Summer 2024

An Initial Investigation of Measuring Change in Reading Proficiency During LETRS Implementation

Beverly Hart
Gardner-Webb University, bhart3@gardner-webb.edu

AN INITIAL INVESTIGATION OF MEASURING CHANGE IN READING PROFICIENCY DURING LETRS IMPLEMENTATION

By
Beverly Hart

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

Gardner-Webb University
2024

Approval Page

This dissertation was submitted by Beverly Hart under the direction of the persons listed below. It was submitted to the Gardner-Webb University College of Education and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Gardner-Webb University.

Mitch Porter, PhD Committee Chair	Date
Jessica Van Cleave, PhD Content Specialist	Date
Mitch Porter, PhD Methodologist	Date
Jennifer Putnam, EdD College of Education Representative	Date
Prince Bull, PhD Dean of the College of Education	Date

Acknowledgments

First, I would like to thank God for always being by my side. I would also like to thank my husband Tommy. I would have never accomplished this lifelong goal without his love and support. The many hours he spent alone taking care of our responsibilities and family did not go unnoticed.

I thank my committee for their role in this long process. Dr Putnam's voice echoed encouragement through trials and triumphs. During my journey, I faced unspeakable loss. Dr. Putnam never lost faith in me, and that confidence gave me the strength to continue.

I would also like to acknowledge Dr. Van Cleave. When the need arose, she stepped in to join my committee without delay. Her feedback and advice were appreciated, and she was an asset to my team.

Dr. Porter served as the chair of my committee. He went above and beyond his responsibilities to mentor me. He answered texts and phone calls at random times and was always quick to find solutions to my problems. His calm demeanor brought comfort when needed, and his wisdom proved invaluable.

Abstract

AN INITIAL INVESTIGATION OF MEASURING CHANGE IN READING PROFICIENCY DURING LETRS IMPLEMENTATION.

Hart, Beverly, 2024: Dissertation, Gardner-Webb University.

This study aimed to investigate the impact of LETRS professional development on reading proficiency. I used ANOVA tests to analyze third- through fifth-grade EOG assessment scores for changes in proficiency from 2021-2023. The results indicated a statistically significant change in the means for all grades. Third-grade results reflected the most impact from the professional development. The analysis also showed that the third-grade achievement gap had declined slightly. A dependent *t* test was conducted to ascertain any change in the growth index scores of schools. The test results showed no statistically significant change. The study's limitations include the lingering impact of the COVID-19 pandemic. The study could not show how much or how little the changes in instruction due to the pandemic influenced the results. I attributed the increase in proficiency to the professional development being grounded in the Cognitive Foundations of Reading Framework. Stakeholders should continue to invest in professional development that is based on reading research.

Keywords: phonics, decoding, phonemic awareness, professional development, assessment, word recognition, language comprehension, reading comprehension

Table of Contents

Figures

Chapter 1: Introduction

Few people in the world would argue the value of learning to read. In 2018, the International Literacy Association (ILA) declared that reading was a crucial right of all children. According to ILA (2018), reading enables individuals to function adequately in society and enriches all aspects of life. Skilled readers can develop broad background knowledge on a variety of subjects. They have wide-ranging vocabularies and critical thinking skills, making them essential to thriving communities (Hougen & Smartt, 2020). For most children, however, this right is never realized.

Children who struggle to read often live without equality and opportunity (ILA, 2018; Moats & Tolman, 2019). They live in poverty at higher rates and are more likely to become a statistic in the criminal justice system (Hanford, 2018; Moats & Tolman, 2019; Seidenberg, 2017). Illiteracy prevents individuals from obtaining basic knowledge to adequately address physical and mental health needs (Castles et al., 2018; ILA, 2018). With the stakes so high, why are there so many children who struggle to read?

The advantages of learning to read may be well established, but how we teach students to read has been at the center of debate for many decades. This debate is frequently referred to as the reading wars (Castles et al., 2018; Riley, 2020). Kilpatrick (2015) proclaimed that these so-called wars have discouraged educators from paying attention to reading research and contributed to the reading crisis in America today. Educators who were split in opinion argued for varied approaches; thus, through the years, educators have been trained through opposing viewpoints, and the argument around the best method of teaching reading continues (Castles et al., 2018).

The controversy over what constitutes best practice in reading instruction lies

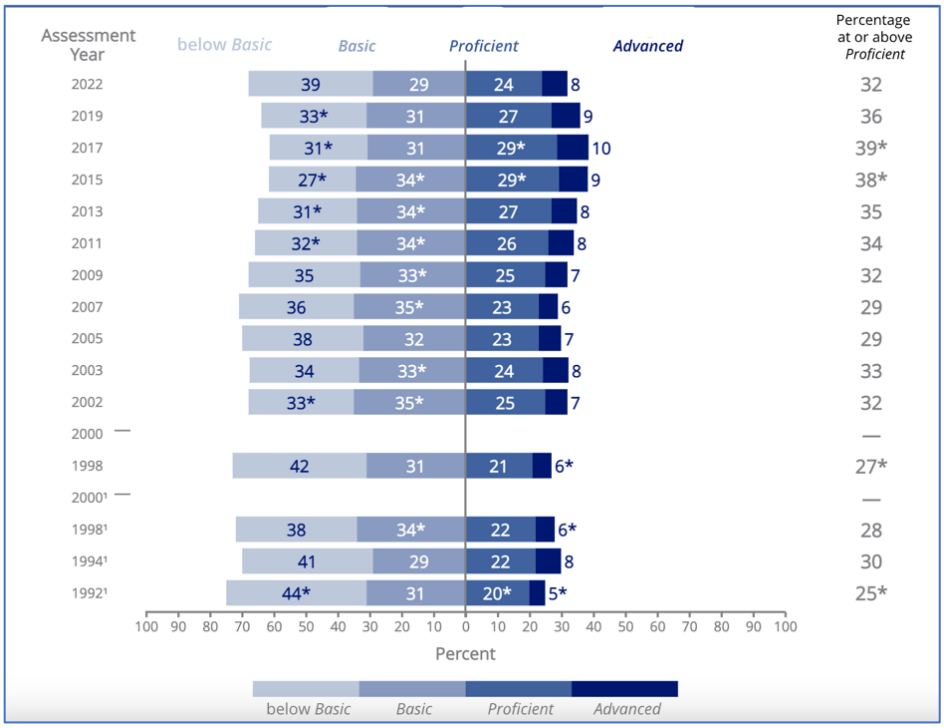
partly with how learning to read occurs. Learning to read has been proven to be a complex assembly of processes in the brain (Moats, 2020a, 2020b; Seidenberg, 2017). Teachers often have difficulty understanding this complicated process, and high-quality professional development is necessary (Moats, 2020b; Rebora, 2021).

The National Assessment of Educational Progress (NAEP) has measured the reading achievement of the nation for decades. NAEP is the most extensive national assessment of students. NAEP was administered by the National Center for Education, a United States Department of Education division. These assessments are administered to fourth- and eighth-grade students in every state at 2-year intervals. As the results of these assessments are published biennially, policymakers scramble to endorse strategies aimed at raising reading performance (Hiebert, 2022).

In 2022, only 33% of the nation's fourth graders scored at or above proficiency on the NAEP assessment. Scores for North Carolina were similar, with 32% of fourth graders scoring at or above proficiency. During the history of NAEP, North Carolina's results have shown that less than half of fourth graders are reading at or above a proficient level. Students who perform below proficiency are poorly equipped to comprehend the complex texts they encounter in school (Hennessy, 2021). Figure 1 displays the longitudinal results of the NAEP fourth-grade reading assessments.

Figure 1

Fourth-Grade NAEP Reading Assessments for North Carolina



Note. The figure was taken from NAEP (2022).

Figure 1 shows the results of the national fourth-grade NAEP assessments from 1992-2022. During the history of the assessments, the percentage of students scoring at or above proficiency has been less than 40%. In fact, the overall scores have changed very little since the assessments were first administered in 1992 (Hoover & Tunmer, 2020; Seidenberg, 2017). Moreover, the 2022 assessment scores marked a significant decline from the 2019 scores.

Hiebert (2022) noted that students scoring below proficiency lack the essential abilities needed to participate fully in a global world. Students with deficits need interventions focusing on word recognition and fluency to become proficient readers (Hiebert, 2022). Even with interventions, these students are unlikely to ever catch up to

their peers (Hougen & Smartt, 2020). These large-scale national reading assessments painted an alarming picture of the number of illiterate students in our nation and forced many states to examine instructional practices and implement new initiatives (Hennessy, 2021; Seidenberg, 2017).

Legislation

Through the years, legislators have enacted laws in an effort to reform educational systems across the nation. The first federal effort to improve outcomes for students was the Elementary and Secondary Education Act brought to action by President Lyndon Johnson in 1965. The act intended to minimize the disadvantages of poverty. The Title 1 provision was created, which distributed $1 billion in funds to low-income schools (Backstrom, 2019). The Elementary and Secondary Education Act was set to be reviewed and possibly modified every 5 years.

One of the most prominent modifications came in 2001 due to the lack of improvement of students in impoverished schools (Backstrom, 2019). The No Child Left Behind Act (NCLB) was enacted in 2001 by the George W. Bush administration. NCLB moved the definition of the targeted population beyond impoverished students to include students who were said to be at risk of failure (Backstrom, 2019). In addition, NCLB emphasized testing and placed a critical focus on teacher accountability. States were now required to administer assessments in math and English in Grades 3-8 each year. Moreover, NCLB included mandated reporting of disaggregated results among minority and disadvantaged subgroups. With these new reports, achievement gaps were brought to light (Backstrom, 2019). The act was highly criticized for extensive testing and the belief that these measures would improve student achievement. The overemphasis on testing

left little room for much-needed professional development (Dennis, 2017).

In 2015, NCLB was replaced by the Every Student Succeeds Act (ESSA). The act called for a decrease in standardized testing and a shift to more formative testing by teachers. Teachers had more autonomy to use assessments to guide instruction. Overall, ESSA placed an emphasis on professional development for teachers and literacy instruction (Dennis, 2017).

In addition to comprehensive literacy instruction, ESSA included measures to simultaneously address student well-being; however, the measures became overshadowed by the overwhelming trauma caused by the COVID-19 pandemic (Rebora, 2022a). Many states enacted laws to combat the learning deficits exacerbated by the pandemic, including the science of reading. In 2021, North Carolina passed the Excellent Public Schools Act (EPSA), which focused on early literacy instruction. EPSA (2021) adds to the legislation known as the Read to Achieve Act (RtA) established in 2012 (Hensley et al., 2017). The RtA ended social promotion and required students to be reading proficiently by the end of third grade. EPSA amended the RtA (2012) by mandating that literacy instruction and assessments be grounded in the science of reading. At the start of the 2021-2022 school year, North Carolina began implementing training in Language Essentials for Teachers of Reading and Spelling (LETRS). The training was set to occur over a period of 2 years, with a goal of all cohorts being fully trained by the end of 2024. The bill required that all prekindergarten through fifth-grade teachers of reading participate in the training.

Statement of the Problem

Despite widespread efforts to address reading deficits across the nation, many

students still find reading challenging. Students who struggle to read have difficulty in life and school; in fact, poor reading skills negatively affect almost all academic subjects (ILA, 2018; Kilpatrick, 2016; Moats, 2020a, 2020b; Wexler, 2019). A global pandemic exacerbated the reading crisis illuminated by the NAEP assessments. In 2020, the COVID-19 pandemic interrupted education and closed schools for in-person learning. As a result, students fell further behind academically, and achievement gaps widened (Donnelly & Patrinos, 2022). While significant gaps in reading have been in place across the United States for quite some time, the increase in the number of students falling behind in the lower grades is dramatic. It has created a sense of urgency and calls for immediate action (Amplify, 2022).

Deficits in reading are much more easily addressed in the early years (Ehri & Flugan, 2018; Moats, 2020a). As students continue through upper elementary and middle school, reading gaps widen and become harder to close (Kilpatrick, 2016; Scarborough, 2001). In his groundbreaking article on the science of reading, Scarborough (2001) wrote that as many as 75% of students labeled as having reading difficulty in the early grades will continue to struggle in the later school years. Reading instruction in the early years is critical (Duke & Ward, 2021; Foorman et al., 2016). Primary reading instruction demands that teachers thoroughly understand foundational reading skills (Piasta & Hudson, 2022). Literacy reform and initiatives must focus on the years leading up to the fourth-grade NAEP assessment. Hudson et al. (2021) noted that focusing teacher training and prep programs on foundational reading skills grounded in the science of reading is essential to producing knowledgeable and skilled teachers.

The science of reading is a large body of research that combines the work of

cognitive psychologists and educational researchers. This research combines multiple disciplines to describe how the brain learns to read (The Reading League, 2022). The Reading League (2022) noted that the scientific research that supports effective reading instruction has existed for decades. Even with a vast amount of scientific research, the reading crisis remains. For many educators, the total volume of reading research is overwhelming and inaccessible (Kilpatrick, 2015).

An old adage boasts, “When you know better, you do better”; however, there is often a disconnect between teacher beliefs, knowledge, and practice. For years, educators have failed to apply effective practices to improve student outcomes (Wexler, 2019). Closing the knowledge-to-practice gap remains an ongoing process (Lewis & Goodwin, 2021). This task is a difficult one. Not only do educators often reject science, but many scholars also oppose using methods that have been proven effective by an enormous amount of scientific research (Seidenberg, 2017).

Purpose and Significance of the Study

The training in LETRS, required by the state of North Carolina, comes at a time when the nation faces increasing teacher shortages that are straining the educational system (Darling-Hammond, 2022; Rebora, 2022b). The state has never required professional development of this magnitude previously, and many educators are optimistic. Intensive and structured professional development can positively affect teachers and early literacy instruction (Ehri & Flugman, 2018); however, professional development often fails to change instructional practice and increase student achievement (Gui, 2020). Moreover, required professional development can inhibit enthusiasm and leave teachers feeling disempowered from improving personal instructional practices

(Tassell, 2014).

Collecting and analyzing data is one way educators can measure the impact of professional learning (Glatthorn et al., 2019). This study assessed the impact of LETRS professional development on early reading achievement in North Carolina and contributes to the existing research in this field. Early literacy skills develop across grades kindergarten through third grade (Foorman et al., 2016; Shanahan et al., 2010). Third grade marks the cumulation of skills learned in the previous grades (Vernon-Feagans et al., 2019). In North Carolina, accountability begins with third-grade reading proficiency and continues through elementary school.

The impact of the COVID-19 pandemic continues to be examined. Studies have shown that the disruption to learning particularly impacted elementary students and minority groups (Crosson & Silverman, 2021; Curriculum Associates, 2021; Donnelly & Patrinos, 2022; Hammerstein et al., 2021; Molnar & Hermann, 2023). Most research has examined the assessment scores for the school year directly following the pandemic (Almasi & Yuan, 2023; Schult et al., 2022; Sun et al., 2023). There is a need for additional research examining any effects of efforts to address learning loss and reading proficiency beyond the 2020-2021 school year following the pandemic.

According to the North Carolina Department of Public Instruction (NCDPI) Office of Early Learning (2023), overall student growth and proficiency should align with the LETRS professional development implementation. This study analyzed the proficiency data collected from the end-of-grade (EOG) assessments in third through fifth grades. The study analyzed any changes in reading proficiency after the implementation of the intensive professional development mandated by the state of North Carolina. In

addition, a quantitative analysis was conducted to ascertain any changes in the school growth index scores. The study focused on the districts from Cohort 1 that had completed the professional development at the time of the research study. Previous research regarding LETRS is limited, and this study contributed to the research on the impact of this professional development. The following research questions were developed to address the problem of the study.

1. How have assessment scores changed since the beginning of LETRS training implementation?
2. How have growth index scores changed since the beginning of LETRS training implementation?

Summary

The ability to read is powerful and life-altering for society (Hanford, 2018; Hougen & Smartt, 2020; ILA, 2018; Moats & Tolman, 2019; Seidenberg, 2017). Experts from many disciplines have produced a vast amount of research on the development of reading. The field has no shortage of evidence-based instructional strategies and practices. With so much research comes disagreement. Throughout history, there has been debate on what constitutes best practice (Castles et al., 2018; Kilpatrick, 2015; Riley, 2020). Amid this disagreement, many students across the nation fail to learn to read. The impact of a global pandemic has contributed to the sense of urgency in addressing the reading crisis across the nation.

In 2019, Mississippi was the only state to show statistically significant gains on the NAEP fourth-grade reading assessments. These gains were contributed to a state-wide literacy improvement effort that began in 2013 (Burk, 2022). At that time, Mississippi

passed the Literacy-Based Promotion Act. This legislation sought to align instruction to the science of reading, and the state subsequently implemented LETRS. The LETRS professional development had already been recognized as improving student achievement in a large research study completed by Folsom et al. (2017). With the attention to the 2019 NAEP scores being high, states like North Carolina hoped to replicate Mississippi's reading growth.

Legislation in many states, including North Carolina, aims to increase student reading proficiency. LETRS is a thorough professional training on how students learn to read and the best instructional practices based on the science of reading (Moats & Tolman, 2019). This study investigated the reading achievement of students before and after the implementation of LETRS by analyzing reading proficiency scores on the EOG assessments. In addition, the study analyzed school growth index scores for any changes from Year 1 to Year 2 of the professional development.

Definition of Terms

Clarifying certain terms aids in the understanding of concepts and ideas presented in research and studies. The list of words is not exhaustive and only represents some of the overarching concepts explored. The following terms are used throughout this study.

Decoding

The ability to understand the relationship between letters and sounds and apply that understanding to print. Decoded words are sounded out and translated into speech (Moats, 2020a).

Grapheme

A term used to apply to a letter or letter combination representing a single

phoneme (Lane & Contesse, 2022).

Language Comprehension

"The ability to extract and construct literal and inferred meaning from linguistic discourse represented by speech" (Hoover & Tunmer, 2020, p. 26). Language comprehension is an essential component of reading comprehension, as illustrated in the equation represented by the Simple View of Reading (SVR; Gough & Tunmer, 1986).

LETRS

An in-depth professional development in reading that is not a curriculum. LETRS provides teachers of language arts with foundational knowledge that can be applied to any reading program (Moats & Tolman, 2019).

Morphology

The knowledge of each meaningful part of a word, including roots, prefixes, and suffixes (Spear-Swerling, 2022). Morphology is further described as the "study of meaningful units of language and how they are combined in word formation" (Moats, 2020a, p. 300).

Orthographic Mapping

The mental process of storing words in memory for automatic and accurate retrieval. This process requires proficiency in phonemic awareness and decoding (Moats, 2020a).

Phonemes

"A speech sound that combines with others in a language system to make words" (Moats, 2020a, p. 301).

Phonemic Awareness

The conscious awareness of speech sounds in spoken words and the ability to manipulate those sounds (Moats & Tolman, 2019). There are six levels of phonemic awareness: phoneme isolation, phoneme blending, phoneme segmentation, phoneme addition, phoneme deletion, and phoneme substitution. The most difficult levels are necessary for successful reading development (Kilpatrick, 2016).

Phonics

"The study of the relationships between letters and the sounds they represent; also used to describe reading instruction that teaches sound-symbol correspondences, such as the phonics approach or phonic reading" (Moats, 2020a, p. 302). The National Reading Panel (NRP, 2000) described two main types of phonics instruction: synthetic and analytic. A synthetic approach teaches students to identify each sound in a word and blend the sounds together. The analytic approach teaches students to recognize whole word units first, and then students are taught to link sounds to letters.

Professional Development

Training or instruction that provides educators with information on pedagogical practices or content matters. Professional development is also referred to as professional learning (Glatthorn et al., 2019). ESSA defined professional development as the activities schools and local agencies provide to educators (Dennis, 2017). These activities are intensive and sustained. They are also collaborative and data-driven. In addition, the activities are job-embedded and applicable to the classroom.

Semantics

Refers to deriving meaning at word level, sentence level, or paragraphs in longer texts (Spear-Swerling, 2022). Semantics is defined as "the study of word and phrase

meanings" (Moats, 2020a, p. 304).

Sight Words

Often referred to as high-frequency words. Sight words are known words that do not have to be sounded out and are recognized automatically upon sight (Moats, 2020a).

Chapter 2: Literature Review

A review of the literature in this chapter provides background information and insight into the problem of illiteracy among America's youth. During the review, it was necessary to examine the current theories on how students learn to read. As a result, descriptions of the SVR (Gough & Tunmer, 1986) and Scarborough's (2001) Reading Rope are provided. These frameworks provide the basis of the current scientific model of reading development.

The seminal research of Ehri (1995) was significant to the description of the reading development processes. There are four distinct stages of progression according to Ehri (1995). These stages are known as the phases of reading development, and they are detailed in this chapter.

A search regarding what comprises the best instructional reading practices resulted in a summary of the NRP (2000) report. This report described the primary instructional areas of focus for efficient reading comprehension to develop. In addition, I looked at instructional practices and shifts in reading education throughout the last few decades. This history is relevant to how reading instruction has evolved to its current state. The most recent push to emphasize scientific research in instruction produced a search and subsequent explanation of the term the science of reading.

How well students are taught to read in schools can impact their long-term well-being (Shanahan, 2020); therefore, I researched effective professional development pedagogy and practices. The research described essential criteria for successful teacher learning. In addition, I describe the guiding standards for professional development set forth by the Professional Learning Association.

Finally, Chapter 2 reviews the literature on assessments. This section provides relevant information on the assessment components analyzed in the study. The literature search also provided information on how the assessment is aligned with the instructional areas presented under the section on the National Reading Panel (2000).

Background to the Problem

In 1983, the National Commission on Excellence in Education published a report titled *A Nation at Risk.* The famed report stated that the American educational system was in a state of severe decline, and economic failure was looming. It caused great alarm and sparked state and federal reforms (Floden et al., 2020; Hirsch, 2020). The subsequent NCLB and ESSA called for teacher transformation and increased accountability. Decades later, the reading crisis remains as national assessments show little improvement, and improvements to the educational system have become stagnant (Fogarty et al., 2020; Hirsch, 2020; Seidenberg, 2017).

As noted in Chapter 1, there have been many contributors to the reading crisis. One example is the disconnect between research and practice. Pfeffer and Sutton (2000) identified the knowing-doing gap. Pfeffer and Sutton stated that a gap occurs when actions are inconsistent with the research that identifies effective practice. This theory applies to reading instruction and is often referred to as the research-to-practice gap in education (Kilpatrick, 2015; Seidenberg, 2017; Wexler, 2019).

Cognitive scientists have produced research that supports effective reading instruction for many decades; however, few classrooms represent this research (Kilpatrick, 2015, 2016; Seidenberg, 2017; Wexler, 2019). Many teachers lack the motivation and peer support to conduct effective research (Lewis & Goodwin, 2021).

Other teachers cannot readily access existing research or fail to understand it (Hoover & Tunmer, 2020; Kilpatrick, 2015; Seidenberg, 2017). The matter is exacerbated by many teacher preparation programs not teaching reading research (Solari et al., 2020).

Educators are often presented with research and practices that conflict with what they have done for years and believe to be effective. This contrasting information causes confusion, known as cognitive dissonance (Wexler, 2019). Veteran educators may fear embracing research signifies that their instructional practices have been ineffective or harmful to students (Wexler, 2019). This state contributes to the knowledge-doing gap (Pfeffer & Sutton, 2000).

Through professional training in reading development, teachers are immersed in research and pedagogical skills. As teachers acquire knowledge and research through the LETRS professional development, they apply this information to their instructional practice. Ideally, the gap between knowing and doing will begin to close, and student achievement will rise.

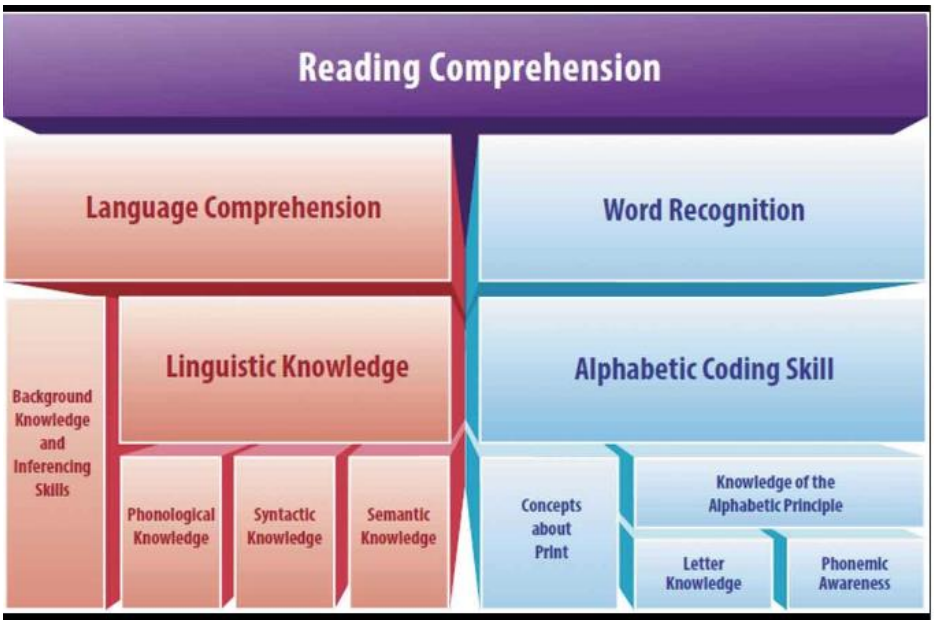
Theoretical Framework

Frameworks are considered tools used to analyze complex concepts. The cognitive foundation, in relation to reading, is defined as the "distinct framework component, element, or module that reflects a mental ability, capacity, competency, function, or state" (Hoover & Tunmer, 2020, p. 14). This Cognitive Foundations of Reading Framework provides the theoretical groundwork for this study (Hoover & Tunmer, 2020). The Cognitive Foundations of Reading Framework explains the psychological complexities and processes required for learning to read, and it is built on the SVR (Hoover & Tunmer, 2020). Figure 2 shows the conceptual model of the

Cognitive Foundations of Reading Framework as a hierarchically composed structure.

Figure 2

The Cognitive Foundations of Reading Framework Model



Note. This figure was taken from Hoover and Tunmer (2020).

Figure 2 depicts Hoover and Tunmer's (2020) representation of the Cognitive Foundations of Reading Framework. In this model, the cognitive components of reading comprehension are seen as independent blocks that build on one another. Figure 2 further illustrates the hierarchical relationship between the skills. The red and blue colors indicate the independence of the language comprehension and word recognition strands. The color purple represents the connectivity when combined to produce reading comprehension.

Hoover and Tunmer (2020) described each of the components of the framework. The overarching reading comprehension block was defined as the ability to construct literal and inferential meaning from printed text. I begin with Hoover and Tunmer's (2020) explanation of the language comprehension strand and the subsequent blocks

beneath this strand. Language comprehension was described as the ability to construct meaning from the discourse in speech. The background knowledge and inferencing skills block was described as the knowledge of pertinent content, preceding linguistic discourse, context, and the ability to derive meaning beyond the literal. Linguistic knowledge was delineated as the unconscious knowledge of grammar. Phonological knowledge was characterized as in-depth knowledge about speech sounds. Syntactic knowledge was defined as a thorough understanding of phrases, clauses, and sentences. Semantic knowledge referred to the ability to understand units of language at the word and sub-word levels and how they contribute to building meaning at the sentence and discourse levels.

Hoover and Tunmer (2020) then explained the word recognition strand and the succeeding blocks below it. Hoover and Tunmer (2020) noted that word recognition was the ability to recognize printed words accurately and fluently to acquire meanings stored in the mental lexicon. The subsequent alphabetic coding skills component was described as the ability to map letters and patterns onto phonemes. Next, the concepts about print section was characterized by the knowledge of how print works. Print knowledge includes letter knowledge. Letter-name knowledge plays an important role in letter-sound knowledge (Kilpatrick, 2015). In fact, knowledge of letter names is necessary for learning the alphabetic principle (Hoover & Tunmer, 2020). There are several important ways that letter-name knowledge contributes to beginning reading development. First, many letter names are used to represent speech sounds in words. Second, many names of letters contain the actual phoneme that the letter typically refers to. Third, letter-name knowledge facilitates phonemic awareness development (Hoover & Tunmer, 2020;

Kilpatrick, 2015).

As students master letter knowledge and print concepts, they gain understanding of the alphabetic principle. Knowledge of the alphabetic principle was described as the conscious of how letters represent phonemes. Finally, letter knowledge was noted as the ability to recognize and manipulate letters in print, and phonemic awareness was the ability to recognize and manipulate phonemes in spoken words.

The SVR

The SVR was first introduced in the empirical work by Gough and Tunmer in 1986 as a theoretical model for learning to read. This model has become the guiding scientific theory for how skilled reading evolves (Kilpatrick, 2020). It is used as a theoretical framework for researchers and provides a valuable framework for educators to understand the reading process (Kilpatrick, 2015). Although the model itself is simple, the SVR processes are very complex (Kilpatrick, 2015; Moats, 2020a, 2020b; Moats & Tolman, 2019).

The SVR stated that skilled reading was the product of decoding and linguistic comprehension. A very important component in the equation is that the components are combined multiplicatively and not additively (Hoover & Tunmer, 2020). According to the SVR, the two processes are of equal importance, and skilled, proficient reading cannot develop without one or the other (Gough & Tunmer, 1986; Hoover & Tunmer, 2018). A deficiency in one of these domains will result in a deficiency in reading comprehension (Gough & Tunmer, 1986; Petscher et al., 2020).

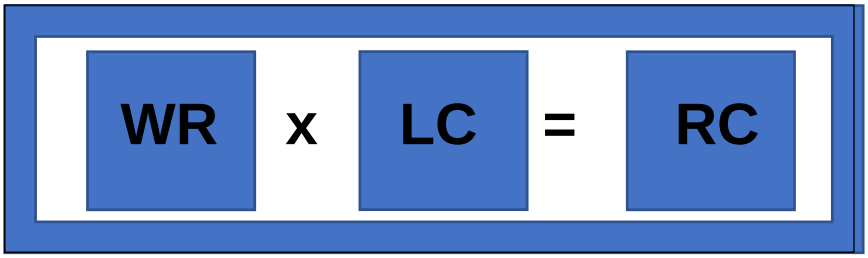
As Gough and Tunmer (1986) first proclaimed, reading comprehension demands decoding and linguistic skills. The SVR was represented as a simple equation. Gough and

Tunmer's equation stated that decoding (D) multiplied by comprehension (C) equals reading (R) or D X C = R. Comprehension in the equation was also referred to as linguistic comprehension (LC) and reading as reading comprehension (RC; Duke & Cartwright, 2021).

Often, the term decoding is interchanged with the term word recognition. Linguistic comprehension is often referred to as listening comprehension or language comprehension. All three terms have equal meaning (Hoover & Tunmer, 2018). The terms word recognition and language comprehension are broader terms that better reflect the skills needed for reading comprehension and still hold true to the original construct of the SVR (Duke & Cartwright, 2021; Hoover &Tunmer, 2020). Representations of the SVR using broader terms are used in most contemporary models, as shown in Figure 3 (Duke & Cartwright, 2021).

Figure 3

Visual Model of SVR



Note. Adapted from Moats and Tolman (2019).

The equation shown in Figure 3, with the terms word recognition and language comprehension, is how the SVR is presented to teachers during LETRS training (Moats & Tolman, 2019); therefore, teachers are most familiar with this representation of the SVR.

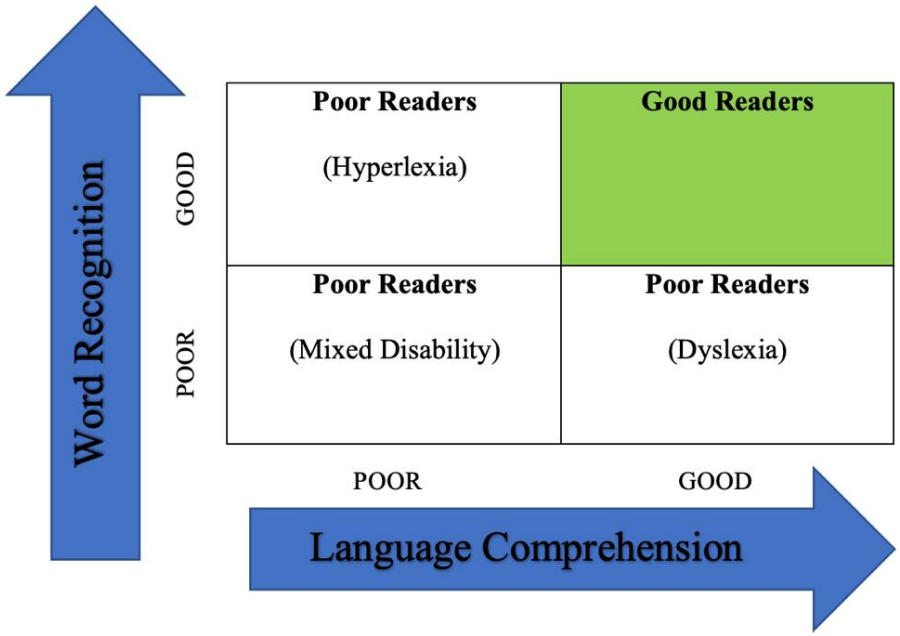
Decoding refers to the ability to understand the printed word (Hoover & Tunmer, 2018; Moats & Tolman, 2019). As readers decode, they convert the printed word to speech through their knowledge of the sound-symbol relationships of letters (Moats, 2020a). When decoding occurs quickly and accurately, word recognition becomes automatic (Moats & Tolman, 2019; Moats, 2020a). Automatic word recognition frees up the brain to process the meaning of text, while poor word recognition showcases short-term memory limitations as the reader forgets what was just read and misinterprets the meaning of sentences (Hoover & Tunmer, 2018; Moats & Tolman, 2019).

The other domain of the SVR, language or linguistic comprehension, requires the understanding of oral language (Moats & Tolman, 2019). Linguistic comprehension is known as the ability to understand spoken language (Hoover & Tunmer, 2018; Moats & Tolman, 2019). Readers with adequate skills in this area can understand texts that are read aloud and retell them in their own words (Hoover & Tunmer, 2018). People are unlikely to understand what they try to read if they fail to understand what is being said to them (Moats & Tolman, 2019).

Skilled readers only develop when skills are mastered in language comprehension and word recognition. Good language comprehension coupled with good word recognition skills predicts good reading comprehension (Moats, 2020a). Poor readers can lack skills in word recognition, language comprehension, or both areas. Hoover and Tunmer (2020) displayed the relationships of the strands with a visual grid that indicates where poor reading occurs. Figure 4 shows this grid.

Figure 4

Matrix Showing the Intersection of Poor and Good Readers in Word Recognition and Language Comprehension



Note. Adapted from Hoover and Tunmer (2020).

Figure 4 shows the dependency between the vertical domain of word recognition and the horizontal domain of language comprehension. This model shows good readers are found in the green quadrant, where skilled readers intersect. Poor readers exist in the other three quadrants. We see poor readers in the quadrant where language comprehension is good, but word recognition is not developed. This category includes people with dyslexia who show deficits in phonological processing despite solid language instruction. On the other side, the quadrant showing where word recognition is good and language comprehension is lacking, we find poor readers, including those known as hyperlexics. Hyperlexics have excellent word recognition but poor comprehension. Finally, where readers lack word recognition and language comprehension skills, we find

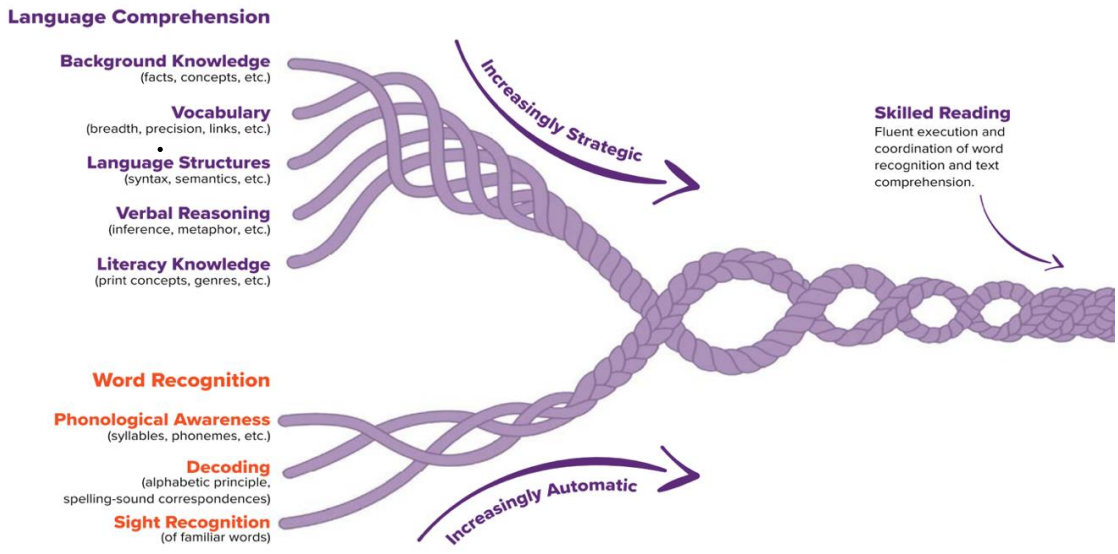
poor readers with mixed disabilities (Hoover &Tunmer, 2020). Figure 4 indicates the multiplicative nature of the SVR, showing that good readers only result from the combination of both strands.

Scarborough's Visual Model

In 2001, Scarborough created a visual model that embodied the concepts represented by the SVR. The model displays the two domains of the SVR and the subsequent skills within each domain. This model was termed Scarborough's Reading Rope and is shown in Figure 5. Figure 5 shows that as each strand becomes entwined, the rope strengthens and reading becomes progressively automatic. Scarborough's Reading Rope has become one of the most important depictions of reading development since its conception. It portrays the steady and continuous development of reading over the elementary years (Moats, 2020a).

Figure 5

Scarborough's Reading Rope



Note. Image taken from Scarborough (2001).

Figure 5 displays the subskills of the language comprehension and word recognition domains. The subskills under the word recognition strand are the early literacy foundational skills typically measured in the primary grades. Phonological awareness is the skill area associated with understanding spoken sounds in words. Decoding refers to the ability to understand the relationship between sounds and letters. As readers associate sounds with words, they automatically recognize words by sight (Moats, 2020a).

The language comprehension strand houses increasingly complex subskills. Background knowledge refers to the existing understanding of concepts, ideas, and information related to texts. Vocabulary is defined as an individual's level of understanding of words. This understanding includes the breadth, or size, of an individual's mental lexicon and the depth or how well an individual understands a word. The subskill of understanding language structures overlaps vocabulary in semantics or knowledge of word meanings and word choice. Vocabulary encompasses morphology, the understanding of the meaningful parts of words. Language structure also includes knowledge of the syntactic structure of sentences. Verbal reasoning incorporates the comprehension skills of understanding figurative language and inferring meaning from texts. Finally, literacy knowledge includes such concepts as print awareness and the understanding of overall text structures and various genres of texts (Hennessy, 2021; Moats, 2020a).

Scarborough (2001) depicted the subskills of each domain as strains interwoven to produce skilled reading. Readers at the end of the rope handle text fluently and automatically (Moats, 2020a). Although the strands can be defined and measured

independently, each strand has an interrelated influence on the development of reading, and each is necessary for fluent reading and comprehension to evolve (Moats & Tolman, 2019). Skilled reading becomes jeopardized when one of the strands becomes frayed (Hennessy, 2021).

Phases of Reading Development

The components of skilled reading visualized by Scarborough's Reading Rope are described as behaviors through reading phases developed by Ehri (Moats, 2020a). The development of learning to read and spell words is considered to occur in phases. In the late 1980s, Ehri (1995) began to write about the development of word reading as occurring in somewhat of a linear progression of phases. Ehri noted that spelling and reading were intricately related (Ehri, 2020; Moats & Tolman, 2019). In 1995, Ehri published a paper on the phases of reading development that became widely recognized and provided important implications for theories of reading development (Beech, 2005; Moats, 2020a). Figure 6 depicts a representation of these phases of reading development.

Figure 6

Ehri's Stages of Word Reading

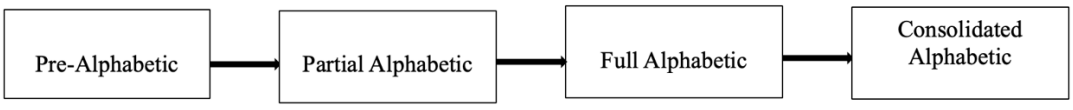


Figure 6 shows Ehri's (1995) four phases of reading: pre-alphabetic, partial alphabetic, full alphabetic, and consolidated alphabetic. The behaviors associated with each phase occur across a continuum of learning that blends as growth transpires (Moats, 2020a). Ehri named the stages for the predominant behavior characteristic of each phase

(Ehri, 2020).

Pre-Alphabetic

The pre-alphabetic stage is also known as the prereading stage. At this beginning stage, students have not yet learned the alphabetic principle and do not understand the letter-sound relationship in words (Ehri, 2020; Moats, 2020a). A student in this stage may be able to recognize a few letters but not fully understand their purpose (Lane & Contesse, 2022; Moats, 2020a). They rely heavily on visual cues or pictures to derive meaning in context (Ehri, 2020; Lane & Contesse, 2022; Moats, 2020a; Moats & Tolman, 2019). For example, the shape of the golden arches helps a child "read" the word McDonald's during this developmental phase (Lane & Contesse, 2022). In addition, there is a lack of phonemic awareness, the ability to identify and manipulate individual phonemes in spoken words (Kilpatrick, 2015; Lane & Contesse, 2022).

Partial Alphabetic

The partial alphabetic phase is also referred to as the early alphabetic stage (Moats, 2020a; Moats & Tolman, 2019). There is a distinctive shift in student insight into spelling and reading at this stage. Students know most letters and are beginning to associate phonemes and graphemes. Phonemic awareness is increasing, and students are starting to read and spell by sounding out words (Lane & Contesse, 2022; Moats, 2020a). Reading at this stage involves guessing from the recognition of the initial consonants in words (Moats, 2020a).

Full Alphabetic

This stage is also known as the later alphabetic stage (Moats, 2020a; Moats & Tolman, 2019). In contrast to the partial alphabetic phase, students now use all the letters

in a word to decode. They have a growing knowledge of phonemic awareness and can blend and segment sounds in words (Lane & Contesse, 2022; Moats, 2020a). In addition, students can recognize letter changes in any position in a word (Ehri, 2020). Although they know most phoneme-grapheme correspondences by this point, students lack automaticity and read disfluently because so much attention is needed to match sounds to symbols (Moats, 2020a).

Consolidated Alphabetic

The consolidated alphabetic phase is labeled the full orthographic mapping phase by Moats (2020a). Orthographic mapping is the process by which words are recognized instantly upon sight with automaticity (Kilpatrick, 2015; Moats, 2020a). This process facilitates fluency in reading. At the consolidated stage, students have reached an advanced level of phonemic awareness. They can reverse, exchange, delete, and add sounds in words (Moats, 2020a). In addition, students can decode syllables and recognize meaningful parts of words. Knowledge of the alphabetic principle can now be applied to unfamiliar words encountered in a text (Ehri, 2020; Lane & Contesse, 2022; Moats, 2020a; Moats & Tolman, 2019).

The Reading Debate or "Reading Wars"

The debate over reading instruction, commonly known as the reading wars, has continued for over 100 years (Petscher et al., 2020). During the past century, educators have gone back and forth between primary phonics instruction and whole language (Castles et al., 2018). Phonics is a bottom-up approach where letters and sounds are considered before words. This code-based system explicitly teaches letters and their sounds (Castles et al., 2018). In contrast, whole language, also known as the "look and

say" method, is considered a top-down approach where words come first.

Whole language proclaims that students do not need explicit instruction and can learn to read by exposure to print. Advocates for whole language do not see a need to focus on the cognitive process of learning to read (Hoover & Tunmer, 2020). In 1950, whole language was the predominant method for reading instruction. This method began to be challenged in 1955 when Rudolf Flesch published *Why Johnny Can't Read.* Flesch criticized memorization and whole-word guessing in this well-known book and advocated for a phonics approach. Flesch contended that most third graders could not understand 90% of the text encountered and the way to remedy the problem was to teach students to decode words. According to Flesch, many reading professors failed to adhere to their own research. This book ignited the phonics versus whole-language debate, also known as the reading wars.

As the debate heated up nationally, the Carnegie Corporation enlisted the head of Harvard's Reading Laboratory, Jeanne Chall, to conduct a study from 1962 until 1965. Chall analyzed existing research on beginning reading instruction while conducting her own. In 1967, Chall summarized her findings in *Learning to Read: The Great Debate*. Chall found that phonics instruction in the early grades was superior to the whole language approach and programs lacking direct phonics. Chall maintained that students who were taught using this method became better spellers and eventually were better able to comprehend text.

Despite the recognition and validity of Chall's (1967) work, researchers continued crediting the whole language approach. Goodman (1967) contradicted the explicit phonics approach by advocating that students could learn to read by being immersed in a

literacy-rich environment. Goodman (1967, as cited in Seidenberg, 2017) referred to learning to read as a psycholinguistic guessing game. Goodman maintained that good readers used background knowledge and context clues to guess unfamiliar words. Goodman's work became the basis of what is known as the three-cueing system. With the three-cueing systems, readers use knowledge of graphics, semantics, and syntax to determine a word (Castles et al., 2018). This system became a popular way for teachers to record and code errors in student reading (Stouffer, 2021).

Goodman's (1967) work fueled further research by cognitive psychologist Frank Smith. Smith (1971) supported the idea that skilled readers used context to construct meaning. Smith contended that phonics instruction was built on a complicated set of rules that were unnecessary in learning to read. Moreover, Smith proclaimed that reading occurred naturally, similar to speaking. Reading as a natural act contradicts cognitive science (Hoover & Tunmer, 2020; Moats, 2020a), and this idea would later surface as a focal point in the science of reading movement. Seidenberg (2017) asserted that although Goodman's and Smith's works were popular, they lacked scientific evidence to support their claims.

Balanced Literacy

In the early 2000s, a new term for literacy instruction surfaced. Balanced literacy emerged as a compromise to the debate in the 1990s (Fisher et al., 2021; Seidenberg, 2017). The balanced literacy approach was defined as a framework that included both phonics and whole language (Chai et al., 2020; Fisher et al., 2021). The comprehensive approach could meet student needs while respecting teacher beliefs and autonomy (Chai et al., 2020).

However, the balanced literacy approach was criticized for a multitude of reasons. Opponents stated that the method devoted too much time to whole-group instruction while overlooking the need for instruction in specialized skills (Fisher et al., 2021). Balanced literacy was also condemned for allowing teachers too much freedom to pick and choose whatever method of instruction they felt worked best for them (Seidenberg, 2017). Other claims hailed the approach as a political smokescreen to calm the debate or nothing more than a slogan (Fisher et al., 2021; Seidenberg, 2017). For many teachers, balanced literacy became a vague, undefined term (Fisher et al., 2021).

NRP Report

As the nation continued its heated discussions concerning literacy instruction, Congress asked the National Institute of Child Health and Human Development to form a panel of experts to weigh in on the debate. NRP was charged with reviewing critical research to identify the most effective instructional practices in early literacy. The report of NRP (2000) became one of the most prominent comprehensive reviews of reading pedagogical practices. The findings would be validated in later comprehensive studies (Foorman et al., 2016). NRP identified five major areas of instruction, or subgroups, needed for skilled reading to develop. These areas, phonemic awareness, phonics, vocabulary, comprehension, and fluency are reflected in Scarborough's (2001) model and are significant to the science of reading. The instructional recommendations presented in NRP in these areas are directly related to the Cognitive Foundations of Reading Framework (Hoover & Tunmer, 2020).

Phonemic Awareness. NRP (2000) found that explicit and systematic phonemic awareness instruction was necessary. Phonemic awareness is a skill that falls under the

larger umbrella term of phonological awareness. While phonological awareness includes the awareness of all units of sounds in spoken words, phonemic awareness is the ability to recognize and manipulate individual phonemes or sounds (Kilpatrick, 2016; Moats & Tolman, 2019). NRP noted that skills associated with the larger units of phonological awareness are easier for students to master than those associated with the smaller units of phonemic awareness. Phonemic awareness develops along a continuum from identifying to manipulating speech sounds. Skilled reading includes the basic phonemic awareness skills of blending and segmenting phonemes, and the more advanced skills of deleting, substituting, and reversing phonemes are also necessary (Kilpatrick, 2015, 2016).

Phonemic awareness is a critical skill in learning to read (Hoover & Tunmer, 2020; Kilpatrick, 2015, 2016; Moats, 2020a; Moats & Tolman, 2019). Words cannot be permanently stored in memory for easy retrieval without the mastery of skills at the phoneme level (Kilpatrick, 2015, 2016). Moreover, most children with difficulty reading have deficits in phonemic awareness, making it critical for developing accurate and fluent reading (Hoover & Tunmer, 2020; Kilpatrick, 2015, 2016; Moats, 2020a; Moats & Tolman, 2019). Phonological awareness and understanding the alphabetic principle are skills that become intertwined as students begin to read words (NRP, 2000; Otaiba et al., 2022).

Phonics. While phonemic awareness is associated with oral language, phonics is associated with printed language. Phonics instruction teaches the relationship between letters and their sounds (Moats, 2020a). Phonics instruction facilitates understanding of the alphabetic principle. The understanding that phonemes are represented by a grapheme or unit of print is the initial process of understanding the alphabetic principle

(Seidenberg, 2017). Knowledge of the alphabetic principle is also described as the awareness of the connection between the sounds of spoken letters and the letters in an orthography (Kilpatrick, 2015; Moats, 2020a). Reading development depends on acquiring the alphabetic principle (Ehri, 2014). This awareness typically develops in kindergarten during the early alphabetic phase of reading development (Moats, 2020a).

Readers do not come to understand the alphabetic principle naturally. Rather, they must receive explicit instruction throughout the various stages of development (Castles et al., 2018). During this process, students learn to decode. Decoding is the ability to convert a printed word from print to speech (Moats & Tolman, 2019). Like phonemic awareness, phonics instruction is essential to reading development and should be intentional and follow a system that progresses from easier to harder skills (Hoover & Tunmer, 2020; Kilpatrick, 2016; Moats, 2020a; Moats & Tolman, 2019; NRP, 2000).

Vocabulary. Scarborough's (2001) Reading Rope portrays vocabulary as one of the interconnected strands that leads to proficient reading. The SVR suggests that students use oral language abilities to become skilled readers. A student's oral language ability is greatly influenced by vocabulary (Coyne & Loftus-Rattan, 2022). NRP (2000) found that direct and indirect vocabulary instruction correlated to increased gains in comprehension. The relationship between vocabulary instruction and reading comprehension increases as students age. As students progress through the grades, their knowledge of words becomes more extensive, leading to a deeper understanding of the text (Coyne & Loftus-Rattan, 2022). A vocabulary-rich curriculum should contain varied language experiences and word-learning strategies (Hennessy, 2021).

Comprehension. Another area studied by NRP (2000) was comprehension. The

SVR (Gough & Tunmer, 1986) and the Reading Rope (Scarborough, 2001) provided insight into the complicated processes involved in reading comprehension. A solitary process cannot explain reading comprehension but instead is the product of many complex systems interacting simultaneously (Castles et al., 2018). NRP found that teaching specific skills and strategies increased reading comprehension. This instruction should include comprehension monitoring, cooperative learning, graphic or semantic organizer, questioning, story structure, and summarization. While these skills can be taught in isolation, teaching multiple skills together is more effective (Hennessy, 2021; NRP, 2000).

Fluency. NRP (2000) found that fluency was dependent on accurate word recognition skills as well as other factors. Fluency is defined by NRP as the ability to read orally with accuracy, automaticity, and expression. Hudson et al. (2022) stated that fluent readers have a strong understanding of the alphabetic principle and store many high-frequency words in memory. NRP noted that fluency was often overlooked in instruction despite its significance to reading comprehension. Effective instructional practices that increase reading fluency include repeated readings and prosody lessons (Hudson et al., 2022; NRP, 2000).

Erroneously, many people characterize fluency as speed reading while guessing or skipping words. There are several key characteristics of a fluent reader. Accuracy is one element of fluency. Fluent readers have mastered the alphabetic principle. They can blend sounds into words accurately. Another prominent element of fluency is automaticity. Automaticity is defined as being able to read a large number of words effortlessly. Finally, fluency involves reading with prosody. Prosody involves reading

with expression and appropriate pacing. Reading comprehension is reciprocally related to reading prosody (Hudson et al., 2022; NRP, 2020).

National Early Literacy Report

While the NRP (2000) report became one of the most influential pieces of research in the field of reading, it was limited and did not reflect on many other areas of reading development (Shanahan & Lonigan, 2010). Another report became a part of the seminal research on reading development and educational policies. In 2002, the National Institute for Literacy assembled the National Early Literacy Panel to create a research synthesis of empirical studies on early literacy development. In 2008, the National Early Literacy Panel released *Developing Early Literacy: Report of the National Early Literacy Panel*.

The National Early Literacy Panel (2008) conducted a meta-analysis of approximately 500 studies related to children from birth to age 5. The National Early Literacy Panel addressed the following questions:

1. What skills and abilities in young children predict later outcomes in reading, writing, and spelling?
2. What programs, interventions, and instructional approaches contributed to or inhibited gains in the skills and abilities linked to later outcomes?
3. Which environments or settings added or inhibited the skills and abilities associated with later outcomes?
4. What characteristics of young children contributed to or inhibited the skills and abilities connected to later outcomes?

The key findings of the report included the identification of the realm of early

literacy skills as precursors and instructional practice that enhanced these skills. Six early literacy skills were found to have moderate to strong predictability on later literacy achievement. An additional five variables were identified as being moderately associated.

Precursor Literacy Skills. The first six included alphabet knowledge, phonological awareness, rapid automatic naming of letters or numbers, rapid automatic naming of colors or objects, writing, and phonological memory. Alphabet knowledge and phonological awareness were identified and defined by NRP (2000). Rapid automatic naming was noted as the ability to rapidly name a sequence of objects, colors, numbers, or letters that were randomly presented. Writing denoted the capability of writing isolated letters or one's name when requested. Finally, phonological memory was defined as being able to remember spoken information for a brief amount of time. These primary skills remained predictive when the variables of IQ or socioeconomic status were added.

The other five early literacy skills were not evaluated for predictability against IQ or socioeconomic status variables. Nonetheless, they were considered to show a moderate correlation with at least one measure of later literacy achievement (Shanahan & Lonigan, 2010). These additional precursors were noted as concepts about print, print knowledge, reading readiness, oral language, and visual processing. Concepts about print included a knowledge of print conventions such as the directionality of text. Print knowledge encompassed alphabet knowledge, concepts of print, and early decoding skills. Reading readiness was defined as a combination of alphabet knowledge, concepts of print, vocabulary, memory, and phonological awareness. Oral language refers to the ability to produce and understand spoken language, vocabulary, and grammar. The final skill of visual processing was defined as the ability to match or categorize symbols presented

visually.

Instructional Practices. The National Early Literacy Panel (2008) acknowledged five categories of instructional practices or strategies correlated to the identified conventional precursor skills. These categories were labeled as code-focused interventions, shared-reading interventions, parent and home programs, preschool and kindergarten programs, and language-enhancement interventions. Interventions that focused on the alphabetic code or phonics were defined as code-focused. These interventions often included phonological awareness instruction as well. Shared-reading interventions were characterized by reader and child interactions with text material being read aloud. Interventions described as parent and home programs involved training parents in practices stimulating linguistic and cognitive development at home. The preschool and kindergarten programs category comprised educational programs, curricula, and policies that dealt with kindergarten-age students. The final category of language-enhancement interventions was defined as studies on the effectiveness of instructional practices designed to improve language development.

Current State of Literacy

In 2019, ILA published a report that acknowledged the continuous debate over reading instruction despite the large existence of reading research. According to this organization, interpretations of the research vary greatly, which results in multiple ineffective practices. Many educators rely on phonics instruction that is incomplete and random, while others use practices that completely leave out application to authentic reading and writing (ILA, 2019). As educators skirmish to find consensus in research, students struggle to find the path to reading proficiency.

In the spring of 2020, the COVID-19 global pandemic shut down schools nationwide and caused significant disruptions in the national education system. Over the next 2 years, schools would toggle between full closures, partial in-person instruction, and remote learning. Hammerstein et al. (2021) found that the remote learning that took place during COVID-19 had a similar effect of learning loss associated with summer closures. Not only did the pandemic change the delivery of instruction, but the focus and time spent on reading instruction were significantly reduced (Crosson & Silverman, 2021; Hammerstein et al., 2021). The disruptions had an immediate and long-term negative impact on learning (Hammerstein et al., 2021; Molnar & Hermann, 2023). Students in the primary grades, where early literacy instruction is critical, suffered the most significant learning loss (Crosson & Silverman, 2021; Hammerstein et al., 2021; Molnar & Hermann, 2023). Still, there have been other studies showing that students in grades beyond third grade suffered from learning losses (Kuhfeld et al., 2023; Relyea et al., 2023).

A flurry of legislative laws erupted nationwide to address the learning loss associated with the pandemic. Many of these laws were reading-specific and focused on the basic literacy skills defined by NRP (Neuman et al., 2023). As the regulations took effect, the correct way to teach reading became a public debate once again. The discussion was fueled by widespread media coverage, including newspapers, journals, and podcasts, which criticized balanced literacy and called for a renewed focus on phonics instruction and the science of reading (Neuman et al., 2023; Sanden et al., 2022; Shanahan, 2020). The current discourse around the science of reading strongly shapes policy and instruction in education (Semingson & Kerns, 2021).

The NAEP assessment is given biennially to fourth graders to assess reading ability. These results consistently show that approximately one third of students are reading below a fundamental level. According to the 2022 NAEP results, 39% of North Carolina's fourth graders were reading below the basic standard. Students reading at this level lack word recognition and decoding skills and exhibit poor reading fluency. These students lack the skills needed to fully participate in the global 21st century world and require direct intervention in these skill areas (Hiebert, 2022; White et al., 2021).

The Science of Reading

According to Shanahan (2020), the term science of reading has been used in education for more than 200 years; however, the term has reached newfound popularity over the last few years. The Reading League (2022) reported that the increased use of the term has led to misconceptions and published a document dedicated to providing a detailed definition that would lead to consensus in the field. The Reading League noted that specifying a common definition would aid the educational transformation to effective reading practice. The guide defines the science of reading as an immense interdisciplinary body of research based on science that explains reading. One of the prominent positive elements emerging from the definition is that the science of reading includes research from experts in numerous disciplines, which results in varied opinions and multiple solutions to the reading crisis (Solari et al., 2020). The Reading League went on to explain the science of reading has roots in the SVR (Gough & Tunmer, 1986) and Scarborough's (2001) Reading Rope.

The push for the science of reading comes primarily from the continuous low literacy scores across the nation (Seidenberg et al., 2020). In addition, the popularity of

the science of reading is attributed to Seidenberg's (2017) book, *Language at the Speed of Sight*, which describes learning to read through a vast amount of research based on cognitive and neurosciences (Shanahan, 2020). Other sources, such as Hanford's (2018, 2019) public journalism, also contributed to the rise in attention. In fact, Riley (2020) and Shanahan (2020) proclaimed that American Public Media's reports by Hanford (2018, 2019) reignited the reading wars. Hanford's (2018, 2019) reports profiled schools with failing readers and brought attention to the bleak NAEP scores. Hanford (2018) praised Seidenberg and pointed to the science of reading as a solution to the reading crisis. Hanford (2019) highly criticized the work of Goodman (1967) and the popular three-cueing system used in schools. There were many critics who pushed back, stating that noneducators and journalists are not credible reading scholars, which fueled opposition to the science of reading (Shanahan, 2020).

In 2022, Hanford continued to draw national attention to the reading crisis with her last work titled "Sold A Story." In this series of podcasts, Hanford declared that educators were sold a flawed idea about teaching reading by large publishers who made an enormous profit. As a result, students have been taught to read incorrectly for many years (Hanford, 2022).

There are some key takeaways gleaned from the science of reading. First, reading is a complicated act that is not natural (Hoover & Tunmer, 2020; Moats, 2020a, 2020b; Seidenberg, 2017; Willingham, 2017). Next, explicit instruction in decoding is essential to learning to read, and students must become fluent in the process (Castles et al., 2018; Moats, 2020a; Seidenberg, 2017). Finally, comprehension strategies should be modeled and taught explicitly to students (Duke & Ward, 2021; Hennessy, 2021; Riley, 2020).

Over the last few years, many states have adopted the science of reading into their educational systems. North Carolina passed EPSA in 2021. This legislation mandated schools provide professional development in the science of reading and ensure that it is incorporated into instructional practice. EPSA (2021) also stated that teacher preparation programs train teachers in the science of reading.

Seidenberg et al. (2020) welcomed the renewed focus on the science of reading but noted that educators should cautiously embrace the movement. First, there is a need for translation of the research. Teachers must understand the research to apply it to instructional practice (Seidenberg, 2017; Seidenberg et al., 2020). Second, the science of reading may be oversimplified by educators, leading to the support of instructional practices that are only vaguely connected to science. Finally, the science of reading is a vast amount of research constantly growing and changing, and educators must diligently keep up with the changes (Seidenberg et al., 2020).

Professional Development

Professional development is also referred to as professional learning and is used to increase teacher knowledge and improve pedagogical skills. Throughout history, instructional programs have failed to improve student achievement, leading many states to push for reform and adopt new practices (Glatthorn et al., 2019). Such reforms require educators to learn new skills and adopt current research-based practices.

Mandated professional development initiatives have failed in the past (Hall & Hord, 2020; Tassell, 2014); however, mandates are frequently used in reform efforts. They set clear expectations and remove ambiguity around implementation. Mandates automatically create a surge in personal concerns, and failure to pay attention to these

concerns throughout the process will result in an unsuccessful initiative (Hall & Hord, 2020). Furthermore, top-down approaches to professional development often fail because they do not consider each educator's varying levels of knowledge (Knight, 2021). Effective professional development demands attention to these issues that are rooted in change and require a high level of support (Hall & Hord, 2020).

Effective professional development can create gains in student reading proficiency. Hattie (2009) reported that professional development had a moderate impact on student achievement, according to his analysis. Basma and Savage (2018) conducted a meta-analysis of studies relating professional development to student reading achievement. Basma and Savage's (2018) review found that of 17 studies examining the correlation between professional development and student achievement, 10 found a significant positive increase in student achievement; however, Basma and Savage (2018) could not establish that longer professional development had better results due to a lack of studies involving sustained length. The literature review produced two more meta-analyses regarding professional development and student reading achievement. Didion et al. (2020) examined studies primarily focused on elementary school. Didion et al. reported that student achievement was higher for students whose teachers participated in professional development as compared to students whose teachers did not. Later, Basma and Savage (2023) completed another meta-analysis and found that professional development had a significant positive effect on reading achievement in middle and high school.

Elements of Effective Professional Development

Ineffective professional development is attributed to training that is brief in

nature, impersonalized, and evaluative. Reeves (2021) noted that it is imperative these elements of professional development are eliminated to assist teachers in a post-pandemic world. Professional development must now consider the individual needs and knowledge base of adult learners (Reeves, 2021). Darling-Hammond et al. (2017) suggested that educators regularly assess the needs of teachers through surveys and ensure professional development is aligned with what teachers feel they need most.

In addition to personalized learning, Darling-Hammond et al. (2017) conducted an extensive review of professional development studies and created a set of criteria needed for professional development to be productive. According to the review, high-quality professional development is thorough and content-focused. It involves active learning and supports collaboration. Teachers are provided with exemplar models to follow. A successful model incorporates coaching and allows for feedback and reflection. Finally, the professional development must be of sustained duration. Pieces of training brief in nature are insufficient (Darling-Hammond et al., 2017).

Darling-Hammond et al. (2017) elaborated on the criteria. Professional development should be aligned with content and pedagogies of specific disciplines such as literacy. There should also be activities that call for active participation and allow teachers to practice the strategies they are learning. Professional development that is defined by lectures does not connect the learning to teacher classrooms (Darling-Hammond et al., 2017).

High-quality professional development supports collaboration. Collaborative communities change cultures, build trust, and boost student achievement (Darling-Hammond et al., 2017; DuFour et al., 2016). In addition, professional development

should provide teachers with models of effective practices. Teachers should be able to view exemplar strategies as they collaborate (Darling-Hammond et al., 2017).

Professional development should be constant and include expert coaching and support (Darling-Hammond et al., 2017; Reeves, 2021). Effective coaching creates deeper relationships that have a powerful impact on student achievement. These types of supported relationships can transform schools and enact positive change (Knight, 2018).

Finally, effective professional development is driven by feedback and reflection. As teachers reflect and solicit expert feedback, they can refine and change their practice (Darling-Hammond et al., 2017). Feedback should be intentional and clarify understandings (Knight, 2018). Students also directly benefit when teachers internalize expert feedback and utilize those strategies in the classroom (Hattie & Clarke, 2019).

Scarparolo and Hammond (2018) conducted a study to measure the impact of an evidenced-based professional model of training. The study included providing teachers with instruction in the explicit teaching of decoding and included modeling, feedback, and coaching. Teachers were administered a knowledge survey before and after the training. There was a statistical increase in knowledge following the training. Scarparolo and Hammond also noted that classroom observations showed an improvement in classroom instructional practices.

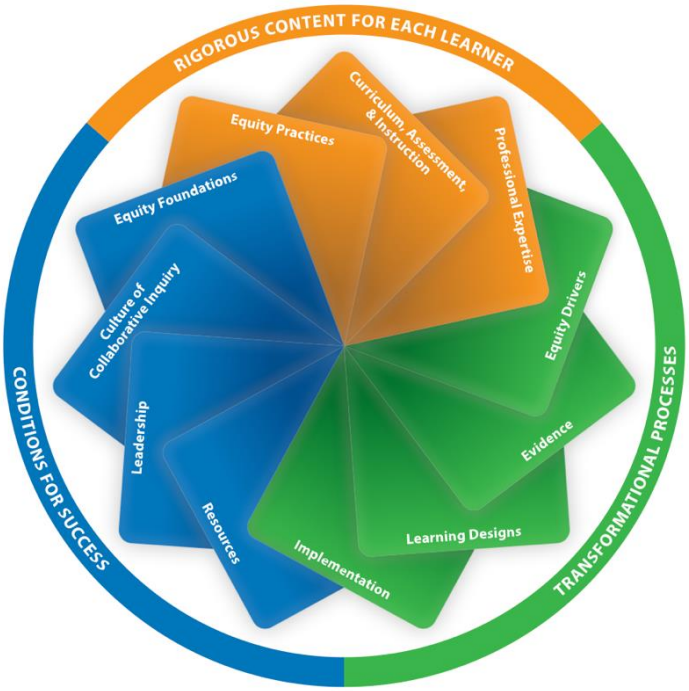
Learning Forward's Standards for Professional Learning

In 2022, the Professional Learning Association, Learning Forward, revised its Standards for Professional Learning. The Standards for Professional Learning were built on extensive research. They were designed to provide a comprehensive model for educators and stakeholders involved in professional learning (Foster, 2022). These

standards contain three overarching categories: rigorous content for each learner, transformational processes, and conditions for success. Each category contains standards to further guide educators and stakeholders. Figure 7 illustrates the categories and related standards.

Figure 7

Standards for Professional Learning



Note. Image from Crow (2022).

Figure 7 displays the standards as overlapping parts. There are three standards under the Rigorous Content for Each Learner frame. The standards in this category are equity practices, curriculum assessment, and professional expertise. This category focuses on content for adult learners and varies depending on the expertise of each learner. This category asserts that teachers should understand student culture and nurture strong relationships with families. Additionally, teachers should aim to implement high-

quality curricula and materials. Finally, teachers should develop expertise by applying standards and research to their learning (Crow, 2022).

The next category, Transformational Processes, contains four standards: equity drivers, evidence, learning designs, and implementation. These standards state that professional learning should include learners' cultures and beliefs. Professional learning should be planned and designed using multiple data sources. In addition, learning should be aligned with evidence about best practices and include goals and priorities. Lastly, this category denotes that professional learning should be sustained over time and incorporate feedback and follow-up (Crow, 2022).

The standards of the Conditions for Success frame are equity foundations, culture of collaborative inquiry, leadership, and resources. Standards in this category focus on system-wide change. They stress establishing expectations and equitable conditions for learning. The importance of collaboration is also emphasized. In addition, these standards call for intentional leadership that prioritizes and maintains resources for professional learning (Crow, 2022).

LETRS

Moats and Tolman (2019) maintained that LETRS training includes many of Learning Forward's professional standards. The training is designed to be collaborative and contains built-in journals for reflection. In addition, the training continues over an extended period of time; however, the professional development does not provide much personalization due to the mandated requirement that all K-5 reading teachers participate regardless of their background knowledge or prior training. Moats and Tolman stated that LETRS is founded on the science of reading and the basic early literacy skills identified

by NRP (2000). Much of the work of Moats is associated with code-based phonics instruction, and she is often cited in research pertaining to the science of reading (Semingson & Kerns, 2021). In addition, several state education agencies have indicated that LETRS has been instrumental in their teacher preparation programs (Semingson & Kerns, 2021).

North Carolina planned to deliver LETRS professional development in cohorts throughout the state. The state divided the implementation for schools into three cohorts representing districts from all regions in the state. The implementation of LETRS began with Cohort 1 in the fall of 2021. The implementation plan ends with Cohort 3 finishing in the spring of 2024. The data from the planned Cohort 1 were used in this study. Table 1 shows the original regions and districts across the state that comprise the first planned cohort to complete the professional development.

Table 1

School Regions and Districts of Cohort 1

Region	District
North Central	Edgecombe
	Granville
	Nash
	Vance
	Warren
	Wilson
	Durham
Northeast	Bertie
	Martin
	Northampton
	Washington
Northwest	Burke
	Catawba
	Newton
	Wilkes
Piedmont-Triad	Mount Airy
	Stokes
	Winston Salem/ Forsyth
Sandhills	Bladen
	Clinton
	Columbus
	Hoke
	Scotland Whiteville
Southeast	Duplin
	New Hanover
	Onslow
Southwest	Anson
	Cabarrus
	Charlotte-Mecklenburg
	Cleveland
	Mooresville
	Union
Western	Asheville
	Clay
	Haywood
	Madison
	McDowell
	Polk

Table 1 indicates the public school districts represented by Cohort 1. The cohort

samples the eight geographical regions designated by the state. Cohort 1 represents approximately one third of the 116 school districts across the state.

Each cohort was to complete two volumes: Volume 1 in Year 1 and Volume 2 in the second year. The training consisted of eight self-paced online modules or units estimated to take 58 hours to complete and 26 hours for additional readings. The modules also contained Bridge to Practice activities designed to apply the knowledge obtained to classroom instruction. The time to complete these activities was estimated at approximately 25 hours (Moats & Tolman, 2019). In addition, each of the eight units required a 1-hour live training session.

The two manuals contain eight units each. Volume 1 focuses on the word recognition strand of Scarborough's (2001) Reading Rope. This strand concentrates on the foundational reading skills that are typically developed in the early years. Volume 1 begins with background information and research correlating to the science of reading. The SVR and Scarborough's Reading Rope are discussed in detail at the beginning of the first volume. The first volume contains four units. Each unit is comprised of eight sessions. Table 2 presents an overview of the material covered in Volume 1.

Table 2

Overview of LETRS Volume 1 Content

Units	Topics covered
Unit 1: The Challenge of Learning to Read	Causes of reading difficulties, research on learning to read, reading assessments
Unit 2: The Speech Sounds of English	Phonological skill development, teaching and assessing phonological skills
Unit 3: Teaching Beginning Phonics, Word Recognition, and Spelling	Essentials of early reading instruction, English orthography, assessing phonics and word recognition
Unit 4: Advanced Decoding, Spelling, and Word Recognition	Morphology, spelling, fluency

Table 2 shows an outline of the information from Volume 1. This volume focuses on word recognition skills, including phonological awareness and phonics. Unit 1 contains research on reading theories, including an overview of the SVR (Gough & Tumner, 1986) and Scarborough's (2001) Reading Rope. Additionally, teachers are taught how to utilize several reading assessments. Moats and Tolman (2019) stated that the course incorporates research from the disciplines of cognitive psychology, neuroscience, and linguistics. Much of this research is embedded in Unit 1.

The technical tone of Unit 1 continues in Unit 2. This unit focuses on the role of the phonological processing system in learning to read. The continuum of phonemic awareness skills and how to instruct at each level are presented. The unit is heavy in linguistics. The phonemes of the English language and various dialects are described.

Unit 3 and Unit 4 dive deep into the word recognition strand of Scarborough's (2001) Reading Rope. There is an emphasis on decoding and phonics instruction, also

referred to as code-emphasis instruction (Moats & Tolman, 2019). The process of building a sight word vocabulary is presented. In addition, these units link morphology to spelling conventions. Volume 1 ends with a rationale for the structured literacy approach (Spear-Swerling, 2018) and the importance of explicit and systematic instruction.

Volume 2 shifts the focus to the language comprehension strand of Scarborough's (2001) Reading Rope and the related subskills (Moats & Tolman, 2019). This volume is made up of four units comprised of six sessions per unit. Table 3 presents an overview of the content found in the units and sessions of this volume.

Table 3

Overview of LETRS Volume 2 Content

Units	Topics covered
Unit 5: The Mighty Word: Oral Language and Vocabulary	Explores background knowledge, types of vocabulary, and instructional routines
Unit 6: Digging for Meaning: Understanding Reading Comprehension	Comprehension of products and processes, poor comprehension causes, syntax, and text structures
Unit 7: Text-Driven Comprehension Instruction	Effective comprehension strategies, lesson planning, and special needs populations
Unit 8: The Reading-Writing Connection	Reading-Writing link, types of writing, and instructional strategies

Table 3 outlines the material in Units 5-8. This volume focuses on the relationship between language comprehension and reading comprehension. Unit 5 explores research supporting the need for vocabulary development and the negative impacts of inadequate development (Hennessy, 2021). The unit explains the differences in receptive and expressive vocabulary. Effective vocabulary strategies and routines are presented.

Unit 6 and Unit 7 concentrate on reading comprehension. The cognitive skills

supporting comprehension are explained. Moats and Tolman (2019) explore the causes of comprehension difficulties and present effective instructional strategies. A lesson plan for teaching comprehension strategies is provided, and the needs of special populations are described.

Unit 8 ends Volume 1 with a focus on how to teach writing. The connections between writing development and reading development are explored. Narrative and informational writing strategies are presented. The stages of the writing process are discussed as well as writing assessments.

Research Regarding LETRS

A review of the research produced two published studies regarding LETRS. First, the Institute of Educational Sciences (IES) funded a randomized control trial comparing a group of teachers who received an eight-day seminar training in LETRS to a group who did not receive the training. In this study, Garet et al. (2008) found an increase in teacher knowledge of the treatment group. However, the study reported that there was no statistical increase in student achievement on the district literacy assessments.

The Institute of Educational Sciences later funded a larger study of the state-wide implementation of LETRS in Mississippi. Teachers completed eight modules from the second edition of LETRS. Teacher knowledge was assessed through the Teacher Knowledge of Early Literacy Skills survey. The instructional practice was measured through classroom observations. Folsom et al. (2017) found that LETRS professional development increased Mississippi's K-3 teacher knowledge of early literacy practices and positively impacted instructional practices in the classroom. Moreover, teacher knowledge increased with the professional development teachers received (Folsom et al.,

2017).

ESSA (Dennis, 2017) stressed that educational systems use evidence-based strategies and programs that are grounded in empirical research (Folsom et al., 2017). ESSA defined and characterized evidence-based practices and programs by tiers. Each tier requires that the impact on student achievement be statically significant. Tier 1 is characterized as showing strong evidence, Tier 2 practices show moderate evidence, and Tier 3 is defined as showing promising evidence. LETRS professional development is a Tier 3 program (Folsom et al., 2017).

Assessment

Assessments of all kinds are used to measure student progress and mastery of skills. When used appropriately, assessments are essential to effective evidence-based reading programs (Hougen & Smartt, 2020). Formative assessments are informal and generally teacher-created. These types of assessments measure student understanding of concepts and guide teacher instruction in the classroom (Hougen & Smartt, 2020). Summative assessments determine knowledge of skills and standards. Standardized tests used by states are a form of summative evaluation. Stiggins (2018) noted that assessments should gather information on student achievement and use it to make informed educational decisions.

Most expectations for student learning and achievement can be directly related to curriculum. Glatthorn et al. (2019) noted that professional learning programs can be evaluated and analyzed for success. Assessing student growth allows insight into curriculum mastery as well as the effectiveness of teaching practices (Glatthorn et al., 2019).

North Carolina Assessment and Accountability

Assessments, however, may have an additional purpose. The NCLB Act of 2001 placed a great deal of emphasis on standardized tests (Backstrom, 2019; Dennis, 2017). Teachers were held accountable for how their students performed. These high-stakes tests contributed to teacher stress and created fragmented instruction (Ainsworth & Donovan, 2019). ESSA called for a reduction in standardized testing; however, these tests still carry much weight among states and educators (Dennis, 2017).

In 2012, North Carolina established the RtA. The RtA created a comprehensive reading policy focusing on early literacy improvement for students in Grades K-3. Citing research that struggling readers in third grade were much more likely to become high school dropouts, RtA's primary goal was to have all students reading on grade level by the end of Grade 3 (Hensley et al., 2017). This legislation requires that third-grade students demonstrate proficiency on the EOG reading assessment. Alternatively, third graders can show mastery on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment or through a test-based reading portfolio (Hensley et al., 2017).

At the end of the 2013-2014 school year, North Carolina began issuing school performance grades also known as a report card for each school. The report card gives a letter grade ranging from A to F based on standardized EOG test results and combined student growth. The grade includes overall performance on EOG tests and the amount of growth students are expected to show from year to year; however, the distribution is not equal. Twenty percent of the grade comes from the growth score, while 80% comes from the overall score of the tests.

North Carolina uses software based on a value-added model to assess student

growth as well as teacher impact on student learning. Value-added models are a controversial way of using test data to gauge the effectiveness of individual teachers (Glatthorn et al., 2019). Despite little evidence to support the validity of value-added models (Glatthorn et al., 2019), the performance of North Carolina's teachers and students is still linked to them.

Although standardized tests pressure educators and students through evaluation, NCDPI maintained that the tests are valid, reliable, and fair (Howard, 2020). Howard (2020 noted that the assessments are based on the Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014). These assessments measure what students know and can do and measure progress (Howard, 2020).

Primary Literacy Assessments in North Carolina

North Carolina uses multiple assessments in grades kindergarten through third grade to measure literacy growth and proficiency. Teachers administer both formative and summative assessments. All students in Grades K-3 are given the eighth edition of DIBELS. Students in third through fifth grades are also given a standardized EOG summative assessment.

DIBELS. DIBELS began as an assessment of early literacy skills in kindergarten and first grade. The assessment was inspired by the curriculum-based measurements defined by psychologists in 1986. Curriculum-based measures are used to monitor student learning and are also considered general outcome measures (Hougen & Smartt, 2020).

DIBELS subtests are criterion-referenced, evaluating students against a set of

predetermined benchmarks or standards. Beginning in 1996, DIBELS has released updated editions every several years. The latest edition, DIBELS 8, expands on decades of reading research to offer a valid and reliable measure of reading development for Grades K-8 (University of Oregon, 2021).

DIBELS 8 assesses the early reading domains identified by the NRP (2000) report and connects with the Cognitive Reading Framework (Hoover & Tunmer, 2020). The assessment produces an overall composite score that indicates each student's level of risk for reading difficulty. The composite scores are categorized by well-below benchmark, below benchmark, at benchmark, and above benchmark. The benchmark score represents students on track to be reading on grade level.

The assessment consists of a series of six subtests. Five subtests are short, 1-minute fluency measures, and the sixth subtest is a 3-minute measure of reading comprehension. Automaticity and fluency in early reading skills are necessary for the development of proficient reading and adequate comprehension. When a reader processes text with automaticity and fluency, a larger portion of the cognitive space and working memory are available for comprehension (Ehri, 2014; Lane & Contesse, 2022). Table 4 shows each DIBELS 8 subtest, the associated skill, and the grade level administered.

Table 4

DIBELS 8 Subtests and Aligned Skills and Grade Levels

Subtest	Skill assessed	Grade level
Letter naming fluency (LNF)	Letter knowledge	K-1
Phonemic segmentation fluency (PSF)	Phonemic awareness	K-1
Nonsense word fluency (NWF)	Alphabetic principle	K-3
Word reading fluency (WRF)	Accuracy and fluency of text	K-3
Oral reading fluency (ORF)	Accuracy and fluency of text	1-3
Maze	Comprehension	2-3

Table 4 shows the connection between the assessment and reading research. DIBELS covers four of the five essential reading components identified by NRP (2000): phonemic awareness, alphabetic principle, fluency, and comprehension. The change in the administration of the subtests across grade levels mirrors the continuum of reading development (University of Oregon, 2021).

Composite Score. DIBELS 8 generates a composite score to represent the overall reading status for students at each grade level. The composite score is derived from the raw scores of the subtests. The subtest scores are based on benchmark cut scores that were set through a national norming process. The benchmark cut scores increase throughout the year for each grade level. All scores fall into one of four categories: well below benchmark, below benchmark, at benchmark, or above benchmark. These categories are color-coded for teachers. The color red marks well below benchmark scores, below benchmark scores are yellow, benchmark scores are green, and those scores that fall above benchmark are blue.

Students performing below benchmark cut scores show poorer performance in the early reading skills and are considered at risk for reading difficulties, including dyslexia. The composite scores were determined through a confirmatory factor analysis

approach, where relations among subtests are accounted for, and each test is optimally weighted (University of Oregon, 2021). Table 5 illustrates how each subtest contributes to the overall composite score. These subtests are directly linked to the components of language comprehension and word recognition.

Table 5

Composite Score Calculations

Grade	Subtest	Weight
Kindergarten	LNF BOY	35.44
	LNF MOY/EOY	8.86
	PSF	4.13
	NWF-Correct Letter Sounds	14.93
	NWF-Words Recoded Correctly	3.56
	WRF	5.62
First	LNF	10.72
	PSF	2.13
	NWF-Correct Letter Sounds	23.13
	NWF-Words Recoded Correctly	7.79
	WRF	13.51
	ORF- Words Correct Per Minute	25.36
	ORF-Accuracy	0.25
Second	NWF-Correct Letter Sounds	32.74
	NWF-Words Recoded Correctly	10.95
	WRF	21.26
	ORF-Words Correct Per Minute	35.36
	ORF-Accuracy	0.15
	Maze	4.28
Third	NWF- Correct Letter Sounds	40.02
	NWF- Words Recoded Correctly	11.80
	WRF	19.83
	ORF- Words Correct Per Minute	39.42
	ORF-Accuracy	0.09
	Maze	4.79

Note. BOY = beginning of year; MOY= middle of year; EOY = end of year; LNF = letter naming fluency; PSF = phoneme segmentation fluency; NWF = nonsense word fluency;

ORF = oral reading fluency.

The weights for each subtest indicated in Table 5 are consistent across each time of year the assessment is given, except for letter naming fluency (LNF). LNF is weighted more at the beginning of the year than it is at the middle of the year and end of the year. Nonsense word fluency (NWF) and oral reading fluency (ORF) produce two scores. For NWF, there is a score for correct letter sounds and a score for words read or recoded correctly. ORF generates a fluency score for words read correctly per minute and a score for accuracy. According to the University of Oregon (2021), NWF-correct letter sounds and ORF-words correct per minute are stronger predictors of reading ability and thus carry significantly more weight than their counterpart scores. This weight is significant when interpreting composite scores as well as individual subtest scores.

LNF. LNF is a standardized test that measures risk for reading achievement. It is administered individually from the start of kindergarten through the end of first grade. Each LNF page contains 100 upper and lowercase letters randomly arranged. Kindergarten forms contain the 40 most frequently seen alphabet letters. First-grade forms contain all of the alphabet letters with the exceptions of upper and lowercase W, and lowercase L. W was omitted because it is a letter made up of three syllables, which takes longer to pronounce. Lowercase L was eliminated because students often confuse this letter with the letter I or the numeral 1. Students are given 1 minute to identify as many letter names as possible (University of Oregon, 2021).

Phonemic Segmentation Fluency. Phonemic segmentation fluency (PSF) is a standardized assessment and is individually administered. PSF measures phonemic awareness, which falls under the phonological awareness umbrella and is a strong

indicator of reading difficulties (Hoover & Tunmer, 2020; Kilpatrick, 2015, 2016; Moats, 2020a; Moats & Tolman, 2019); however, PSF is considered a limited measure of phonemic awareness because it only assesses segmentation and not the other skills across the phonemic awareness continuum (Hoover & Tunmer, 2020). Like LNF, PSF is given at the beginning of kindergarten and throughout first grade. During the PSF assessment, students must segment words into their individual phonemes in 1 minute. The words are presented orally to the student by the assessor. To minimize the effect of vocabulary, the final selection of words comes from the 2,500 most frequent words in English. This word list was established by the English Lexicon Project (Balota et al., 2007). Words that are more frequent appear earlier in the assessment, with less frequent words appearing later. Moreover, the words progress in difficulty, beginning with two phonemes and increasing to terms with six phonemes (University of Oregon, 2021).

NWF. NWF is another standardized, 1-minute, individually administered assessment. This assessment measures a student's ability to understand the alphabetic principle or phonics skills. Administration of NWF begins at the start of kindergarten and continues through the end of third grade. Students are given a sheet of made-up words, which they can segment by phonemes or read the whole word. A score is provided for both approaches to reading the words. NWF-correct letter sounds reflects the number of sound segments produced correctly, and NWF-words recoded correctly reflects the number of whole words read correctly. The forms reflect phonetically regular letter combinations found in English. For example, there are no-nonsense words that end in v or j.

Difficulty increases across each form. Kindergarten forms contain only the

consonant vowel consonant pattern, with the most frequent letter combinations appearing at the beginning. In addition to increasing difficulty by frequency of letter combinations, forms at other grade levels also add additional phonetic patterns. Assessing nonsense word reading is considered a "pure" approach to measuring student phonics skills because vocabulary and sight word knowledge are removed (University of Oregon, 2021). Good readers can read nonsense words at a similar rate to reading actual words (Kilpatrick, 2015).

Word Reading Fluency (WRF). WRF is another standardized, individually administered assessment. Frequency was also a consideration in determining the pool of words for this measure. Four different word lists were used to create the final selection of words. The 1-minute evaluation measures student ability to read actual words in isolation. The assessment measures how fluently students decode or read high-frequency words. This subtest contains a combination of decodable and irregular words that show up most often at a particular grade level. Students are given a page of words and are asked to read as many as possible. Unlike NWF, no credit is given for segmenting a word by phonemes. The subtest measures fluency and accuracy. Students are given one score for the number of words read correctly in 1 minute (University of Oregon, 2021).

ORF. ORF is administered at the beginning of first grade and continues through the end of third grade. This standardized assessment measures student ability to read real words in the context of a passage. Passages were written by experts from a variety of backgrounds. In Grades 1-5, narrative text structure accounts for 60% of the passages, while nonfiction text structure constitutes the remaining 40%. During the development of the passages, measures were taken to reduce bias and create consistency in the difficulty

level across passages at each grade level. This assessment is another fluency subtest measuring basic to advanced phonics skills. Students are given a score for the number of words read correctly per minute, including self-corrected words. The measure also provides an accuracy score. This score reflects the percentage of correct words read of the total number of read words. For example, a student may read 110 words per minute with 90% accuracy (University of Oregon, 2021).

Maze. Like ORF, Maze passages were written by a diverse group of experts. Additionally, measures were taken to reduce bias within the passages; however, Maze contrasts with the other subtests in a couple of ways. First, students take this assessment online without an assessor.

Additionally, students are given 3 minutes to complete the assessment. This assessment is not available in kindergarten or first grade. It is administered at the beginning of second grade through the end of second grade. Students must read a passage where every seventh word in a sentence is removed. Three word choices are presented, and students must choose the appropriate word to replace the missing word. In second grade, the first two sentences and the last sentence in each paragraph are left intact. In third grade, the first and last sentences are left intact. A score is calculated by subtracting half of the number of errors from the number of correct responses.

EOG. The EOG is administered at the end of the year for Grades 3-5 to assess reading comprehension. Comprehension is dependent upon mastery of the early literacy skills assessed through DIBELS (Hoover & Tunmer, 2020). Vernon-Feagans et al. (2019) examined the quality of instruction in early literacy classrooms. Students who were successful in these classrooms performed higher on reading comprehension passages in

third grade. Another study accurately identified 88.7% of students who would pass the EOG assessment in third grade by the proficiency of DIBELS assessments in previous grades (Smith et al., 2020). The correlation between the assessments demonstrates the movement from mastery of foundational literacy skills to mastery of reading comprehension.

The EOGs in North Carolina assess comprehension based on literacy standards. North Carolina adopted its current standards in 2017. The literacy standards are divided into the domains of Speaking and Listening, Reading Foundations, Reading for Literature, Reading for Informational Text, and Language. The Reading Foundations and Speaking and Listening standards can be mapped onto the Cognitive Foundations of Reading Framework to illustrate the connection to the strands. These two types of standards support the standards from the other domains (Hoover & Tunmer, 2020).

The third- through fifth-grade EOGs assess standards from Reading for Literature, Reading for Informational Text, and Language. In these grades, 38% to 42% of the assessment items are based on the Reading for Literature strand, 46% to 50% correlate to the Reading for Informational Text strand, and the remaining 13% to 15% of test items represent the Language strand. The weight distribution across the domains for test items is the same for all three of these grade levels.

EOG scores are divided into five levels. A score falling in Levels 1-2 is considered not proficient. Scores that fall into Level 3 are said to be grade-level proficient. Levels 4 and 5 are categorized as college and career ready. This study analyzed data at Level 3. For all grade levels, a Level 3 score indicates that students have an adequate understanding of grade-level content. Just as reading development is

expected to change through progression from year to year, the criteria for Level 3 increase somewhat from third grade to fifth grade. Table 6 indicates the differences for Level 3 proficiency at each grade.

Table 6

Level 3 Achievement Descriptors for Grades 3-5

Domain	Grade 3 criteria	Grade 4 criteria	Grade 5 criteria
Reading for literature	Identify words that impact meaning	Identify words that impact tone	Understand word choices that contribute to meaning, and tone
Reading for informational text	Determine the meaning of academic and domain-specific Grade 3 words, and phrases	Determine the meaning of academic and domain-specific Grade 4 words, and phrases	Determine the meaning of academic and domain-specific Grade 5 words, and phrases
Language	Distinguish the literal and/or nonliteral meanings of words, and phrases	Understand figurative language, similes, and metaphors	Interpret language, similes, and metaphors

Table 6 does not list all the criteria that are required for proficiency. Rather, the primary skills that can be directly related to the word recognition or language strands of the Cognitive Foundations of Reading Framework (Hoover & Tunmer, 2020) are shown.

Webb's Depth of Knowledge

Webb's depth of knowledge (DOK) is a scale that is often used in rigorous curriculum and assessment design (Ainsworth & Donovan, 2019). The EOGs were created using Webb's DOK. In the DOK model, there are four levels. The four levels increase in the complexity of cognitive processes. Level 1 demands recall and reproduction. Level 2 is related to skills and concepts. Level 3 ramps up cognitive requirements to include strategic thinking and reasoning. Finally, Level 4 requires

extending thinking and application. The EOGs contain questions from Levels 1-3. Table 7 depicts the distribution of questions across the levels.

Table 7

EOG Assessment Items by DOK Distribution

Grade	DOK Level 1	DOK Level 2	DOK Level 3
3	20-40%	60-80%	N/A
4	12-25%	50-75%	5-10%
5	N/A	75-90%	10-25%

Table 7 shows the percentage of questions at each of the DOK levels for Grades 3-5. The largest percentage of questions are written at DOK Level 2 for each grade level. The Grade 4 EOG contains no questions at Level 3, while the Grade 5 EOG contains no questions written at Level 1; therefore, the EOGs increase in complexity somewhat from grade to grade.

Literacy Achievement Gap

When looking at literacy proficiency data, it is important to note the achievement gap that exists between gender, socioeconomic status, and race. The term achievement gap usually refers to the disproportions in academic performance between White students and minority populations (Seidenberg, 2017). Achievement gaps in education were made apparent when NCLB required states to report assessment scores of subgroups annually (Backstrom, 2019).

While gaps exist across all subjects, the achievement gap in reading is particularly profound and complex (Horowitz & Samuels, 2017). Standardized tests have shown that this gap in performance has existed for many decades, and it has not narrowed despite great effort (Dintersmith, 2018; Merga, 2020; Seidenberg, 2017). In fact, as

students move through the grades, the achievement gap tends to widen. The widening achievement gap contributes to economic and educational inequalities (Merga, 2020).

Impact of the COVID-19 Pandemic

The COVID-19 pandemic contributed to the literacy gap. During the 2019-2020 academic year, face-to-face learning was disrupted as many schools transitioned to remote learning. The impact of the pandemic is still being assessed. Although evaluating the impact has been difficult, there is no doubt the learning loss left an indelible mark on academic achievement (Almasi & Yuan, 2023; Kuhfeld et al., 2020).

Multiple studies in the United States indicated that the pandemic had a negative impact on reading comprehension (Crosson & Silverman, 2021; Curriculum Associates, 2021; Donnelly & Patrinos, 2022; König & Frey, 2022). König and Frey (2022) found that the decline in reading comprehension caused by the pandemic was similar to that of summer learning loss. Not surprisingly, younger learners, minority groups, and students from economically oppressed backgrounds were impacted the most (Crosson & Silverman, 2021; Curriculum Associates, 2021; Donnelly & Patrinos, 2022; Hammerstein et al., 2021; Molnar & Hermann, 2023).

In a recent publication, Kuhfeld et al. (2023) used scores on the Northwest Evaluation Association MAP assessment to determine how student achievement had been impacted by the pandemic in Grades 3-8. This study compared fall 2019, 2020, and 2021 assessment scores. The data from the fall of 2019 were labeled as pre-pandemic and used as a baseline. Kuhfeld et al. (2023) found that students in Grades 3-5 showed more significant declines in growth scores compared to students in Grades 6-8. In addition, students from high-poverty schools and students of color were particularly impacted.

Amplify (2022) compared data from the 2019-2020 school year to the 2020-2021 school year data. This study analyzed K-5 student scores with the DIBELS 8 assessment. Amplify's results found that the percentage of students on track to reading proficiency had significantly dropped in Grades K-2; however, in Grades 3-5, students were making adequate progress. Like Kuhfeld et al. (2023), Amplify found that minority students showed the largest declines in scores. Moreover, the results of this study indicated that the achievement gap between students of color and White students had widened.

Summary

For decades, educators have worked to end the reading crisis. Experts have touted various instructional methods as best practices in raising student performance. Proponents have argued for phonics-based approaches, while others maintain a balanced literacy method is best (Castles et al., 2018). In recent years, instructional practice has been pushed to focus on the science of reading. The science of reading includes decades of research and the NRP (2000) report. This report shows that reading instruction must focus on basic early literacy skills. The Cognitive Foundations of Reading Framework (Hoover & Tunmer, 2020) was founded on the SVR and guides the understanding of reading development.

Assessments provide an avenue to measure growth and proficiency in reading. In addition, assessments aid in determining the effectiveness of curriculum (Glatthorn et al., 2019). Analysis of assessments has shown that there are gaps between subgroups, and these gaps may have worsened because of learning loss during the global pandemic.

North Carolina attempted to align literacy instruction with the science of reading and address learning loss by mandating LETRS training across the state. Through LETRS

professional learning, teachers devote many hours to learning how students learn to read and what research shows is best practice (Moats & Tolman, 2019). The early literacy assessments used in the state of North Carolina include the DIBELS assessment and the EOG summative assessments. These assessments are directly aligned with basic early literacy skills, and results taken across the year show how students grow in each area. Data collection can show how curriculum and instructional changes impact student growth and achievement (Glatthorn et al., 2019). This study analyzed data for change throughout the implementation of the professional learning imposed by the state.

The literature review contributed to the study's methodology and addressed the following research questions:

1. How have assessment scores changed since the beginning of LETRS training implementation?
2. How have growth index scores changed since the beginning of LETRS training implementation?

Chapter 3 explains the quantitative methodology chosen to address these research questions. ANOVA tests were completed to answer Research Question 1. A dependent *t* test was conducted to answer Research Question 2. The data collection process and procedures are included. In addition, the chapter describes the measures, variables, and instruments used in the study.

Chapter 3: Methodology

This study investigated the impact of the LETRS professional development mandated by the state of North Carolina. This professional development is a rigorous 2-year process meant to improve teacher early instructional reading practices. North Carolina measures the early reading progress of students with the DIBELS 8 and EOG assessments. In theory, if professional development leads to an increase in teacher pedagogical knowledge and improved instructional practices, students will strengthen their foundational literacy skills (Scarparolo & Hammond, 2018).

I chose a quantitative design for this study. Quantitative research examines the relationship among variables and analyzes data using statistical procedures (Creswell & Creswell, 2018). This study compared EOG assessment data across 3 years to determine the impact of the LETRS professional development. The 2020-2021 school year data are considered Year 0, 2021-2022 is Year 1, and 2022-2023 data are Year 2 of the professional development. The study answered the following research questions:

1. How have assessment scores changed since the beginning of LETRS training implementation?
2. How have growth index scores changed since the beginning of LETRS training implementation?

Population

I analyzed third- through fifth-grade EOG data from schools in 39 districts that comprised Cohort 1. This cohort began implementation of the LETRS professional development in 2021 and finished in 2023. For Research Question 1, the sample population numbered 6,502 for third grade. There were 4,667 data for fourth grade and

4,626 for fifth grade. Where demographic data are disaggregated by subgroup, the N represents the number of schools that reported for that subgroup. For Question 2, the population numbered 942.

Measures

The LETRS professional development was expected to improve instructional practices and increase reading proficiency and growth. I analyzed any change in reading proficiency, as indicated by EOG assessment scores, during and after the professional development. Reading proficiency can be defined as students who have obtained comprehension skills as well as the preceding literacy skills that lead to an adequate understanding of the text.

Reading Comprehension

Comprehension refers to the ability to extract and construct meaning from text and is the primary purpose of reading (Hennessy, 2021; Moats, 2020a). Skills in both strands of Scarborough's (2001) Reading Rope, word recognition and language comprehension, are necessary for understanding the text. Reading comprehension requires the activation of complicated mental processes and skills acquisition (Castles et al., 2018). A reader's background knowledge and reasoning skills contribute to their ability to comprehend text (Moats, 2020a).

While the subskills that contribute to adequate comprehension are not directly assessed on the EOG assessment, they are necessary for proficiency. Knowledge of letters, phonemic awareness, the alphabetic principle, vocabulary, and fluency must be mastered to be able to comprehend text (Hoover & Tunmer, 2020; Kilpatrick, 2015, 2016; Moats, 2020a). For this reason, the subskills are incidentally measured in the EOG

assessments.

Instrumentation

The instruments used in this study included the EOG assessments for third, fourth, and fifth grades. The EOGs were given in the following school years: 2020-2021, 2021-2022, and 2022-2023. Growth index scores generated from the reading EOG assessment data were also used for the same years.

EOG Assessment Scores

North Carolina requires that all third- through fifth-grade students take an EOG assessment to measure comprehension and growth. EOG assessments in Grades 3-5 contain five reading selections. Each reading selection contains eight assessment questions with a four-response option multiple-choice format. In addition, each assessment contains one field test passage and eight additional field test items that are not calculated in the final score (NCDPI, 2023).

NCDPI Accountability Services (2021) noted that the test development process was rigorous, with multiple stages containing checks and balances. Items for EOG assessments undergo a process that has 19 stages, from creation to final approval. Trained educators from a variety of backgrounds are involved throughout the entire process.

Proficiency on the EOGs is reported in five levels. Levels 1 and 2 are considered not proficient. Students achieving Levels 3-5 are proficient in reading comprehension. Level 3 scores are labeled as grade-level proficient. Scores falling in Levels 4 and 5 are further categorized as being college and career ready. Levels are determined by cut scores at each grade level. Table 8 indicates the required scores for students in third through fifth grade to meet each level.

Table 8

Reading Cut Scores

Grade	Not proficient (Levels 1 & 2)	Level 3	Level 4	Level 5
3	≤ 539	540-545	546-550	≥ 551
4	≤ 543	544-547	548-555	≥ 556
5	≤ 549	550-553	554-559	≥ 560

I analyzed scores that fell between Levels 3 and 5. These levels were considered grade-level proficient and provided insight into the students who grew developmentally as a result of the professional development.

School Growth Index Scores

A school's growth score is calculated by the Education Value-Added Assessment System (EVAAS), which measures student growth. Growth is defined as the amount of academic progress students make in a subject during a school year. EVAAS uses a value-added growth model to calculate student growth using EOG assessment data. Current assessment achievement is compared to all previous achievements on quality assessments. A composite index of growth is generated. The growth index scores are divided into three categories. Schools with growth index scores less than -2 are considered to have not met growth. Scores anywhere from -2 to 1.99 are proficient for growth. Finally, scores that are greater or equal to 2 have exceeded growth expectations.

Data Collection Process

EOG data are made available to the public each year by NCDPI. The data were taken from NCDPI data and accountability website and were contained in a public Excel file titled *School Assessment and Other Indicator Data* for Years 2021-2023. The data came from the 39 districts represented in Cohort 1 and were uploaded to JASP software

for analysis. They were sorted by the EOG test scores for Grades 3-5 and disaggregated by demographics and year of implementation.

Data Analysis

I analyzed the data through quantitative methodology. I ran three different two-way analysis of variance (ANOVA) models for each grade level, 3-5, to answer Research Question 1. I investigated mean differences in EOG scores and growth scores among the different levels of implementation as well as the interactions between demographics. Post hoc comparison tests using the Tukey method were run when the ANOVAs resulted in statistical differences. Independent variables are manipulated in experimental studies and affect the outcomes. Dependent variables are the outcomes of the study and are influenced by the independent variables in the study (Creswell & Creswell, 2018). Table 9 shows the alignment of the research questions, the dependent and independent variables, and the instrument used in the study. For Research Question 2, I ran a *t* test to compare the school growth index scores after 1 year of implementation to the scores after the final year of implementation.

Table 9

Alignment of Research Questions and Variables

Research question	Method	Dependent variable	Independent variable	Instrument
1	Two-way ANOVA	Overall reading proficiency	Time, demographics, grade levels	EOG
2	Dependent *t* test	Reading growth	Time	Growth index score

The variables and instrumentation depicted in Table 9 illustrate how each research question was addressed in the study.

Research Question 1: How Have Assessment Scores Changed Since the Beginning of LETRS Training Implementation?

In order to answer Research Question 1, I ran three two-way ANOVA models. This allowed me to analyze the data longitudinally to investigate whether there were statistically significant changes in student overall reading development during the LETRS implementation. The data came from Grades 3-5 EOGs from the 2021-2023 school years.

Research Question 2: How Have Growth Index Scores Changed Since the Beginning of LETRS Training Implementation?

In order to answer Research Question 2, I ran a dependent *t* test. This allowed me to analyze the data longitudinally to investigate whether there were statistically significant changes in reading growth scores during the LETRS implementation. The data compared the 2021-2022 school year data to the 2022-2023 school year data. The 2021-2022 school year marked the completion of Year 1 of the professional development for Cohort 1. The 2022-2023 school year marked the end of the completion of Year 2 of the professional development.

Summary

I used a quantitative methodology for this study. I analyzed EOG assessment data for Grades 3-5. In addition, the growth index scores for the reading EOG were also utilized in the analysis. ANOVA tests were conducted to answer Research Question 1. The tests included the variables of time and race. For Research Question 2, I used a dependent *t* test.

Chapter 4 presents the results of the quantitative tests. The results of the tests run to answer Research Question 1 are presented first. The results are divided and detailed for each grade level: third, fourth, and fifth respectively. A summary follows each grade-level description of results. Next, the results of the tests run to answer Research Question 2 are presented.

Chapter 4: Results

This study focused on the relationship between professional development and reading proficiency. The intensity and rigor of LETRS professional development had the potential to change instructional practices and improve reading comprehension (Moats & Tolman, 2019). One way I chose to investigate the impact of the professional development on reading comprehension was to analyze reading EOG assessment data. The data were taken from Years 2021-2023 for Grades 3-5. Only data from Cohort 1 were analyzed. This cohort began the LETRS professional development in 2021 and finished in 2023. A series of ANOVA tests were run, and the results are presented in this chapter.

Another way I chose to investigate the impact of the professional development was to analyze growth index scores for reading. Each school reports a growth index score for the EOG reading assessments combined. Research Question 2 investigated any changes in these scores from the completion of the first year of professional development to the completion of the second year of professional development. A dependent *t* test was completed to ascertain any changes.

Research Question 1: How Have Assessment Scores Changed Since the Beginning of LETRS Training Implementation?

A two-way ANOVA was run to investigate changes in EOG assessment scores by year, demographics, and the interactions between year and demographics for each individual grade. The analysis focused on changes to grade-level proficiency. After a statistically significant result, a Tukey method for the post hoc comparisons was conducted to discover which means were significantly different from each other. Prior to

interpreting the results, Levene's test was used to check the normal distribution assumption. The result showed that the assumption had been met, $p = > 0.05$.

Third-Grade Reading Proficiency

A two-way ANOVA test was run for year and demographics for Grade 3 grade-level proficient EOG scores. Table 10 shows the results of the omnibus test.

Table 10

ANOVA Omnibus Results for Third Grade

Cases	Sum of squares	*df*	Mean square	*F*	*p*	η^2
Year	82358.411	2	41179.206	123.226	< .001	0.028
Race	696590.959	5	139318.192	416.900	< .001	0.235
Year ∗ Race	6161.096	10	616.110	1.844	0.048	0.002
Residuals	2.173e+6	6502	334.176			

Note. Type 3 Sum of Squares.

Table 10 indicates that there were statistical differences by year of implementation and demographics. In addition, there was a statistically significant interaction between the year of implementation and demographics.

Years. The test examined any differences by year of implementation. The test was run with year as the independent variable and grade-level proficient EOG as the dependent variable. The sample size for this portion of the study was 6,520 participants divided into three groups based on year of implementation. The two-way ANOVA result indicated that there existed statistically significant differences in terms of the years of implementation, $F(2, 6502) = 123.23$, $p = < 0.001$ with an effect size of $\eta^2 = .03$. The results are shown in Table 10. Table 11 displays the descriptive statistics by year.

Table 11

Descriptives – Grade-Level Proficient EOG Scores by Year

Year	Mean	*SD*	*N*
2021	41.492	22.057	2126
2022	48.927	20.867	2172
2023	51.818	20.070	2222

The results indicated that Year 2023 had the highest average of $M = 51.82$, SD = 20.07. Year 2022 followed with an average of $M = 48.93$, $SD = 20.87$. Year 2021 had the lowest average of $M = 41.49$, $SD = 22.06$.

Furthermore, to investigate which pairs had statistically significant mean differences, a Tukey post hoc test was conducted. The post hoc analysis revealed that there were statistical differences between all years. Table 12 indicates the post hoc comparisons between years.

Table 12

Post Hoc Comparisons by Years

Variable		Mean difference	SE	t	p
2021	2022	-7.435	0.641	-11.605	< .001
	2023	-10.326	0.637	-16.209	< .001
2022	2023	-2.891	0.634	-4.563	< .001

Table 12 indicates that third-grade students scored an average of 7.44 points higher in Year 2022 than in Year 2021. In addition, third-grade students scored an average of 10.33 points higher in Year 2023 compared to Year 2021. There was also a statistical difference between Year 2022 and Year 2023. Students scored an average of 2.89 points higher in 2023 compared to 2022.

Demographics. The ANOVA test examined any differences by demographics. The demographic categories were divided into the following races: Asian, Black, Hispanic, Multi-Race, and White. The test was run with demographics as the independent variable, along with the grade-level proficient EOG as the dependent variable. The results are shown in Table 10. The sample size for this portion of the study was 4,920 participants divided into five groups based on demographics. The two-way ANOVA result indicated that there existed statistically significant differences in terms of demographics, F (5, 6502) = 416.90, p = < 0.001, with an effect size of η^2 =.24. Table 13 indicates the means by demographics.

Table 13

Means by Race

Race	Mean	*SD*	*N*
Asian	67.782	21.424	368
Black	34.657	17.451	1314
Hispanic	40.050	18.747	1297
Multiracial	50.266	21.222	634
White	61.421	17.463	1307

The results in Table 13 indicated that the Asian subgroup had the highest average of M = 67.78, *SD* = 21.42. The White subgroup had the next-to-highest average of *M* = 61.42, *SD* = 17.46. Following the White group was the Multiracial subgroup at *M* = 50.27, *SD* = 21.22. The next-to-lowest average was the Hispanic group at *M*= 40.05, *SD* = 18.75. The Black demographic group had the lowest average of *M* = 34.66, *SD* = 17.45. Except for the Asian subgroup, the means of all the other minority races were significantly behind the White demographic group.

Furthermore, to investigate which pairs had statistically significant mean differences, a Tukey post hoc test was conducted. The post hoc analysis revealed that there were statistical differences among all demographics. Table 14 indicates the post hoc comparisons between each of the five demographic groups.

Table 14

Post Hoc Comparisons by Demographics

Race		Mean difference	*SE*	*t*	*p*
Asian	Black	33.125	1.108	29.906	< .001
	Hispanic	27.732	1.109	25.001	< .001
	Multiracial	17.516	1.231	14.232	< .001
	White	6.361	1.108	5.739	< .001
Black	Hispanic	-5.393	0.735	-7.336	< .001
	Multiracial	-15.609	0.908	-17.188	< .001
	White	-26.764	0.734	-36.479	< .001
Hispanic	Multiracial	-10.217	0.910	-11.226	< .001
	White	-21.371	0.736	-29.034	< .001
Multiracial	White	-11.155	0.909	-12.272	< .001

Table 14 indicates the mean statistical differences between all student demographics on the post hoc comparison test. The largest difference occurred between the Asian and Black demographic groups. The Asian demographic group scored 33.13 average points higher than the lowest average of the Black demographic group. In contrast, the smallest difference can be seen between the Black and Hispanic groups, with the Hispanic group averaging 5.39 points higher.

Demographics and Year. The ANOVA test examined any differences by demographics and years. The test was run with demographics and years as the categorical variables along with the grade-level proficient EOG, and the results are shown in Table 10. The sample size for this portion of the study was 4,920 participants divided into three

groups per each of the five demographic groups for a total division of 15 equal groups. The two-way ANOVA result indicated that statistically significant differences existed in terms of demographics and year, F (10, 6502) = 1.84, p = 0.048, with an effect size of η^2 = 0.002. Table 15 indicates the means by demographics and years.

Table 15

EOG Means by Demographics and Years

Race	Year	Mean	*SD*	*N*
Asian	2021	63.075	22.775	119
	2022	70.329	20.738	120
	2023	69.756	20.185	129
Black	2021	26.931	16.198	434
	2022	36.272	16.712	442
	2023	40.683	16.580	438
Hispanic	2021	33.752	19.095	428
	2022	41.754	17.561	428
	2023	44.509	17.908	441
Multiracial	2021	45.087	20.832	193
	2022	50.886	22.012	210
	2023	54.030	19.997	231
White	2021	57.153	17.430	425
	2022	62.609	16.877	439
	2023	64.339	17.318	443

In Table 15, each demographic group's average is listed by each of the 3 years. Table 15 indicates that the Asian demographic group had the highest average in Year 2022: M = 70.33, SD = 20.74. The Black demographic had the highest average in Year 2023: M = 40.68, SD = 16.58. Year 2023 ranked highest among each of the other demographic groups as well: Hispanic (M = 44.51, SD = 17.91), Multi (M =54.03, SD = 20.00), White (M = 64.34, SD = 17.32). The means can be seen in rank order in Table 16.

Table 16

Descending Order EOG Means by Demographics and Years for Third Grade

Race	Year	Mean	*SD*	*N*
Asian	2022	70.329	20.738	120
Asian	2023	69.756	20.185	129
White	2023	64.339	17.318	443
Asian	2021	63.075	22.775	119
White	2022	62.609	16.877	439
White	2021	57.153	17.430	425
Multiracial	2023	54.030	19.997	231
Multiracial	2022	50.886	22.012	210
Multiracial	2021	45.087	20.832	193
Hispanic	2023	44.509	17.908	441
Hispanic	2022	41.754	17.561	428
Black	2023	40.683	16.580	438
Black	2022	36.272	16.712	442
Hispanic	2021	33.752	19.095	428
Black	2021	26.931	16.198	434

Table 16 shows the variation in groups and years when ranking the means. Table 16 depicts the Asian demographic group in Year 2022 with the highest average when ranked: $M = 70.33$, $SD = 20.74$. The Black demographic in Year 2021 resulted in the lowest average of $M = 26.93$, $SD = 16.20$.

Furthermore, to investigate which pair(s) had statistically significant mean differences, a post hoc test was conducted. The post hoc analysis revealed that there were statistical differences among some of the demographics and years. Table 17 indicates the post hoc comparisons between the demographic groups and years.

Table 17

Post Hoc Comparisons for Third-Grade Demographics

Variables		Mean difference	*SE*	*t*	*p*
Asian 2021	Asian 2022	-7.254	2.365	-3.067	0.169
	Asian 2023	-6.681	2.324	-2.875	0.267
Black 2021	Black 2022	-9.341	1.235	-7.561	< .001
	Black 2023	-13.751	1.238	-11.107	< .001
Hispanic 2021	Hispanic 2022	-8.002	1.250	-6.404	< .001
	Hispanic 2023	-10.757	1.240	-8.672	< .001
Multiracial 2021	Multiracial 2022	-5.799	1.823	-3.181	0.125
	Multiracial 2023	-8.943	1.783	-5.017	< .001
White 2021	White 2022	-5.455	1.244	-4.385	0.002
	White 2023	-7.185	1.241	-5.789	< .001

The post hoc comparisons in Table 17 depict the statistical differences for each demographic group at the end of Year 2022, which would have ended Year 1 of the professional development, and the statistical differences after Year 2023, which would have concluded Year 2 of the professional development. Table 17 indicates that there were statistical differences for some groups across the years but not all.

For the Asian demographic group, there were no statistical differences between the years. The Black and Hispanic student demographic groups showed statistical significance for both years. The Multiracial demographic group showed a statistical difference from Year 1 of the professional development (Year 2021) to the end of the professional development (Year 2023); however, there was no statistical difference when looking at this group from Year 2021 to Year 2022. The White demographic group showed statistically significant differences for both years.

Third-Grade Achievement Gap. The post hoc comparison showed that there were achievement gaps in the grade-level proficient EOG scores across the years. Table 18 shows the differences in EOG scores for minority demographic groups and the White demographic group for each year. In 2021, the Black demographic student group scored 30.22 average points behind the White demographic group. By the end of the professional development in 2023, the Black demographic scored 23.66 average points behind the White group. The Hispanic group's average mean was 23.40 behind the White group in 2021 and 19.83 points behind the White group in 2023. In 2021, the Multi-Race group scored 12.07 points behind the White group and 10.31 points behind the White demographic group in 2023.

Table 18

Post Hoc Comparisons Pairwise for Race and Year

Variables		Mean difference	*SE*	*t*	*p*
Black 2021	White 2021	-30.222	1.248	-24.226	< .001
Black 2022	White 2022	-26.337	1.232	-21.381	< .001
Black 2023	White 2023	-23.656	1.232	-19.205	< .001
Hispanic 2021	White 2021	-23.402	1.252	-18.694	< .001
Hispanic 2022	White 2022	-20.855	1.242	-16.794	< .001
Hispanic 2023	White 2023	-19.830	1.230	-16.126	< .001
Multiracial 2021	White 2021	-12.066	1.587	-7.604	< .001
Multiracial 2022	White 2022	-11.723	1.534	-7.643	< .001
Multiracial 2023	White 2023	-10.309	1.484	-6.948	< .001
Asian 2021	White 2021	5.921	1.896	3.123	0.147
Asian 2022	White 2022	7.720	1.883	4.100	0.005
Asian 2023	White 2023	5.417	1.829	2.962	0.219

In contrast to these groups, the Asian demographic group had an average mean of 5.92 above the White demographic group in 2021. In 2023, the Asian demographic

student group averaged 5.42 points above the White demographic student group; however, there was no statistical difference between the Asian group and the White group for Year 2021 or Year 2023. On the other hand, there was a statistical difference between the Asian and White demographic groups for the 2022 year. During this year, the Asian group scored 7.72 points higher than the White group, and it was statistically significant, $p = 0.005$. Year 2022 was the end of Year 1 of the professional development. Table 18 shows the comparisons between the groups for all years.

Table 18 indicates that the largest achievement gap exists between the Black demographic group and the White demographic group. The gap between the Hispanic and the Multiracial demographic groups compared to the White demographic group follows. The achievement gap between each minority demographic group and the White demographic group declined each year. Conversely, the Asian demographic group performed better than the White demographic group, with a significant value for Year 2022.

Third-Grade Summary. ANOVA tests showed that there were statistically significant differences in grade-level proficiency EOG scores for third-grade students for the categorical variables of year and demographics. The size indicates that demographics had the largest impact on the scores. The statistical differences were apparent across most years and demographic groups. Figure 8 illustrates the EOG means for each demographic group from year to year.

Figure 8

Third-Grade EOG Change in Means

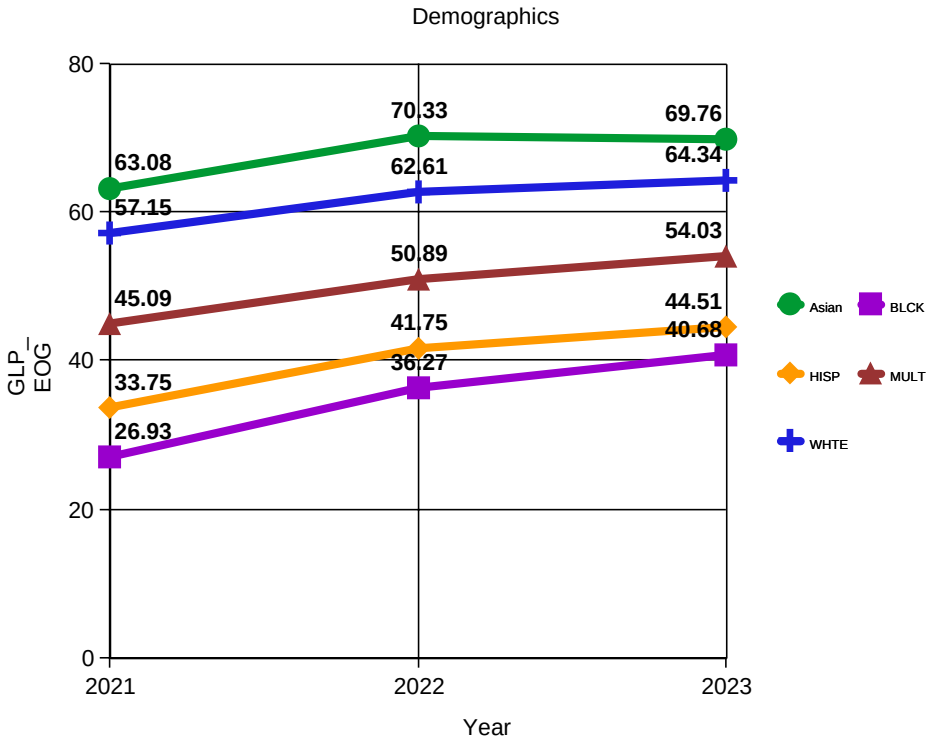


Figure 8 indicates that the Black, White, and Hispanic demographic groups improved each year. These results were statistically significant from year to year as well as from the start of the professional development in Year 2021 to the end of the professional development in Year 2023. The Multiracial demographic group also presented an increase in EOG scores from year to year; however, there was only statistical significance from Year 2021 to Year 2023 when analyzing data for the Multiracial demographic group. Therefore, the professional development increased scores after Year 1 and after completion for the Black, White, and Hispanic demographic groups. For the Multiracial group, the impact of the professional development was only evident after completion. For the Asian demographic group, any changes were not statistically significant, indicating the professional development did not impact this group

after Year 1 or Year 2.

When the post hoc comparison test was run, gaps in achievement were apparent and significant for all minority groups except the Asian demographic group. The achievement gap for each of the other groups declined each year. The comparisons between the Black to White demographic groups were significant for all years. The Hispanic and Multiracial comparisons to the White demographic were also all significant. Only Year 2022 was significant when comparing the Asian and White demographic groups.

Fourth-Grade Reading Proficiency

A two-way ANOVA test was run for year and demographics for Grade 4 grade-level proficient EOG scores. Table 19 shows the results of the omnibus test.

Table 19

ANOVA Omnibus Results for Fourth Grade

Cases	Sum of squares	*df*	Mean square	*F*	*p*	η^2
Year	29224.734	2	14612.367	51.731	< .001	0.015
Race	568134.605	5	113626.921	402.266	< .001	0.299
Year ∗ Race	3294.089	10	329.409	1.166	0.309	0.002
Residuals	1.302e+6	4609	282.467			

Note. Type 3 Sum of Squares.

Table 19 indicates that there were statistical differences by year of implementation and demographics. There were no statistically significant interactions between the year of implementation and demographics.

Years. The test examined any differences by year of implementation. The test was run with year as the independent variable and grade-level proficient EOG scores as

the dependent variable. The sample size for this portion of the study was 4,627 participants divided into three groups based on year of implementation. The two-way ANOVA result indicated that there existed statistically significant differences in terms of the years of implementation, F (2, 4609) = 51.73, p = < 0.001, with an effect size of η^2 =.02. The results are shown in Table 19. Table 20 displays the descriptive statistics by year.

Table 20

Descriptives – Grade-Level Proficient EOG Scores by Year for Fourth Grade

Year	Mean	*SD*	*N*
2021	41.336	20.501	1491
2022	47.000	20.132	1557
2023	51.487	19.772	1579

The results indicated that Year 2023 had the highest average of M = 51.49, SD = 19.77. Year 2022 followed with an average of M = 47.00, SD = 20.13. Year 2021 had the lowest average of M = 41.34, SD = 20.50.

Furthermore, to investigate which pairs had statistically significant mean differences, a Tukey post hoc test was conducted. The post hoc analysis revealed that there were statistical differences between all years. Table 21 indicates the post hoc comparisons between years.

Table 21

Post Hoc Comparisons by Years for Fourth Grade

Variables		Mean difference	*SE*	*t*	*p*
2021	2022	-6.920	1.072	-6.458	< .001
	2023	-10.705	1.057	-10.129	< .001
2022	2023	-3.785	0.997	-3.798	< .001

Table 21 indicates that fourth-grade students scored an average of 6.92 points higher in Year 2022 than in Year 2021. In addition, fourth-grade students scored an average of 10.70 points higher in Year 2023 compared to Year 2021. There was also a statistical difference between Year 2022 and Year 2023. Students scored an average of 3.79 points higher in 2023 compared to 2022.

Demographics. The ANOVA test examined any differences by demographics. The demographic categories were the following ethnicities: Asian, Black, Hispanic, Multi-Race, and White. The test was run with demographics as the independent variable and the grade-level proficient EOG as the dependent variable. The results are shown in Table 19. The sample size for this portion of the study was 4,920 participants divided into five groups based on demographics. The two-way ANOVA result indicated that there existed statistically significant differences in terms of demographics, F (5, 4609) = 402.27, p = < 0.001, with an effect size of η^2 =.30. Table 22 indicates the means by demographics.

Table 22

Fourth-Grade Means by Demographics

Race	Mean	*SD*	*N*
Asian	72.148	20.267	107
Black	34.249	16.200	995
Hispanic	37.771	17.511	900
Multiracial	54.419	21.150	106
White	62.598	16.039	1034

The results in Table 22 indicate that the Asian subgroup had the highest average of M = 72.15 SD = 20.27. The White subgroup had the next-to-highest average of M = 62.60, SD = 16.04. Following the White group was the Multiracial subgroup at M =

54.42, *SD* = 21.15. The next-to-lowest average was the Hispanic group at *M*= 37.77, S*D* = 17.51. The Black demographic group had the lowest average of *M* = 34.25, *SD* = 16.20.

Furthermore, to investigate which pairs had statistically significant mean differences, a Tukey post hoc test was conducted. The post hoc analysis revealed that there were statistical differences among all demographics. Table 23 indicates the post hoc comparisons between each of the five demographic groups.

Table 23

Post Hoc Comparisons by Race for Fourth Grade

Race		Mean difference	*SE*	*t*	*p*
Asian	Black	37.988	1.711	22.203	< .001
	Hispanic	34.503	1.720	20.062	< .001
	Multiracial	19.180	2.335	8.213	< .001
	White	9.442	1.708	5.529	< .001
Black	Hispanic	-3.484	0.774	-4.503	< .001
	Multiracial	-18.808	1.759	-10.690	< .001
	White	-28.546	0.747	-38.235	< .001
Hispanic	Multiracial	-15.323	1.768	-8.667	< .001
	White	-25.061	0.767	-32.688	< .001
Multiracial	White	-9.738	1.756	-5.545	< .001

Table 23 indicates the mean statistically significant differences between all student demographics on the post hoc comparison test. The largest difference occurred between the Asian and Black demographic groups. The Asian demographic group scored 37.99 average points higher than the lowest average of the Black demographic group. In contrast, the smallest difference can be seen between the Black and Hispanic groups, with the Hispanic group averaging 3.48 points higher.

Demographics and Year. The test examined any differences by demographics and years. The test was run with demographics and years as the categorical variables and

with the grade-level proficient EOG as the dependent variable. The results are shown in Table 19. The sample size for this portion of the study was 3,142 participants divided into three groups per each of the five demographic groups for a total division of 15 groups. The two-way ANOVA result indicated that no statistically significant differences existed in terms of demographics and year, $F(10, 4609) = 1.16$, p 0.31. Table 24 indicates the means by demographics and years.

Table 24

Fourth-Grade EOG Means by Demographics

Race	Year	Mean	*SD*	*N*
Asian	2021	68.626	22.711	34
	2022	73.611	19.667	36
	2023	73.959	18.531	37
Black	2021	27.473	13.901	317
	2022	34.296	15.520	335
	2023	40.465	16.362	343
Hispanic	2021	32.179	16.369	281
	2022	37.582	16.710	303
	2023	42.926	17.740	316
Multiracial	2021	42.900	22.951	25
	2022	56.324	18.795	38
	2023	59.433	19.954	43
White	2021	58.043	16.275	348
	2022	63.194	15.148	344
	2023	66.634	15.528	342

In Table 24, each demographic group's average is listed by each of the 3 years. All demographic subgroups had an increase in average scores from year to year. Each group had the highest mean average in 2023: Asian ($M = 73.92$, $SD = 18.53$), Black ($M = 40.47$, $SD = 16.36$), Hispanic ($M = 42.93$, $SD = 17.74$), Multiracial ($M = 59.43$, $SD = 19.95$), White ($M = 66.63$, $SD = 15.52$). The means can be seen in rank order in Table 25.

Table 25

Descending Order EOG Means by Race and Years for Fourth Grade

Race	Year	Mean	*SD*	*N*
Asian	2023	73.959	18.531	37
Asian	2022	73.611	19.667	36
Asian	2021	68.626	22.711	34
White	2023	66.634	15.528	342
White	2022	63.194	15.148	344
Multiracial	2023	59.433	19.954	43
White	2021	58.043	16.275	348
Multiracial	2022	56.324	18.795	38
Hispanic	2023	42.926	17.740	316
Multiracial	2021	42.900	22.951	25
Black	2023	40.465	16.362	343
Hispanic	2022	37.582	16.710	303
Black	2022	34.296	15.520	335
Hispanic	2021	32.179	16.369	281
Black	2021	27.473	13.901	317

Table 25 shows the variation in groups and years when ranking the means. Table 25 depicts the Asian demographic group in Year 2023 with the highest average when ranked (M = 73.96, SD = 18.53). The Black demographic in Year 2021 resulted in the lowest average (M = 27.47, SD = 13.90).

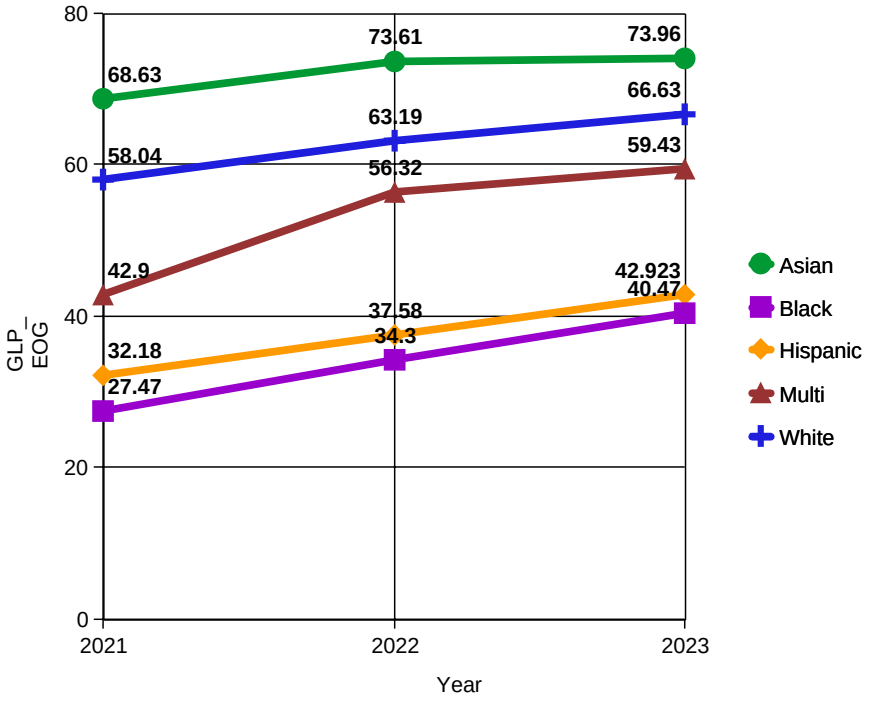
Fourth-Grade Summary. There were statistically significant differences for the entire fourth-grade population across all years. Fourth-grade students had the largest gain from the start of the professional development in 2021 to the end of the professional development in 2023. There was also an increase from Year 2021 to Year 2022. In addition, there was a statistical difference from the end of Year 1 of the professional development in 2022 to the final year in 2023. This result would suggest a positive impact overall; however, the effect size indicates a small impact.

When comparing the population by race, there were significant differences as well. The White demographic group performed higher than the Black, Hispanic, and Multiracial demographic groups. These results indicate that an achievement gap exists when looking at the means of the groups for all of the years. The effect size for demographics was greater than that of years.

Figure 9 shows the means for each race at the end of each year. The scores show an increase for demographic groups from year to year.

Figure 9

Fourth-Grade EOG Change in Means



Although the means increased across the years, a post hoc analysis showed no statistical difference when the year was combined with race.

Reading Proficiency in Fifth Grade

A two-way ANOVA test was run for year and demographics for Grade 5 grade-

level proficient EOG scores. Table 19 shows the results of the omnibus test.

Table 26

ANOVA Omnibus Results for Fifth Grade

Cases	Sum of squares	*df*	Mean square	*F*	*p*	η^2
Year	6886.239	2	3443.120	11.710	< .001	0.003
Race	671102.823	5	134220.565	456.498	< .001	0.329
Year ∗ Race	2449.326	10	244.933	0.833	0.597	0.001
Residuals	1.360e+6	4626	294.023			

Note. Type 3 Sum of Squares.

Table 26 indicates that there were statistical differences by year of implementation and demographics. There were no statistically significant interactions when the year of implementation and demographics were combined.

Years. The test examined any differences by year of implementation. The test was run with year as the independent variable and grade-level proficient EOG as the dependent variable. The sample size for this portion of the study was 4,644 participants divided into three groups based on year of implementation. The two-way ANOVA result indicated that there existed statistically significant differences in terms of the years of implementation, F (2, 4626) = 11.71, p = < 0.001. The results are shown in Table 26. Table 27 displays the descriptive statistics by year.

Table 27

Descriptives – Grade-Level Proficient EOG Scores by Year for Fifth Grade

Year	Mean	*SD*	*N*
2021	38.862	20.599	1527
2022	41.903	21.202	1539
2023	44.989	20.981	1578

The results indicated that Year 2023 had the highest average of *M* = 44.99, *SD* = 20.98. Year 2022 followed with an average of *M* = 41.90, *SD* = 21.20. Year 2021 had the lowest average of *M* = 38.86, *SD* = 20.60.

Furthermore, to investigate which pairs had statistically significant mean differences, a Tukey post hoc test was conducted. The post hoc analysis revealed that there were statistical differences between Year 2021 and Year 2023 and between Year 2022 and 2023; however, there was no statistical difference between Year 2021 and Year 2022. Table 28 indicates the post hoc comparisons between years.

Table 28

Post Hoc Comparisons by Years for Fifth Grade

Variable		Mean difference	*SE*	*t*	*p*
2021	2022	-1.529	1.052	-1.454	0.313
	2023	-4.709	1.003	-4.693	< .001
2022	2023	-3.180	1.003	-3.169	0.004

Table 28 indicates that fifth-grade students scored an average of 4.8 points higher in Year 2023 compared to Year 2021. There was also a statistical difference between Year 2022 and Year 2023. Students scored an average of 3.18 points higher in 2023 compared to 2022. There was no statistical difference noted between Year 2021 and Year 2022.

Demographics. The test examined any differences by demographics. The demographic categories were sorted by the following ethnicities: Asian, Black, Hispanic, Multi-Race, and White. The test was run with demographics as the independent variable, along with the grade-level proficient EOG as the dependent variable. The results are shown in Table 26. The sample size for this portion of the study was 4,644 participants

divided into five groups based on demographics. The two-way ANOVA result indicated that the demographic data were statistically significant: F (5,4626) = 456.50, p = < 0.001. These data had an effect size of η^2 =.33. Table 29 indicates the means by demographics.

Table 29

Fifth Grade Means by Race

Race	Mean	*SD*	*N*
Asia	74.002	21.300	117
Black	28.640	15.517	1011
Hispanic	32.606	16.306	910
Multiracial	46.792	21.419	115
White	58.751	17.059	1026

The results in Table 29 indicate that the Asian subgroup had the highest average of M = 74.00 SD = 21.30. The White subgroup had the next-to-highest average of M = 58.75, SD = 17.06. Following the White group was the Multiracial subgroup at M = 46.79, SD = 21.42. The next-to-lowest average was the Hispanic group at M= 28.64, SD = 15.52. The Black demographic group had the lowest average of M = 28.64, SD = 15.52.

Furthermore, to investigate which pairs had statistically significant mean differences, a Tukey post hoc test was conducted. The post hoc analysis revealed that there were statistical differences among all demographics. Table 30 indicates the post hoc comparisons between each of the five demographic groups.

Table 30

Post Hoc Comparisons by Race for Fifth Grade

Race		Mean difference	*SE*	*t*	*p*
Asian	Black	45.435	1.680	27.038	< .001
	Hispanic	41.494	1.690	24.552	< .001
	Multiracial	27.671	2.267	12.208	< .001
	White	15.312	1.679	9.119	< .001
Black	Hispanic	-3.941	0.784	-5.029	< .001
	Multiracial	-17.764	1.702	-10.440	< .001
	White	-30.123	0.760	-39.642	< .001
Hispanic	Multiracial	-13.823	1.711	-8.079	< .001
	White	-26.182	0.781	-33.527	< .001
Multiracial	White	-12.359	1.700	-7.269	< .001

Table 30 indicates the mean statistically significant differences between all student demographics on the post hoc comparison test. The largest difference occurred between the Asian and Black demographic groups. The Asian demographic group scored 45.44 average points higher than the lowest average of the Black demographic group. In contrast, the smallest difference can be seen between the Black and Hispanic groups, with the Hispanic group averaging 3.94 points higher.

Demographics and Year. The test examined any differences by demographics and years. The test was run with demographics and years as the categorical variables and with the grade-level proficient EOG as the dependent variable. The results are shown in Table 26. The sample size for this portion of the study was 3,179 participants divided into three groups per each of the five demographic groups for a total division of 15 groups. The two-way ANOVA result indicated that no statistically significant differences existed in terms of demographics and year, $F(10, 4626) = .83$, $p = 0.597$. Table 31 indicates the means by demographics and years.

Table 31

Fifth-Grade EOG Means by Demographics

Year	Race	Mean	*SD*	*N*
2021	Asian	75.219	17.556	36
	Black	25.829	15.017	334
	Hispanic	29.855	15.805	300
	Multiracial	45.277	20.179	35
	White	54.963	16.855	340
2022	Asian	73.630	23.781	37
	Black	27.954	15.197	337
	Hispanic	32.233	16.302	297
	Multiracial	43.497	21.038	34
	White	59.506	16.497	345
2023	Asian	73.318	22.309	44
	Black	32.081	15.711	340
	Hispanic	35.598	16.332	313
	Multiracial	50.380	22.514	46
	White	61.763	17.165	341

In Table 31, each demographic group's average is listed for each of the 3 years. The highest mean average for the Asian group was Year 2021 *(M* = 75.22, *SD* = 17.56). Year 2023 resulted in the highest average for the remaining demographic groups: Black (*M* = 32.08, *SD* = 15.71), Hispanic (*M* = 35.60, *SD* = 16.33), Multiracial (*M* = 50.38, *SD* = 22.51), White (*M* = 61.76, *SD* = 17.17). The means can be seen in rank order in Table 32. Table 31 indicates which years resulted in the highest and lowest means for each demographic subgroup.

Table 32

Descending Order EOG Means by Race and Years for Fifth Grade

Race	Year	Mean	*SD*	*N*
Asian	2021	75.219	17.556	36
Asian	2022	73.630	23.781	37
Asian	2023	73.318	22.309	44
White	2023	61.763	17.165	341
White	2022	59.506	16.497	345
White	2021	54.963	16.855	340
Multiracial	2023	50.380	22.514	46
Multiracial	2021	45.277	20.179	35
Multiracial	2022	43.497	21.038	34
Hispanic	2023	35.598	16.332	313
Hispanic	2022	32.233	16.302	297
Black	2023	32.081	15.711	340
Hispanic	2021	29.855	15.805	300
Black	2022	27.954	15.197	337
Black	2021	25.829	15.017	334

Table 32 shows the variation in groups and years when ranking the means. Table 32 depicts the Asian demographic group in Year 2021 with the highest average when ranked: M = 75.22, SD = 17.56. The Black demographic group in Year 2021 resulted in the lowest average of M = 25.83, SD = 15.02.

Fifth-Grade Summary. There were statistically significant differences for the entire fifth-grade population from Year 2021 to Year 2023. This comparison showed the largest gain of 4.7 average points. There was also an increase from Year 2022 to Year 2023, which would suggest a positive impact overall.

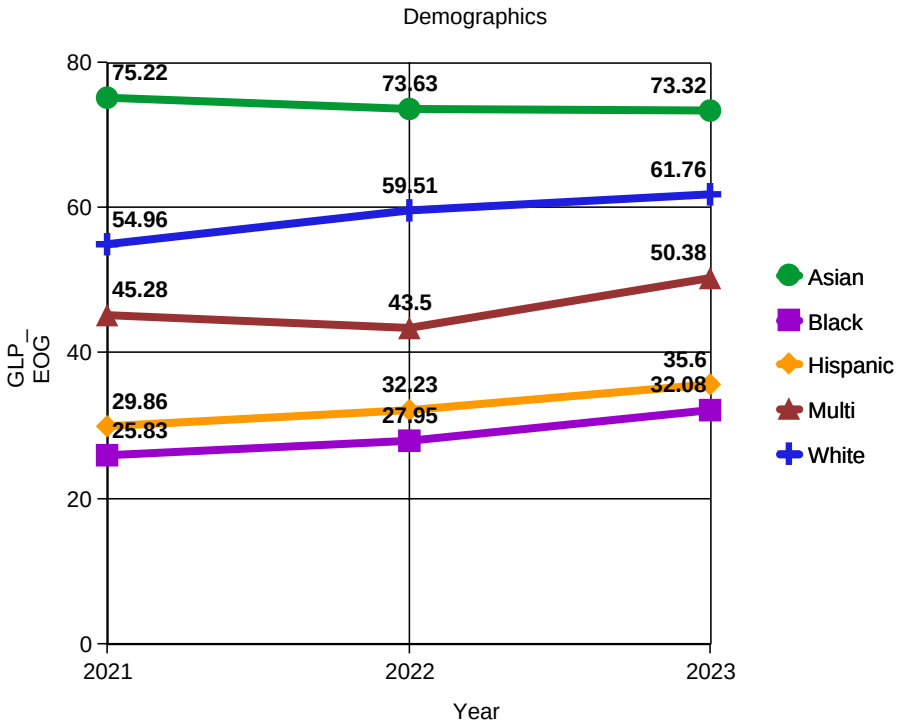
When comparing the population by race, there were significant differences as well. The White demographic group performed higher than the Black, Hispanic, and Multiracial demographic groups. These results indicate that an achievement gap exists

when looking at the means of the groups for all the years. The effect size for demographics was the largest, indicating this category had a greater impact on scores.

Figure 10 shows the means for each race at the end of each year. The scores show an increase for most demographic groups from year to year.

Figure 10

Fifth-Grade EOG Change in Means



Although the means increased across the years, a post hoc analysis showed no statistical difference when the year was combined with race.

Research Question 2: How Have Growth Index Scores Changed Since the Beginning of LETRS Training Implementation?

The purpose of this question was to investigate whether or not the professional development had an impact on the growth index scores from schools in Cohort 1. The reading growth index scores after Year 1 (2022) of the professional development were

compared to the scores after the completion of Year 2 (2023). To investigate the impact through changes to the growth index scores, an independent samples *t* test was conducted. The sample size for this study was 942 students (469 from Year 1, and 473 from Year 2). Levene's test was used to check the equal variance assumption, and the result showed that the assumption had been met, $F(1) = 0.50$, $p = > 0.05$. Table 33 shows the descriptive statistics for this test.

Table 33

Descriptives for Growth Index t Test

Variable	*N*	Mean	*SD*	*SE*	Coefficient of variation
Year 1	469	0.136	2.084	0.096	15.319
Year 2	473	0.235	2.291	0.088	9.767

The *t* test result in Table 33 indicated that the average growth index score from Year 1 was .14 ($M = .14$, $SD = 2.08$). The growth index score for Year 2 was higher ($M = .24$, $SD = 2.29$); however, the difference in average scores did not appear to be statistically significant. Table 34 shows these statistics.

Table 34

Paired Samples t Test

Measure 1	Measure 2	*t*	*df*	*p*
Year 1	Year 2	-1.415	468	0.158

Table 34 indicates that the average score for Year 2 was slightly higher than the average score for Year 1; however, the results were insignificant from Year 1 to Year 2, $t(942) = -1.42$, $p = > 0.05 = 0.158$. These results would indicate that the professional development had no impact on the schools' growth impact scores.

Summary

ANOVA tests indicated that there were increases in EOG assessment data for reading from the beginning of the professional development in 2021 to the completion of the professional development in Year 2023. These increases were apparent for Grades 3, 4, and 5. A post hoc comparison test showed significant interactions between year and race. The dependent *t* test to investigate the impact of the professional development on school growth scores indicated no changes from Year 2022 to Year 2023.

Chapter 5 discusses the implications of the findings, including the disparities among races represented by the data. The results are interpreted through the lens of the Cognitive Foundations of Reading Framework. The chapter also offers recommendations for future research and recommendations for practice. Finally, the limitations and delimitations of the study are considered.

Chapter 5: Discussion

It can take several years to show the impact of a professional development program; however, it is crucial to evaluate any new initiatives for short-term and long-term impact (Glatthorn et al., 2019). North Carolina mandated the LETRS professional development after EPSA (2021), which required that literacy instruction and assessment be based on the science of reading. In August of 2023, NCDPI indicated in a press release that reading proficiency on assessments was improving. NCDPI attributed the gains to the LETRS training received by teachers. Although NCDPI's focus on assessment gains was limited to Grades K-3, this study aimed to analyze the assessments administered in Grades 3-5.

Moats and Tolman (2019) maintained that LETRS is an intensive study of pedagogical practices based on scientific research. This research study aimed to analyze the impact of the LETRS professional development on reading proficiency for schools in Cohort 1 of the professional development. This cohort began the training in 2021 and finished in 2023. The study addressed the following research questions:

1. How have assessment scores changed since the beginning of LETRS training implementation?
2. How have growth index scores changed since the beginning of LETRS training implementation?

I analyzed EOG assessment data for third through fifth grades through ANOVA tests. Year 2022 scores were reflective of the completion of 1 year of professional development. Year 2023 scores came after the full 2 years of training had been completed for Cohort 1. The results from the ANOVAs indicated that there were statistically

significant differences in the grade-level proficient EOG scores taken at the end of 2021, 2022, and 2023. The dependent *t* test that was conducted to determine the impact of the professional development on school growth index scores showed no statistical change from Year 2022 to Year 2023.

Discussion of Findings

The analysis indicated that all grade levels studied increased their means from Year 2021 to Year 2023. The results for these years were all statistically significant. Table 35 details the means further and indicates the average point gains from year to year for each grade level.

Table 35

Average Point Gains Across Years of Implementation

Year	Third grade	Fourth grade	Fifth grade
2021-2022	+ 7	+ 7	NS
2022-2023	+ 3	+ 4	+ 3
2021-2023	+10	+ 10	+ 5

Note. NS = Non-Significant Result

Table 35 indicates that the greatest gains were for Grades 3 and 4. In addition, these grades gained an average of 7 points after Year 1 of the LETRS professional development. Fifth grade showed no significant gains after Year 1.

The results are consistent with the focus of each year of the professional development. Year 1 of LETRS focused on early literacy skills generally taught in Grades K-3. This year focused heavily on the essentials of early reading development and instruction. Teachers were trained in phonemic awareness and decoding while simultaneously learning research on structured literacy practices (Moats & Tolman,

2019).

The results of ANOVA tests indicated that third and fourth grades had more growth than fifth grade. The most statistically significant change was evident in third grade. Table 36 shows the grade level of the population used in this study during the professional development implementation.

Table 36

Grade Levels and Year of Professional Development Implementation

School year	Implementation	Grade		
2022-2023	Year 2	Third	Fourth	Fifth
2021-2022	Year 1	Second	Third	Fourth
2020-2021	Year 0	First	Second	Third

Table 36 shows that the third graders in 2023 were in second grade at the end of Year 1, and the fourth graders were in third grade at the end of Year 1; therefore, these students had more instruction in the foundational skills that were the focus of Year 1 of the professional development. Kindergarten through third-grade teachers emphasize those early literacy skills in instruction (Didion et al., 2020; Kuhfeld et al., 2023). One area that may have influenced the growth in third grade is phonological awareness. Phonological awareness has a significant impact on reading comprehension in second and third grades (Otaiba et al., 2022). Phonological awareness is also extensively taught in Volume 1 of LETRS.

While the results for fourth grade indicated similar gains to those for third grade, the effect size for third grade was higher. When viewing the effect size for all grades with the variable of year, the third grade had the highest: third (effect size = .03), fourth (effect size = .02), and fifth (effect size = .00). This higher effect size indicates that third grade

showed the most impact of the professional development. The content of Volume 1, along with the focus on instruction in early skills development in Grades K-3, contributed to this impact. Third grade represents the cumulation of the precursor skills of the lower grades (Foorman et al., 2016; Shanahan et al., 2010; Vernon-Feagans et al., 2019).

Fifth grade had no statistically significant gains in EOG means after Year 1 of the professional development; however, there was a statistically significant gain for this grade level after Year 2. Reading development for students in kindergarten through third grade typically centers on developing word recognition skills, while students beyond third grade are typically in the stages that focus on comprehension skill development (Hennessy, 2021; Spear-Swerling, 2018). Instruction in fourth grade and beyond focuses more on the complex skills of comprehension (Didion et al., 2020; Kuhfeld et al., 2023). In contrast to Year 1, Year 2 of the professional development focused on the more sophisticated skills of comprehension. These skills included the comprehension processes and instructional strategies that align with the development of fourth- and fifth-grade students.

Reading Proficiency Disaggregated by Race

Third-grade results showed statistically significant changes in the variables of race and year. In addition, the change when race and year were combined as variables was also statistically significant. Post hoc comparisons in third grade revealed that all subgroups, except for the Asian subgroup, had statistically significant growth from the beginning of the professional development in Year 2021 to the end of the professional development in 2023. The increase in the means suggests that the professional development had a positive influence on the White, Black, Hispanic, and Multiracial

subgroups.

The means for the Asian subgroup were not significant, indicating that the professional development had no impact on this group. Despite an abundance of research supporting explicit instruction on skills founded in the science of reading, a small percentage of students do not respond to this type of instruction (Peng & Goodrich, 2020). LETRS is heavily based on this type of instruction, and more research is needed to determine if this professional development adequately addresses literacy development in all races.

In addition to the statistically significant gains that the subgroups of third grade showed over the years, closer examination of the variables of year and race indicated a slight decline in the gap between the minority and White races throughout the professional development. These results were statistically significant. Other studies examining the impact of professional development on reading assessment have failed to show a decline in the achievement gap (Basma & Savage, 2018, 2023; Didion et al., 2020). The result that this decline was only evident in third grade may be due to the higher effect size in the results.

Although third-grade results showed some progress in decreasing the gap among subgroups, the difference in means of minority races remains large across all grades. When the data were analyzed with the demographical variable of race, there were statistically significant differences for all grades. The gaps were most apparent when comparing the Black and Hispanic groups to the White group. Table 37 shows the EOG means of all demographic groups for each grade. These scores reflect the means without the variable of year.

Table 37

EOG Means by Grade and Race

Race	Third Grade	Fourth Grade	Fifth grade
Asian	67.78	72.15	74.02
White	61.42	62.60	58.80
Multi	50.27	54.42	46.80
Hispanic	40.05	37.80	32.61
Black	34.66	34.25	28.64

Table 37 indicates that the largest disparities in scores occurred in fifth grade. In fact, the gap between the White demographic group and the minority demographic groups rose through the grade levels. In third grade, the Multiracial group performed 11 points behind the White group, the Hispanic group was 21 points behind, and the Black group was 26 points behind. By fifth grade, the Multiracial group was 12 points behind the White group, the Hispanic group averaged 26 points behind, and the Black group averaged 30 points behind. These results indicate that after 2 years of professional development, the achievement gap among these groups remains large. This finding is important because the achievement gap contributes to economic and educational system inequalities (Merga, 2020). Understanding the achievement gaps in literacy is particularly complex (Horowitz & Samuels, 2017). The disaggregated results illustrated that the achievement gap had widened from third grade to fifth grade. Merga (2020) noted a wider disparity among subgroups in literacy assessment data as students advance in grade levels.

Theoretical Framework Applications

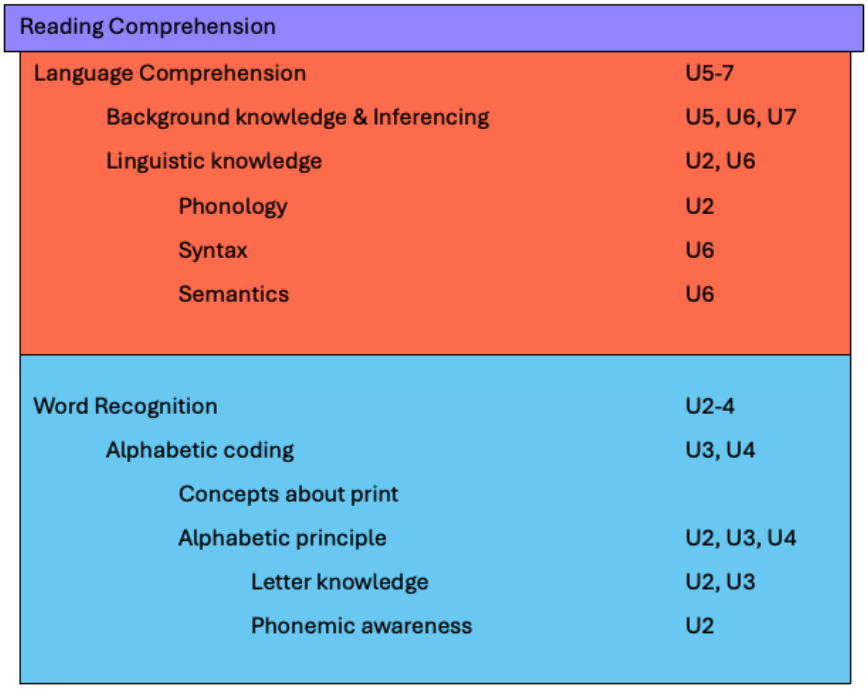
The increase in means from the beginning of the professional development to the end of the professional development indicates a progression in reading development.

Hoover and Tunmer (2020) asserted the benefits of aligning curricula and instruction to the Cognitive Foundations of Reading Framework. The LETRS professional development can be mapped to show the alignment to the cognitive framework as well. The framework is founded on the SVR, indicating that both the word recognition and language comprehension strands play a vital role in reading proficiency. The first year of implementation was from 2021 to 2022 and focused on the word recognition strand. The second year of the implementation, from 2022 to 2023, focused on the language comprehension strand of the Cognitive Foundations of Reading Framework.

Hoover and Tunmer (2020) created visual maps illustrating how the components of essential reading instruction reported by NRP (2000) and Foorman et al. (2016) were aligned with the Cognitive Foundations of Reading Framework. I have already described how LETRS includes components of this seminal research through the description of the units in Chapter 2. To show a more detailed alignment of LETRS to the Cognitive Foundations of Reading Framework, I created a map similar to those of Hoover and Tunmer (2020). Figure 11 illustrates how LETRS maps to the theoretical framework.

Figure 11

Map of LETRS Units in Relation to Theoretical Framework



Note. U stands for units and refers to the units of LETRS.

Figure 11 shows that the content of the LETRS professional development can be directly linked to components of the Cognitive Foundations of Reading Framework. During Year 1 of the professional development, teachers completed Units 1-4; and they completed Units 5-6 during Year 2. The ANOVA tests showed that the greatest positive change occurred in third grade. These students had more instruction in the skills that fall under the word recognition strand. This finding implies that Year 1 and the word recognition components had a greater impact on student achievement. Professional development targeted at the word recognition strand provides teachers with different instruction practices than professional development that targets the language comprehension strand. While teachers need professional development in both areas, the differences in learned techniques may affect the outcomes of student achievement

(Didion et al., 2020).

EOG assessments show measures of how well students have mastered the literacy standards. Hoover and Tunmer (2020) maintained that explicit mapping of literacy standards onto the cognitive framework can enhance educator understanding of the relationship between the two constructs. Reading Foundations standards directly relate to the word recognition component of the cognitive framework (Hoover & Tunmer, 2020). North Carolina's Reading Foundations standard states that students will know and apply phonics and appropriate developmental word analysis skills by decoding. This standard is more in-depth in second and third grades as opposed to fourth and fifth grades. Second grade contains six sub-levels that support the overarching phonics skill, and third grade has four. The same standard for fourth and fifth grades includes only one sublevel; thus, lower grades have a greater concentration of standards related to the word recognition component of the cognitive framework.

The lower grade connection to the word recognition component of the cognitive framework is apparent; however, the cognitive framework maintains that skills of both strands must be present for adequate reading comprehension. The gain in EOG means presented in the results of the analysis indicates a progression in reading development. Hoover and Tunmer (2020) asserted that "reading comprehension depends completely on the lower-order skills of language comprehension and word recognition" (p. 32); therefore, proficiency on the EOG assessment indicates that comprehension developed under these constructs.

Impact of Professional Development on School Growth Scores

The second part of this study investigated whether LETRS's professional

development impacted the school growth index scores. These scores differ from the EOG assessment scores. These scores derive from the EVAAS system used for educator accountability. This system compares student current achievements to all previous achievements on assessments. Growth scores for all grades assessed in reading are combined to generate one growth index score for the school in reading. The analysis comparing the change in growth index scores of schools from Year 2022 to Year 2023 showed no statistically significant change. The results show no impact on reading growth from the professional development when the scores are analyzed in this manner. No statistical change could be correlated to the lack of data for the first year of the professional development. No school growth index scores were issued for the 2020-2021 school year. These results are consistent with studies that indicate value-added systems can be less sensitive to changes in growth as opposed to other measures (Soland, 2021).

Limitations

Restrictions and constraints that are beyond the researcher's control and may impact the study's outcome are considered limitations (Theofanidis & Fountouki, 2019). One limitation was that DIBELS assessment data were not available for kindergarten through second grade. As a result, the study was limited to assessment data for Grades 3-5. Because of this restriction, I presented connections between the skills assessed in the lower grades and those assessed in Grades 3-5 during my literature review. In addition, the relationship between the assessments was noted. The DIBELS assessment is used to track reading proficiency in Grades K-2. It is also used in addition to the EOG to track reading proficiency in Grade 3. Smith et al. (2020) found that there are similarities in the skills measured on the third-grade EOG and the third-grade DIBELS assessment;

however, an assessment that measures comprehension is also a measure of mastery of early literacy skills. If students are not proficient in all skills contributing to comprehension, they will not show proficiency on assessments focusing on more sophisticated comprehension levels (Didion et al., 2020).

Another limitation is the impact of the COVID-19 pandemic. The pandemic closed schools and created a lull in learning. As a result, many assessments of reading proficiency declined (Donnelly & Patrinos, 2022; Hammerstein et al., 2021). Some studies analyzing standardized assessment data for the 2020-2021 school year indicated a slight increase in scores (Almasi & Yuan, 2023). The change in instruction, as many students returned to the classroom for face-to-face instruction, could have affected the scores. It is impossible to differentiate the impact of the instructional environment versus the professional development on the assessment scores; however, this study was not conducted using a causal design. Therefore, direct cause-and-effect relationships cannot be inferred, and the results should be interpreted cautiously.

In addition, Research Question 2 presented its own limitations. North Carolina chose not to collect school growth scores for the 2020-2021 school year; therefore, the study could only analyze professional development growth from Year 1 to Year 2. Adding more years to the analysis could be useful for understanding changes in reading growth at the school level.

Delimitations

In contrast to limitations, delimitations refer to the variables and outcomes that arise from the researcher's decisions and choices (Theofanidis & Fountouki, 2019). Delimitations are meant to make the study manageable and are within the researcher's

control. A delimitation of this study is that I chose to exclude demographic variables of gender and socioeconomic status. Instead, I chose the demographic variable of race. This decision kept the data set manageable and provided more in-depth analysis. In addition, the variable of race offered insight into the historical achievement gap.

Using data only from the schools in Cohort 1 was another delimitation of the study. The sample size was smaller by excluding the other two cohorts involved in the professional development. Including the other two cohorts would increase the generalizability of the study; however, the completion of the professional development varied among regions and schools. According to NCDPI, Cohort 1 completed the professional development at the end of the 2023 school year. This cohort had the greatest number of districts and schools that had finished the professional development.

Recommendations for Practice

The LETRS professional development provides teachers with pedagogy and instructional practices based on the science of reading (Moats & Tolman, 2019). Decades of research have shown that the science of reading should guide literacy instruction (Castles et al., 2018; Moats, 2020b; Petscher et al., 2020; The Reading League, 2022). Specific professional development on evidence and research-based practice can increase teacher knowledge and improve student reading achievement (Scarparolo & Hammond, 2018). The results showing increased means of EOG scores after the LETRS professional development indicate that stakeholders should continue to provide professional development grounded in the science of reading.

The data indicated that the least growth in EOG means occurred in the upper grades. We know that the pandemic caused a greater percentage of students in the later

grades to lose essential instruction in early literacy skills (Crosson & Silverman, 2021; Donnelly & Patrinos, 2022; Hammerstein et al., 2021; Molnar & Hermann, 2023). Educators should ensure that teachers in fourth grade and beyond adequately understand and can teach the skills typically acquired in earlier years (Didion et al., 2020; Kuhfeld et al., 2023).

Finally, educators should also focus professional development on culturally responsive teaching to address the achievement gap. While the third-grade data suggested that research-focused professional development impacted the disparities in means of the subgroups, the gap remained large. Educators must understand and empathize with the differences unique to minority cultures to impact achievement gaps (Kafele, 2021; Mayfield, 2020). Professional development must increase educator understanding of racial inequities (Mayfield, 2021). When teachers understand the impact of culturally responsive practices in reading, student achievement is positively impacted (Cantrell et al., 2023).

Recommendations for Additional Research

There is a strong link between the DIBELS assessment and the EOG assessment in third grade (Smith et al., 2020). North Carolina is in the process of creating new policies regarding the release of assessment scores for research purposes, and the data were not available at the time of this research. Future research investigating changes to reading proficiency on the primary assessment is needed. Access to early literacy data in Grades K-2 would allow longitudinal comparisons to EOG data as students progress through the grades. Kuhfeld et al. (2023) argued that analyzing assessments from lower and upper grades is important and needs research.

Additional qualitative research is also needed. Many educators have difficulty understanding research that could impact their practice (Hoover & Tunmer, 2020). LETRS is heavily founded on research, and much of the professional development is geared toward teacher understanding of it (Moats & Tolman, 2019). Moreover, teachers often fail to apply knowledge and learned research to their instructional practices (Seidenberg, 2017; Wexler, 2019). More research is needed to investigate how well teachers understand and implement the pedagogy learned in the LETRS professional development.

In addition, McKeown et al. (2019) asserted that teachers can provide valuable insight into what constitutes effective professional development. Teacher perceptions and beliefs toward professional development can affect the outcome (Scarparolo & Hammond, 2018). Rodgers et al. (2022) found that teachers linked their students' high academic performance to the knowledge gained from professional development; therefore, a qualitative investigation into teacher perceptions of LETRS and how that professional development contributed to the results of this study that show an increase in student achievement is warranted. In addition, this type of research design can provide knowledge of how well the professional development was understood and applied. Teachers absorb professional development differently and will need different support levels to succeed (Hall & Hord, 2020; Lewis & Goodwin, 2021).

Summary

This study investigated the impact of LETRS professional development on reading proficiency for Cohort 1. The results indicated an increase in reading proficiency for all grade levels from the baseline year of 2021 to the end of the professional

development in 2023. Third- and fourth-grade assessments showed a larger increase in reading proficiency than fifth-grade assessments. This result indicates that professional development had a negligible impact over the years.

Additional results signified that there were significant differences in assessment scores between the White race and the minority subgroups of Black, Hispanic, and Multiracial at each grade level. For third grade, the disparity among means declined somewhat over the course of the professional development; however, this decline was not found in the fourth- and fifth-grade results. These findings indicate that the achievement gap remains evident in reading proficiency assessments after 2 years of LETRS professional development.

Professional development is essential to policymakers and stakeholders who invest substantial resources into improving teacher knowledge. Measuring the change that occurs in reading development as a result of professional development can be difficult; however, professional development that is based on research-based practices can positively impact student achievement (Basma & Savage, 2018, 2023; Didion, 2020). This study showed that explicit professional learning based on science and reading and grounded in the Cognitive Foundations of Reading Framework can positively impact students.

References

Ainsworth, L., & Donovan, K. (2019). *Rigorous curriculum design: How to create curricular units of study that align standards, instruction, and assessment* (2nd ed.). International Center for Leadership in Education.

Almasi, J. F., & Yuan, D. (2023). Reading comprehension and the COVID-19 pandemic: What happened and what can we do about it. *The Reading Teacher*, *77*(3), 383-391.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Amplify. (2022). Amid academic recovery in classrooms nationwide, risks remain for youngest students with least instructional time during critical early years. https://amplify.com/wp-content/ uploads/2022/02/mCLASS_MOY-Results_February-2022-Report.pdf

Backstrom, B. (2019). *School turnaround efforts: What's been tried, why those efforts failed, and what to do now*. Nelson A. Rockefeller Institute of Government. https://rockinst.org/wp-content/uploads/2019/07/7-23-19-School-Turnaround-Efforts.pdf

Balota, D. A, Yap, M. J., Hutchinson, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39,* 445-459. https://doi.org/10.3758/BF03193014

Basma, B., & Savage, R. (2018). Teacher professional development and student literacy growth: A systematic review and meta-analysis. *Educational Psychology Review, 30*(2), 457-481. https://doi.org/10.1007/s10648-017-9416-4

Basma, B., & Savage, R. (2023). Teacher professional development and student reading in middle and high school: A systematic review and meta-analysis. *Journal of Teacher Education, 74*(3), 214-228. https://doi.org/10.1177/00224871231153084

Beech, R. J. (2005). Ehri's model of phases of learning to read: A brief critique. *Journal of Research in Reading*, *28*(1), 50-58.

Burk, K. (2022, December 02). How Mississippi reformed reading instruction. https://excelined.org/2022/12/02/how-mississippi-reformed-reading-instruction/

Cantrell, S. C., Sampson, S. O., Perry, K. H., & Robershaw, K. (2023). The impact of professional development on in service teachers' culturally responsive practices and students' reading achievement. *Literacy Research and Instruction, 62*(3), 233-259.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Chai, H., Elston, A. E., & Kramer, T. A. (2020). Focus on literacy needs: Partnering to support balanced literacy instruction. *Reading Improvement,57*(3), 117-130. https://go.exlibris.link/G4ct901L

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

Coyne, M. D., & Loftus-Rattan, S. M. (2022). Structured literacy interventions for vocabulary. In L. Spear-Swerling (Ed.), *Structured literacy interventions: Teaching students with reading difficulties, grades K-6* (pp. 114-135). The Guildford Press.

Creswell, W. J., & Creswell, D. J. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.

Crosson, A. C., & Silverman, R. D. (2021). Impact of COVID-19 on early literacy instruction for emergent bilinguals. *Reading Research Quarterly, 57*(1), 5-14.

Crow, T. (2022). New look, new standards: As the learning landscape shifts, so do we. *The Learning Professional*, *43*(3), 28-33.

Curriculum Associates. (2021). Academic achievement at the end of the 2020–2021 school year: Insights after more than a year of disrupted teaching and learning. https://www.curriculumassoc iates.com/-/media/mainsite/files/i-ready/iready-understanding-student-needs-paper-spring-resul ts-2021.pdf

Darling-Hammond, L. (2022). Breaking the legacy of teacher shortages. *Educational Leadership, 80*(2), 14-20.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. https://doi.org/10.54300/122.311

Dennis, D. V. (2017). Learning from the past: What ESSA has the chance to get right. *Reading Teacher, 70*(4), 395-400.

Didion, L., Toste, J. R., & Filderman, M. J. (2020). Teacher professional development and student reading achievement: A meta-analytic review of the effects. *Journal of Research on Educational Effectiveness, 13*(1), 29. https://doi.org/10.1080/19345747.2019.1670884

Dintersmith, T. (2018). *What school could be.* Princeton University Press.

Donnelly, R., & Patrinos, H. A. (2022). Learning loss during COVID-19: An early systematic review. *Prospects, 51,* 601–609. https://doi.org/10.1007/s11125-021-09582-6

DuFour, R., DuFour, R., Eaker, R., Many, W. T., & Mattos, M. (2016). *Learning by doing: A handbook for professional learning communities at work* (3rd ed.). Solution Tree.

Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly, 56*(S1), S25-S44. https://doi.org/10.1002/rrq.411

Duke, N. K., & Ward, A. E. (2021). The science of reading comprehension instruction. *The Reading Teacher, 74*(6), 663-672. https://doi.org/10.1002/trtr.1993

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2), 116. https://doi.org/10.1111/j.1467-9817.1995.tb00077.x

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Readin*g, *18*(1), 5-21. https://doi.org/10.1080/10888438.2013.819356

Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly*, 55(S1), S45-S60. https://doi.org/10.1002/rrq.334

Ehri, L. C., & Flugman, B. (2018). Mentoring teachers in systematic phonics instruction: Effectiveness of an intensive year-long program for kindergarten through 3rd grade teachers and their students. *Reading and Writing, 31*, 425 - 456. https://doi.org/10.1007/s11145-017-9792-7

Excellent Public Schools Act of 2021, Stat. 387. (2021). https://www.ncleg.gov/BillLookup/2021/S387

Fisher, D., Frey, N., & Lap, D. (2021). Veteran teachers' understanding of balanced literacy. *Journal of Education, 203*(1), 188-185.

Flesch, R. (1955). *Why Jonny can't read and what you can do about it.* Harper & Bros.

Floden, R. E., Richmond, G., & Salazar, M. (2020). A nation at risk or a nation in progress? Naming the way forward through research in teacher education. *Journal of Teacher Education, 71*(2), 169-171. https://doi.org/10.1177/0022487119900628

Fogarty, R. J., Kerns, G. M., & Pete, B. M. (2020). *Literacy reframed: How a focus on decoding, vocabulary, and background knowledge improves reading comprehension*. Solution Tree.

Folsom, J. S., Smith, K. G., Burk, K., & Oakley, N. (2017). *Educator outcomes associated with implementation of Mississippi's K–3 early literacy professional development initiative* (REL 2017–270). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. https://files.eric.ed.gov/fulltext/ED573545.pdf

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., Furgeson, J., Hayes, L., Henke, J., Justice, L., Keating, B., Lewis, W., Sattar, S., Streke, A., Wagner, R., & Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade* (NCEE 2016-4008). National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.

Foster, E. (2022). The research foundation for standards for professional learning. *The Learning Professional*, *43*(3), 22-26.

Garet, M., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H., Doolittle, F., Zhu, P., & Sztejnberg, L. (2008). *The impact of two professional development interventions on early reading instruction and achievement*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Glatthorn, A. A., Boschee, F., Whitehead, B. M., & Boschee, B. F. (2019). *Curriculum leadership: Strategies for development and implementation* (5th ed.). SAGE.

Goodman, K. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, *6*, 125–126.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6-10. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ee3c58058290e1ff077c96fa87f8fc8775d49b33

Gui, G. E. (2020). Changing the conversation about teacher development–Is it, truly, a "mirage?" An analytical review of the mirage: Confronting the hard truth about our quest for teacher development and a practitioner's perspective on teacher development. *Journal of Education & Social Policy*, *7*(4), 82-89. https://doi.org/10.30845/JESP.V7N4P7

Hall, G. E., & Hord, S. M. (2020). *Implementing change: Patterns, principles, and potholes* (5th ed.). Pearson.

Hammerstein, C., König, T., & Frey, D. A. (2021). Effects of COVID-19 related school closures on student achievement. *Frontiers in Psychology, 12*, 1-8. https://doi.org/10.3389/fpsyg.2021.746289

Hanford, E. (2018, September 10). *Hard words: Why aren't kids being taught to read?* American Public Media. https://www.apmreports.org/story/2018/09/10/hard-words-why-american-kids-arent-being-taught-to-read

Hanford, E. (2019, August 22). *At a loss for words: How a flawed idea is teaching millions of kids to be poor readers.* American Public Media. https://www.apmreports.org/episode/2019/08/22/whats-wrong-how-schools-teach-reading

Hanford, E. (2022, October 20). *Sold a story.* American Public Media. https://features.apmreports.org/sold-a-story/

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J., & Clarke, S. (2019). *Visible learning: Feedback*. Routledge.

Hennessy, N. L. (2021). *The reading comprehension blueprint: Helping students make meaning from text.* Paul H. Brookes Publishing Co.

Hensley, T., Turner, S., Drill, K., Hill, A., & Sharpe, D. (2017). *North Carolina read to achieve: An inside look.* Foundation for Excellence in Education. https://files.eric.ed.gov/fulltext/ED574423.pdf

Hiebert, E. H. (2022). When students perform at the below basic level on the NAEP: What does it mean and what can educators do? *The Reading Teacher, 75*(5), 631-639. https://doi.org/10.1002/trtr.2082

Hirsch Jr., J. D. (2020). *How to educate a citizen*. HarperCollins Publishers.

Hoover, W. A., & Tunmer, W. E. (2018). The simple view of reading: Three assessments of its adequacy. *Remedial and Special Education, 39*(5), 304–312.

Hoover, W. A., & Tunmer, W. E. (2020). *The cognitive foundations of reading and its acquisition: A framework with applications connecting teaching and learning.* Springer.

Horowitz, R., Ed, & Samuels, S. J., Ed. (2017). *The achievement gap in reading: Complex causes, persistent issues, possible solutions*. Routledge, Taylor & Francis Group.

Hougen, M., & Smartt, S. (2020). *Fundamentals of literacy instruction & assessment.* Paul H. Brookes Publishing Co.

Howard, T. (2020, Nov 15). *North Carolina annual testing program* [PowerPoint slides]. https://webservices.ncleg.gov/ViewDocSiteFile/72372

Hudson, A. K., Moore K. A., Han, B., Koh, P. W., Binks-Cantrell, E., & Joshi, R. M. (2021). Elementary teachers' knowledge of foundational literacy skills: A critical piece of the puzzle in the science of reading. *Reading Research Quarterly, 56*(S1), S287-S315.

Hudson, R. F., Anderson, E. M., McGraw, M., Ray, R., & Wilhelm, A. (2022). Structured literacy interventions for reading fluency. In L. Spear-Swerling (Ed.), *Structured literacy interventions: Teaching students with reading difficulties, grades K-6* (pp. 114-135). The Guildford Press.

International Literacy Association. (2018). *The case for children's rights to read.* https://www.literacyworldwide.org/docs/default-source/resource-documents/the-case-for-childrens-rights-to-read.pdf

International Literacy Association. (2019). *Meeting the challenges of early literacy phonics instruction*. https://www.literacyworldwide.org/docs/default-source/where-we-stand/ila-meeting-challenges-early-literacy-phonics-instruction.pdf

Kafele, B. K. (2021). *The equity & social justice education: 50 Critical questions for improving opportunities and outcomes for Black students.* ASCD.

Kilpatrick, D. A. (2015). *Essentials of assessing, preventing, and overcoming reading difficulties.* Wiley.

Kilpatrick, D. (2016). *Equipped for reading success: A comprehensive, step-by-step program for developing phonemic awareness and fluent word reco*gnition. Casey & Kirsch Publishers.

Kilpatrick, D. A. (2020). The article that introduced the simple view of reading. *The Reading League Journal,* 13-14. https://www.thereadingleague.org/wp-content/uploads/2020/06/13-14_Kilpatrick.pdf?mc_cid=a62105b477&mc_eid=6a8c71dc23

Knight, J. (2018). *The impact cycle: What instructional coaches should do to foster powerful improvements in teaching.* Corwin.

Knight, J. (2021). Moving from talk to action in professional development. *Educational Leadership*, *78*(5), 16-21.

König, C., & Frey, A. (2022). The impact of COVID-19 related school closures on student achievement: A meta-analysis. *Educational Measurement: Issues and Practice*, *41*(1), 16–22.

Kuhfeld, M., Lewis, K., & Peltier, T. (2023). Reading achievement declines during the COVID-19 pandemic: Evidence from 5 million U.S. students in grades 3–8. *Reading and Writing*, *36*(2), 245-261. https://doi.org/10.1007/s11145-022-10345-8

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549–565.

Lane, H., & Contesse, V. (2022). *UFLI foundations: An explicit and systematic phonics program.* Ventris Learning.

Lewis, D., & Goodwin, B. (2021). Research matters: Building the implementation bridge. *Educational Leadership, 78*(8), 74-75.

Mayfield, V. (2020). *Cultural competence now: 56 Exercises to help educators understand and challenge bias, racism, and privilege.* ASCD.

Mayfield, V. (2021). Learning to challenge racial colorblindness. *Educational Leadership, 78*(5), 74-75.

McKeown, D., Brindle, M., Harris, K. R., Sandmel, K., Steinbrecher, T. D., Graham, S., Lane, K. L., & Oakes, W. P. (2019). Teachers' voices: Perceptions of effective professional development and classwide implementation of self-regulated strategy development in writing. *American Educational Research Journal, 56*(3), 753-791. https://doi.org/10.3102/0002831218804146

Merga, M. K. (2020). "Fallen through the cracks": Teachers' perceptions of barriers faced by struggling literacy learners in secondary school. *English in Education, 54*(4), 371-395. https://doi.org/10.1080/04250494.2019.1672502

Moats, L. C. (2020a). *Speech to print* (3rd ed.). Paul H. Brookes Publishing Co.

Moats, L. C. (2020b). Teaching reading is rocket science: What expert teachers of reading should know and be able to do. *American Educator, 44*(2), 4-9. https://files.eric.ed.gov/fulltext/EJ1260264.pdf

Moats, L. C., & Tolman, C. A. (2019). *LETRS* (3rd ed.). Voyager Sopris Learning Inc.

Molnar, G., & Hermann, Z. (2023). Short and long-term effects of COVID related kindergarten and school closures on first to eight-grade students' school readiness skills and mathematics, reading and science learning. *Learning and Instruction, 83*, 101706. https://doi.org/10.1016/j.learninstruc.2022.101706

National Assessment of Educational Progress. (2022). *The nation's report card: 2022 Reading snapshot report: North Carolina grade 4 reading.* National Center for Educational Statistics. https://nces.ed.gov/nationsreportcard/subject/publications/stt2022/pdf/2023010NC4.pdf

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* National Institute of Child Health and Human Development.

Neuman, S. B., Quintero, E., & Reist, K. (2023). *Reading reform across America: A survey of state legislation.* Albert Shanker Institute. https://www.shankerinstitute.org/sites/default/files/2023-07/ReadingReform%20ShankerInstitute%20FullReport.pdf

National Early Literacy Panel. (2008). *Developing early literacy: Report of the national early literacy panel.* National Institute for Literacy.

North Carolina Department of Public Instruction. (2023, December 19). *End of grade reading tests at grades 3-8 North Carolina test specifications*. https://www.dpi.nc.gov/documents/accountability/testing/eog/eog-reading-grades-3-8-test-specifications

North Carolina Department of Public Instruction Accountability Services. (2021, February). *Assessment development process.* https://www.dpi.nc.gov/narrative-items-forms-and-selectionspdf/open

North Carolina Department of Public Instruction Office of Early Learning. (2023, August 3). *Year-end literacy assessment shows strong gains continuing for NC early grades students* [Press Release]. https://www.dpi.nc.gov/news/press-releases/2023/08/03/year-end-literacy-assessment-shows-strong-gains-continuing-nc-early-grades-students

Otaiba, A. S., Allor, J. H., & Sterwart, J. (2022). Structured literacy interventions for phonemic awareness and basic word recognition skills. In L. Spear-Swerling (Ed.), *Structured literacy interventions: Teaching students with reading difficulties, grades K-6* (pp. 114-135). The Guildford Press.

Peng, P., & Goodrich, J. M. (2020). The cognitive element model of reading instruction. *Reading Research Quarterly, 55*(S1), S77-S88. https://doi.org/10.1002/rrq.336

Petscher, Y., Cabell, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., Lonigan, C. J., Phillips, B. M., Schatschneider, C., Steacy, L. M., Terry, N. P., & Wagner, R. K. (2020). How the science of reading informs 21st-century education. *Reading Research Quarterly, 55*(S1), S267– S282. https://doi.org/10.1002/rrq.352

Pfeffer, J., & Sutton, R. (2000). *The knowing-doing gap: How smart companies turn knowledge into action*. Harvard Business School Press.

Piasta, S. B., & Hudson, A. K. (2022). Key knowledge to support phonological awareness and phonics instruction. *The Reading Teacher, 76*(2), 201-210.

The Reading League. (2022). *The science of reading: Defining guide*. https://www.thereadingleague.org/wp-content/uploads/2022/03/Science-of-Reading-eBook-2022.pdf

Rebora, A. (2021). Beyond checking boxes: Re-envisioning pd in schools. *Educational Leadership, 78*(5), 9.

Rebora, A. (2022a). Beyond lip service on student well-being. *Educational Leadership, 79*(9), 12.

Rebora, A. (2022b). A profession on the brink. *Educational Leadership, 80*(2), 9.

Reeves, D. (2021). Five professional learning transformations for a post-COVID world. *Educational Leadership, 78*(5), 44-48.

Relyea, J. E., Rich, P., Kim, J. S., & Gilbert, J. B. (2023). The COVID-19 impact on reading achievement growth of grade 3–5 students in a U.S. urban school district: Variation across student characteristics and instructional modalities. *Reading and Writing, 36*(2), 317-346. https://doi.org/10.1007/s11145-022-10387-y

Riley, B. (2020). Drawing on reading science without starting a war. *Educational Leadership, 77*(5), 16-22.

Rodgers, E., D'Agostino, J., Berenbon, R., Mikita, C., Winkler, C., & Wright, M. E. (2022). Teachers' beliefs and their students' progress in professional development. *Journal of Teacher Education, 73*(4), 381-396.

Sanden, S., MacPhee, D. A., Hartle, L., Poggendorf, S., & Zuiderveen, C. (2022). The status of phonics instruction: Learning from the teachers. *Reading Horizons: A Journal of Literacy and Language Arts, 61*(1), 69-92. https://scholarworks.wmich.edu/reading_horizons/vol61/iss1/5

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy*. (pp. 97-110). Guilford Press. https://johnbald.typepad.com/files/handbookearlylit.pdf

Scarparolo, G. E., & Hammond, L. S. (2018). The effect of a professional development model on early childhood educators' direct teaching of beginning reading. *Professional Development in Education, 44*(4), 492-506. https://doi.org/10.1080/19415257.2017.1372303

Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2022). Long-term consequences of repeated school closures during the COVID-19 pandemic for reading and mathematics competencies. *Frontiers in Education, 13*.

Seidenberg, M. (2017). *Language at the speed of sight: How we read, why so many can't, and what can be done about it*. Basic Books.

Seidenberg, M., Borkenhagen, M., C., & Kearns, D. M. (2020). Lost in translation? Challenges in connecting reading science and educational practice. *Reading Research Quarterly, 55*(51), S119-S130.

Semingson, P., & Kerns, W. (2021). Where is the evidence? Looking back to Jeanne Chall and enduring debates about the science of reading. *Reading Research Quarterly, 56*(S1), S157-S169.

Shanahan, T. (2020). What constitutes a science of reading instruction? *Reading Research Quarterly, 55*(S1),1-13.

Shanahan, T., Callison, K., Carriere, C., Duke, N., Pearson, D. P., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: A practice guide* (NCEE 2010-4038). National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.

Shanahan, T., & Lonigan, C. J. (2010). The national early literacy panel: A summary of the process and the report. *Educational Researcher, 39*(4), 279-285.

Smith, C., Amendum, S. J., & Jang, B. G. (2020). Predicting performance on a 3rd grade high-stakes reading assessment. *Reading & Writing Quarterly, 36*(4), 365.

Smith, F. (1971). *Understanding reading a psycholinguistic analysis of reading and learning to read.* Holt Rienhart and Winston.

Soland, J. (2021). Are schools deemed effective based on overall student growth also closing achievement gaps? Examining the Black–White gap by school. *Teachers College Record, 123*(12), 211-236. https://doi.org/10.1177/01614681211070876

Solari, E. J., Terry, N. P., Gaab, N., Hogan, T. P., Nelson, N. J., Pentimonti, J. M., Petscher, Y., & Sayko, S. (2020). Translational science: A road map for the science of reading. *Reading Research Quarterly*, *55*(S1), S347– S360.

Spear-Swerling, L. (2018). Structured literacy and typical literacy practices: Understanding differences to create instructional opportunities. *Teaching Exceptional Children*, *51*(3), 201-211. https://drive.google.com/file/d/1DGh8M3yJkHIbzlJscyzT5fpnPajMOz6T/view

Spear-Swerling, L. (2022). Structured literacy interventions for phonemic awareness and basic word recognition skills. In L. Spear-Swerling (Ed.), *Structured literacy interventions: Teaching students with reading difficulties, grades K-6* (pp. 1-22). The Guildford Press.

Stiggins, R. (2018). Better assessments require better assessment literacy. *Educational Leadership, 75*(5), 18-19.

Stouffer, J. (2021). Seeking middle ground: Analyzing running records from the top and bottom. *The Reading Teacher, 74*(6), 769-784.

Sun, X., Marks, R. A., Eggleston, R. L., Zhang, K., Lau, C., Yu, C.-L., Nickerson, N., & Kovelman, I. (2023). Impacts of the COVID-19 disruption on the language and literacy development of monolingual and heritage bilingual children in the United States. *Reading and Writing, 36,* 347–375. https://link.springer.com/article/10.1007/s11145-022-10388-x?utm_source=getftr&utm_medium=getftr&utm_campaign=getftr_pilot

Tassell, R. V. (2014). The trouble with top-down. *Educational Leadership, 71*(8), 76-78.

Theofanidis, D., & Fountouki, A. (2019). Limitations and delimitations in the research process. *Perioperative Nursing, 7*(3), 155–162. http://doi.org/10.5281/zenodo.2552022

University of Oregon. (2021). *8th Edition of dynamic indicators of basic early literacy skills: Administration and scoring guide*. University of Oregon. https://dibels.uoregon.edu

Vernon-Feagans, L., Mokrova, I. L., Carr, R. C., Garrett-Peters, P. T., Burchinal, M. R., & Family Life Project Key Investigators. (2019). Cumulative years of classroom quality from kindergarten to third grade: Prediction to children's third grade literacy skills. *Early Childhood Research Quarterly, 47,* 531–540.

Wexler, N. (2019). *The knowledge gap.* Penguin Random House LLC.

White, S., Sabatini, J., Park, B. J., Chen, J., Bernstein, J., & Li, M. (2021). *The 2018 NAEP oral reading fluency study* (NCES 2021-025). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021025

Willingham, D. T. (2017) *The reading mind: A cognitive approach to understanding how the mind reads.* Jossey-Bass.