on these findings, improve their likelihood of passing the EOG.

The data from this research can have a positive impact classroom instruction. These findings should provide teachers with a motivation to differentiate instruction to meet the needs of students based on reading levels through daily instruction. The recommendations are below.

1. Teachers can use the TRC scores to differentiate classroom instruction and instructional groups to address reading comprehension deficits by subgroups. Teachers can divide the class into groups based on their TRC level and provide weekly support to each group to make sure all students receive specialized instruction to ensure growth. Instead of ability grouping, teachers can determine the skills (fluency, phonics, or comprehension) that students are

missing to propel them to the next text level and create strategy groups to address these needs (Fountas & Pinnell, 2010).

2. Within the classroom, teachers can tailor whole class teaching points to address concerns based on particular skills and standards as identified by BOG data. Teachers can analyze common assessment data to pinpoint concepts that most students or small groups of students need additional support. Once an assessment is given, a teacher can analyze questions by curriculum standards to identify which standards need to be remediated or enriched and teach these concepts to the class.

Implications for grade level/data teams. Based on these findings, information should be shared with grade-level and data teams to expound upon the impact that each assessment has on the other. At a whole school staff meeting, teams can share where each grade level ended as far as proficiency levels and areas of need. Teams can also use data to design mini professional development sessions for teachers based on student and teacher needs and specific skill deficits. After assessments are given, student data can be analyzed as well as teacher data to identify areas of strength and need in teaching and provide professional development to teachers based on areas of teaching and scaffolding need. Teams can also utilize these data to provide classroom modeling and coaching dependent upon teacher's areas of strength and need. Based on the data, weekly real time teacher coaching can be provided to teachers to support them as they process through learning a new strategy, and this will provide teachers with feedback in the moment to ensure success.

The data from this research can positively impact grade-level data. These findings could provide grade-level/data teams with a desire to shift team data and instruction to meet the needs of students as a collective. According to Visible Learning

(2016), collective teacher efficacy is the collective belief in a teacher's ability to positively affect student outcomes. Student achievement is strongly correlated with collective teacher efficacy, with an effect size of d=1.57 (Visible Learning, 2016). The recommendations are below.

1. Teams can use the data to differentiate grade-level and team instruction by providing a team and/or school intervention time to address reading deficits on a larger scale to utilize the strengths of all teachers on the team. Based on the data received, each grade level can tailor information according to the needs specifically for each grade. Cybulski (2003) noted, "collective efficacy of teachers was found to have a positive, direct effect on student achievement" (p. iii). By implementing a whole grade-level intervention, teachers can utilize their areas of strengths to meet student needs according to these strengths.
2. Teams can use common assessment data to determine the length of time to spend on a topic based on collective student needs. If one grade has a greater need in reading, they can increase the reading time as opposed to a different grade that has a greater need for writing.

Implications for administrators. Based on these findings, administrators can allocate funding for programs and staffing to address student needs. Administrators can also use this information to inform staffing decisions and candidate qualifications based on team and student needs. This information could also be used to plan for professional development courses needed at the school level. The recommendations are below.

1. Administrators can use school data to allocate funding for hiring, professional development, and text book/program purchases to address the needs on each team and the school as a whole. Based on student and teacher needs, funding

can be set aside to attend conferences to learn new topics and funding can be used to buy supplies or books to support the school with new topics.

2. Administrators can use these data to identify next steps for teachers through the use of classroom walk-throughs and the use and implementation of small groups, because this can remind teachers and staff members of a whole school and grade-level need. Based on observations, feedback can be tailored to teachers and grade levels to support teaching best practices to meet student needs.
3. Administrators can use these data to create a common planning and intervention time as a part of the master schedule. In order for teachers and grade levels to effectively and collectively analyze data, there needs to be time set aside for this collaboration (Dufour, Dufour, Eaker, & Many, 2006).

Contributions. This longitudinal research study investigated student TRC, BOG, and EOG scores over 2 school years. This research adds to the body of research in regard to various reading assessments that are emphasized in NC and the correlations among them. This research also adds to the body of knowledge that already exists among TRC and EOG correlations (Bowles, 2014).

Limitations. This research did find a correlation among assessment scores, but there are limitations that should be noted. The schools in this study were a representation of all the schools in the school district. The research results were from one school district in NC, creating a small sample in terms of the state; therefore, the results may not be generalizable throughout the state. TRC second grade data were taken at least a year before the EOG was given and do not take into consideration summer educational loss or general educational fatigue, because these assessments are taken at the end of the school year.

Additional extraneous factors not controlled by the researcher that could have impacted research results are: (a) teachers who conduct one-on-one assessments (TRC) may not reflect all assessment data correctly; however, all teachers must go through an assessment training; (b) the EOG assessment is approximately four hours long and does not take into consideration if a student worked through all questions; however, a test administrator and proctor are in the room to keep students working to their best ability; (c) student work effort and school performance and support systems differ within a school and classroom, which can effect assessment results.

Recommendations for Future Research

This study on the predictability and relationship between TRC and BOG was an average size scale study with limited generalizability. The schools in this study were selected because they each had subgroups in the LEP category, AIG category, and EC category. The data from this research can positively impact classroom instruction. Based on the results from the study several recommendations for additional research were developed. The recommendations are below.

1. Future studies that analyze the impact of CogAT on student achievement as measured by TRC and EOG data.
2. Future replication of this study across all schools within a district to include Title 1 and non-Title 1 schools to increase generalizability of research findings.
3. Future replication of this study across all schools within different school districts to increase generalizability of research findings.
4. Future longitudinal studies that assess EOG scale scores across grade levels to determine if the prediction of student schools change over time.
5. Future studies that analyze the impact of TRC intervention on NC EOG scale

scores.

6. Future studies that analyze the impact of teacher experience on student achievement as measured by TRC and EOG data.

Future research can continue to have an impact on the trajectory of student success as measured by assessment data, because researchers can continue to push the status quo in education through the use of data. Through the integration of reading strategies and classroom instruction based on data-driven decisions, student reading proficiency will continue to increase.

Conclusion

The purpose of this research was to examine predictability of NC assessment. This study found a correlation between NC reading assessments as measured by NC EOG, NC BOG, and TRC. These findings are consistent with previous research conducted on EOG and TRC assessments. Data revealed that the higher the reading comprehension levels in second grade, the higher the projected BOG and EOG reading scaled scores.

The results of this study have a profound impact on NC literacy development and assessments, because the findings are consistent with other research suggesting the subgroup scores are a direct indicator of high stakes EOG assessments. These findings also add to the findings that TRC significantly predicts EOG scores. This research serves as a basis for research that was conducted to determine the predictability of the NC BOG as a measure for EOG reading scores. Results from this study support the need to use the NC BOG reading assessment as an ideal indicator for third-grade pacing and data-driven instruction. This assessment can address possible teaching points to cover based on the assessment standard breakdown. The results of this study should be used by educators to evaluate current instructional practices to make sure students' instructional needs are

being met in classrooms. The results of this study illustrated the importance of increasing student reading levels in second and third grade because they are both predictors of student success on high stakes assessments such as the NC reading EOG.

References

Amplify. (2013). Validity evidence for mClass: Reading 3D and student performance on the 2012-2013 North Carolina end of grade reading comprehension test. Retrieved from http://www.ncpublicschools.org/docs/superintendents/messages/2014/0217/addendum-attachment02.pdf

Anderman, E. M., & Sinatra, G. M. (2009). *The challenges of teaching and learning about science in the 21st century: Exploring the abilities and constraints of adolescent learners.* Paper prepared for the Workshop on Exploring the Intersection of Science Education and the Development of 21st Century Skills, National Research Council. Retrieved from www7.nationalacademies.org/bose/AndersonSinatraPaper.pdf

Annie E. Casey Foundation. (2013). National study shows students who don't read well in third grade are more likely to drop out or fail to finish high school. Retrieved from http://www.aecf.org/Newsroom/NewsReleases/HTML/2011Releases/~/media/PDFFiles/Newsroom/DoubleJeopardyPressrelease47.pdf

Badger, J., & Christmann, E. (2009). Designing elementary instruction and assessment: Using the cognitive domain. Thousand Oaks, CA: Corwin.

Battle, J., & Linville, D. (2006). Race, sexuality and schools: A quantitative assessment of intersectionality. *Race, Gender & Class, 13*(3), 180-188, 190-191, 194-199.

Bennett, A., Bridglall, B., Cauce, A., Everson, H., Gordon, E., Lee, C., … Stewart, J. (2004). *All students reaching the top: Strategies for closing academic achievement gaps. A report of the national study group for the affirmative development of academic ability.* Naperville, IL: North Central Regional Educational Laboratory.

Bond, G. L., & Dykstra, R. (1997). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 32,* 348-427. (Original work published 1967).

Bowles, A. S. (2014). The relationship between mClass reading 3D assessment and the North Carolina end-of-grade assessment of reading comprehension in an elementary school. *Education Dissertations and Projects,* 6. Retrieved from https://digitalcommons.gardner-webb.edu/education_etd/6

Calkins, L. (1994). *The art of teaching writing*. Portsmouth, NH: Heinemann.

Calkins, L. (2011). *A curricular plan for the reading workshop, grade 2*. Portsmouth, NH: Heinemann.

Chin, T., & Phillips, M. (2004). Social reproduction and child-rearing practices: Social class, children's agency, and the summer activity gap. *Sociology of Education*, *77*(3), 185-210.

Clay, M. (2001). *Change over time in children's literacy development.* Portsmouth, NH: Heinemann.

Clay, M. (2013). *An observation survey of early literacy achievement.* Portsmouth, NH: Heinemann.

Clay, M. (2015). *Becoming literate.* Portsmouth, NH: Heinemann.

Cognitive Abilities Test Form 7 Research and Development Guide. (2012). Rolling Hills Meadow, IL: Riverside Publishing Company.

Cortiella, C., & Horowitz, S. (2014). *The state of learning disabilities: Facts, trends and emerging issues.* New York, NY: National Center for Learning Disabilities.

Coyne, M., Kame'enui, E., & Carnine, D. (2011). *Effective teaching strategies that accommodate diverse learners.* Upper Saddle River, NJ: Pearson.

Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches* (XX ed.). Thousand Oaks, CA: Sage.

Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson: Merrill Prentice Hall.

Cybulski, T. (2003). *Investigating the role of the collective efficacy of teachers in fiscal efficiency and student achievement.* (Electronic Thesis or Dissertation). Retrieved from https://etd.ohiolink.edu/

Decker, D. M., & Bolt, S. E. (2008). Challenges and opportunities for promoting student achievement through large-scale assessment results: research, reflections, and future directions. *Assessment for Effective Intervention, 34*(1), 43-51.

Dooley, C., & Matthews, M. (2009). Emergent comprehension: Understanding comprehension development among young literacy learners. *Journal of Early Childhood Literacy, 9*(3), 269-294.

DuFour, R., DuFour, R., Eaker, R., & Many, T. (2006). *Learning by doing: A handbook for professional learning communities at work*. Bloomington, IN: Solution Tree.

Duibhir, P., & Cummins, J. (2012). *Towards an integrated language curriculum for primary schools (3-12 years)* (Report No. 16). National Council for Curriculum and Assessment (NCCA). Dublin.

Eccles, J., Midgley, C., & Adler, T. (1984). *Grade-related changes in the school environment: Effects on achievement motivation.* Greenwich, CT: JAI Press.

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

Fountas, I., & Pinnell, G. (2001). *Guiding readers and writers.* Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (2006). *Guided reading.* Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (2009). *Guiding readers and writers (grades 3-6) teaching, comprehension, genre, and content literacy.* Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (2010). *The continuum of literacy learning, grades preK-8.* Portsmouth, NH: Heinemann.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Boston, MA: Pearson Education, Inc.

Gottfredson, L., & Saklofske, D.H. (2009). Intelligence: Foundations and issues in assessment. *Canadian Psychology/Psychologie Canadienne, 50*(3), 183-195.

Graham, A., & Erwin, K. (2011). “I don't think black men teach because how they get treated as students": High-achieving African American boys' perceptions of teaching as a career option. *Journal of Negro Education, 80*(3), 398-416.

Henshon, S. E. (2014). A big-picture thinker illuminates identification and ability: An interview with David F. Lohman. *Roeper Review, 36*(1), 3-10.

Hoffman, K., Llagas, C., & Snyder, T. (2003). Status and trends in the education of blacks. U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubs2003/2003034.pdf

Huba, M., & Freed, J. (2000). *Learner centered assessment on college campuses: Shifting the focus from tracking to learning.* Needham Heights, MA: Allyn & Bacon.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581-592.

International Literacy Association. (2007). *The case for children's rights to read*. Newark, DE.

International Literacy Association (formerly the International Reading Association). (2014). *Using high-stakes assessments for grade retention and graduation decisions* [Position statement]. Newark, DE.

Kennedy, E., Dunphy, E., Dwyer, B., Hayes, G., McPhillips, T., Marsh, J., & Shiel, G. (2012). *Literacy in early childhood and primary education (3-8 years).* (Report No. 15). National Council for Curriculum and Assessment (NCCA). Dublin.

Kerr, M., & Frese, K. (2017). Reading to learn or learning to read? Engaging college students in course readings. *College Teaching, 65*(1), 28-31.

Lemov, D. (2010). *Teach like a champion: 49 techniques that put students on the path to college*. Indianapolis, IN: Jossey-Bass.

Lohman, D. F. (2006). Belief about difference between ability and accomplishment: From folk theories to cognitive science. *Roeper Review, 29*(1), 32-40.

Lohman, D. F. (2011). *Cognitive abilities test, form 7.* Rolling Meadows, IL: Riverside Publishing.

Lohman, D. F. (2012). *CogAT score interpretation guide.* Rolling Meadows, IL: Riverside Publishing.

Lohman, D. F., Gambrell, J., & Lakin, J. (2008). The commonality of extreme discrepancies in the ability profiles of academically gifted students. *Psychology Science Quarterly, 50*(2), 269-288.

Loveless, T. (2015). *The gender gap in reading.* Retrieved from https://www.brookings.edu/research/the-gender-gap-in-reading/

Lund Research Ltd. (2013). Quick start resources. Retrieved from https://statistics.laerd.com/premium/resources.php

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70*(3, Pt.1), 151-159.

Kratochwill, T. R., Levin, J. R., & Horner, R. H. (2018). Negative results: Conceptual and methodological dimensions in single-case intervention research. *Remedial and Special Education, 39*(2), 67-76.

Madrid, E. (2011). The Latino achievement gap. *Multicultural Education, 19*(3), 7-12.

Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty replication in the education sciences. *Educational Researcher, 43*(6), 304-316.

mClass: Reading 3D. (2010). In mClass reading 3D. Retrieved from https://www.mclasshome.com/wgenhelp/R3D2_Harcourt/Reading_3D/The_mCLASS_Reading_3D_System.htm

National Center for Education Statistics. (2015). Common core of data: 2004 & 2015. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from http://nces.ed.gov/ccd/ccddata.asp

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy. Retrieved from http://www.nifl.gov/earlychildhooc/NELP/NELPreport.html

National Institute of Child Health and Human Development (NICHHD). (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.* Washington, DC: U.S. Government Printing Office.

Nonte, S., Hartwich, L., & Willems, A. S. (2018). Promoting reading attitudes of girls and boys: A new challenge for educational policy? multi-group analyses across four European countries. *Large-Scale Assessments in Education, 6*(1), 1-22.

North Carolina Department of Public Instruction. (2000). ABC program information. Raleigh, NC: North Carolina Department of Public Instruction. Retrieved from http://www.dpi.state.nc.us/accountability/reporting/abc/2000-01/history#dss

North Carolina Department of Public Instruction. (2011). *ABC/AYP 2011 accountability report background packet.* Raleigh, NC: North Carolina Department of Public Instruction.

North Carolina Department of Public Instruction. (2014a). *Accountability brief.* Raleigh, NC: North Carolina Department of Public Instruction.

North Carolina Department of Public Instruction. (2014b). North Carolina read to achieve: Comprehensive reading plan k-12. A supplement to the North Carolina plan. Retrieved August 17, 2015, from http://www.dpi.state.nc.us/docs/k-3literacy/achieve/plan.pdf

North Carolina Department of Public Instruction. (2015). *North Carolina testing program.* Raleigh, NC: North Carolina Department of Public Instruction.

North Carolina Department of Public Instruction. (2016). Data results. Raleigh, NC: North Carolina Department of Public Instruction. Retrieved from http://www.ncpublicschools.org/accountability/reporting/

Parker, D., & Benedict, K (2002). Assessment and intervention: Promoting successful transitions for college students with ADHD. *Assessment for Effective Intervention, 27*(3), 3-24.

Parkin, J. R. (2018). Wechsler individual achievement test–third edition oral language and reading measures effects on reading comprehension in a referred sample. *Journal of Psychoeducational Assessment*, *36*(3), 203-218.

Perkins Greene, T. R. (2014). *Program evaluation of balanced literacy in an urban school district* (Order No. 3641154). Available from ProQuest Dissertations & Theses Global: The Humanities and Social Sciences Collection. (1620418039). Retrieved from http://ezproxy.gardner-webb.edu/login?url=https://search-proquest-com.ezproxy.gardner-webb.edu/docview/1620418039?accountid=11041

Preckel, F., Götz, T., Pekrun, R., & Kleine, M. (2008). Gender differences in gifted and average-ability students comparing girls' and boys' achievement, self-concept, interest, and motivation in mathematics. *Gifted Child Quarterly, 52*(2), 146-159.

Reading 3D Brochure. (2009). Retrieved May 6, 2015, from http://www.wirelessgeneration.com/pdf/brochures/mCLASS_Reading_3D_Brochure.pdf

Reardon, S. F., Valentino, R. A., & Shores, K. A. (2012). Patterns of literacy among U.S. students. *The Future of Children, 22*(2), 17-37.

Riverside Publishing. (2012). *Cognitive abilities test™ FORM 7 research and development guide.* Rolling Meadows, IL: Houghton Mifflin Harcourt.

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, *48*(2), 268-302.

Roskos, K., Strickland, D., Haase, J., & Malik, S. (2009). *First principles for early grades reading programs in developing countries.* Prepared for USAID/EQUIP1 Project. Available at http://www.equip123.net/docs/ e1-EarlyGradesToolkit.pdf

Routman, R. (2003). *Reading essentials*. Portsmouth, NH: Heinemann.

Routman, R. (2005). *Writing essentials: Raising expectations and results while simplifying teaching*. Portsmouth, NH: Heinemann.

Routman, R. (2007). *Teaching essentials: Expecting the most and getting the best from every learner, K-8*. Portsmouth, NH: Heinemann.

Sampson, R. (2002). *Sociology of delinquency. International Encyclopedia of the Social and Behavioral Science.* Oxford: Elsevier Science Limited.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology, 13*(2), 90-100.

Shanahan, T. (2016). Reading research: Importance of replication. *Reading Teacher, 70*(4), 507-510.

Simon, M., & Goes, J. (2013). *Assumptions, limitations, delimitations, and scope of the study.* Retrieved February 18, 2019 from http://www.dissertationrecipes.com/wp-content/uploads/2011/04/Assumptions-Limitations-Delimitations-and-Scope-of-the-Study.pdf

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

Spandel, V. (2008). *Creating young writers: Using the six traits to enrich writing process in primary classrooms* (2nd ed.). Boston: Pearson Allyn & Bacon.

Toncar, M. F., & Munch, J. M. (2008). Meaningful replication: When is a replication no longer a replication? A rejoinder to Stella and Adam. *Journal of Marketing Theory and Practice, 18*(1), 71-80.

U.S. Department of Education. (2008). *Achievement gaps.* Executive summary section. Washington, DC. Retrieved from https://nces.ed.gov/nationsreportcard/studies/gaps/

Visible Learning. (2016). Collective teacher efficacy according to John Hattie. Retrieved from https://visible-learning.org/2018/03/collective-teacher-efficacy-hattie/

Walkington, C., Clinton, V., & Shivrai, P. (2018). How readability factors are differentially associated with performance for students of different backgrounds when solving math word problems. *American Educational Research Journal, 55*(2), 362-414.

Weikert, D. E. (2018). *Effectiveness of reading intervention program types in increasing comprehension for intermediate school students* (Order No. 10830179). Available from ProQuest Dissertations & Theses Global: The Humanities and Social Sciences Collection. (2072463412). Retrieved from http://ezproxy.gardner-webb.edu/login?url=https://search-proquest-com.ezproxy.gardner-webb.edu/docview/2072463412?accountid=11041

Worthen, B., & Spandel V. (1991). Putting the standardized test debate in perspective. *Educational Leadership, 48*(5), 65-69.

Wright, G. (2011). Student-centered learning in higher education. *International Journal of Teaching and Learning in Higher Education, 23*(1), 92-97.

Zehner, F., Goldhammer, F., & Salzer, C. (2018). Automatically analyzing text responses for exploring gender-specific cognitions in PISA reading. *Large-Scale Assessments in Education, 6*(1), 1-26.

Appendix

Permission Letter

Dear Accountability Department

I am currently working on a doctorate in Curriculum and Instruction at Gardner-Webb University. For my dissertation, I have chosen to conduct research to determine how mClass TRC levels and CogAT scores correlate with BOG and EOG reading scores with third grade students.

I am very interested in research on the relationship between assessments with an emphasis on the NC EOG reading scores at the elementary level as part of my curriculum and instruction doctoral work through Gardner-Webb University. I have met and spoken Mrs. D. Powell in the Literacy Department. The study is entitled An Evaluation of the Relationship among North Carolina Reading Assessments. I will not administer any treatment or interact with the students in anyway. I will request existing TRC data, CogAT scores, Beginning and End-of-Grade test scores in reading and statistically analyze them to determine if there is a significant relationship between the TRC text levels, CogAT scores, BOG scores and EOG scores. All data will be completely confidential and anonymous.

Thank you so much for allowing me to collect and analyze the data in nine elementary schools to provide a strong, data-driven foundation for predictive reading text levels, cognitive ability tests, NC reading BOG and EOG scores.

If there are any questions or concerns you may contact the researcher, Sophia Crawford by phone XXXXXXX or by email at XXXXXXX and/or academic advisor, Morgan Blanton by email at XXXXXXXX.

If you approve of this proposed study, please sign below. Sincerely,

Sophia Crawford
Doctoral Candidate, Gardner-Webb University

_____ _____
Superintendent Signature Date

_____ _____
Accountability Department Signature Date