12-2017

Exploring the Relationship Between the Use of a Selected Phonics Curriculum and the Oral Reading Fluency and Nonsense Word Fluency Scores of First-grade Students

Bryce B. Day
Utah State University

EXPLORING THE RELATIONSHIP BETWEEN USE OF A SELECTED PHONICS PROGRAM AND THE ORAL READING FLUENCY AND NONSENSE WORD FLUENCY SCORES OF FIRST-GRADE STUDENTS

by

Bryce B. Day

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

of

DOCTOR OF PHILOSOPHY

in

Education

Approved:

Kathleen A. J. Mohr, Ed.D. Major Professor	Marla Robertson, Ph.D. Committee Member
Susan A. Turner, Ph.D. Committee Member	Sylvia Read, Ph.D. Committee Member
Marietta A. Veeder, Ph.D. Committee Member	Mark R. McLellan, Ph.D. Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2017

ABSTRACT

Exploring the Relationship Between the Use of a Selected Phonics Curriculum and the Oral Reading Fluency and Nonsense Word Fluency Scores of First-grade Students

by

Bryce B Day, Doctor of Philosophy

Utah State University, 2017

Major Professor: Kathleen A. J. Mohr, Ed.D.
Department: Teacher Education and Leadership

The purpose of this quantitative study was to examine the effects, if any, of a supplemental phonics program, Saxon Phonics, on the reading achievement of first-grade students in one mountain-west, semi-rural, school district. The design was casual-comparative and *ex post facto*, and answered the questions: (1) Do students taught using a traditional basal program and students taught using both the traditional basal program and a supplemental phonics program (control vs. treatment) differ on selected end-of-year reading achievement scores (i.e., portions of the DIBELS Next assessment—nonsense word fluency [NWF], oral reading fluency [ORF], and accuracy [ACC])? (2) Do any possible interactions among selected variables (i.e., instructional program and beginning-of-year reading level) exist related to performance differences on end-of-year reading achievement scores among students receiving reading instruction with or without a supplemental phonics program? The independent variables were the instructional

program Saxon Phonics, a traditional/basal reading curriculum and the reading levels of low, medium, and high. The dependent variables were oral reading fluency, accuracy and nonsense word fluency, measured by the DIBELS Next assessment. The 2014-2015 and 2015-2016 DIBELS Next data were collected from the school district database upon approval from the Institutional Review Board in January of 2017. A mixed effects model was utilized to explore the relationship between use of the selected supplemental phonics curriculum and selected reading achievement scores of first-grade students. Results revealed that there was no significant difference between the control and treatment groups, though there was a statistically significant improvement of low readers in the treatment group over the control group.

(117 pages)

PUBLIC ABSTRACT

Exploring the Relationship Between the Use of a Selected Phonics Curriculum and the Oral Reading Fluency and Nonsense Word Fluency Scores of First-Grade Students

by

Bryce B. Day

The purpose of this quantitative study was to examine the effects, if any, of a supplemental phonics curriculum, Saxon Phonics, on the reading achievement of first-grade students in one mountain-west, semirural, school district. The design was casual-comparative and *ex post facto*, and answered the questions: (1) Do students taught using a traditional basal program and students taught using both the traditional basal program and a supplemental phonics program (control vs. treatment) differ on selected end-of-year reading achievement scores (i.e., portions of the DIBELS Next assessment—nonsense word fluency [NWF], oral reading fluency [ORF], and accuracy [ACC])? (2) do any possible interactions among selected variables (i.e., instructional program, gender, and beginning-of-year reading level) exist related to performance differences on end-of-year reading achievement scores among students receiving reading instruction with or without a supplemental phonics program? The independent variables were the instructional program Saxon Phonics, a traditional/basal reading curriculum and the reading levels of low, medium, and high. The dependent variables were oral reading fluency, accuracy and nonsense word fluency, measured by the DIBELS Next assessment. The 2014-2015 and

2015-2016 DIBELS Next data were collected from the school district database upon approval from the Institutional Review Board in January of 2017. A mixed effects model was utilized to explore the relationship between use of the selected supplemental phonics curriculum and selected reading achievement scores of first-grade students. Results revealed that there was no significant difference between the control and treatment groups, though there was a statistically significant improvement of low readers in the treatment group over the control group.

ACKNOWLEDGMENTS

The completion of a doctoral dissertation is a difficult, arduous, and yet somehow a personally gratifying task; however, as I reflect back on the process I know that I did not make it here on my own—only the support, encouragement, and coaching from numerous individuals have brought me to this journey's end. I am grateful to my family, friends, and colleagues who have supported me throughout this process. This accomplishment is not mine alone, but shared with all who have lifted me up along the way.

I express my gratitude to my major professor, Dr. Kathleen A. J. Mohr, for her guidance, encouragement, direction, and firm-hand throughout this process. She took this journey with me and persevered alongside me. I also wish to thank Dr. Marietta Veeder, for her friendship, mentoring, and constant positive encouragement in completing this degree. I also express my appreciation to all those at Utah State University who shaped my experience—the challenges and victories. Without those experiences, I would not stand here today, at the end of this journey. I extend my appreciation to the many educators with whom I work on a daily basis, especially those who continually cheered for me along the way. Without their willingness to allow me into their classrooms, I would not have been able to learn how to become a better educator.

I am immensely grateful to my parents, Kirk and Vickie, for instilling in me a desire and perseverance to set and accomplish goals. They taught me the importance of hard work and the value of family. I am grateful to my wife's parents, Scott and Phyllis, for being my cheerleaders—who always knew I was capable of doing this.

Finally, I am grateful to my wonderful family, my wife Melanie and son Alexander. I would not have completed this goal without the many nights I spent away from them. I wondered continuously if it was going to be worth it, but I was continually reassured, by those I love most, that it was. Their patience and sacrifice of time with me is a debt I can only attempt to repay.

Bryce B. Day

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Perhaps one of the most valuable tasks students undertake in school is learning how to read. This task is a complex skill set, which facilitates access to information and knowledge across an array of subjects and permeates every facet of life. By the end of first grade, students typically master an arsenal of reading skills (e.g., phonics, fluency, spelling, and writing) with which to tackle the remainder of their school career (Bos & Vaughn, 2002; National Reading Panel [NRP], 2000; Snow, Burns, & Griffin, 1998). The ability to meet reading-related goals requires a strong linguistic foundation, exposure to text and motivation to learn literacy, in addition to effective instructional programs that support beginning readers.

Although research indicates that teachers are an essential element in the acquisition of beginning reading skills (Bond & Dykstra, 1967), most schools provide instructors with programs and materials to guide and structure reading instruction. These programs vary with regard to content and emphasis depending on the current trends within the field. However, following the NRP's (2000) meta-analysis of rigorous reading research, the general consensus is that beginning reading instruction should include phonemic awareness and phonics as essential components. The question arises whether a packaged or published programs are an effective means of supporting beginning readers, and which program might yield significant results. The intense scrutiny focused on student achievement as an indicator of teacher quality has intensified the quest for effective and efficient reading methods.

The importance of finding efficient and effective instructional programs is clear given current student proficiency levels. For example, the Nation's Report Card for the State of Utah for 2015 indicated that only 40% of fourth-grade students were at or above proficiency in reading, as reported in the National Assessment of Educational Progress (NAEP; Utah State office of Education [USOE], 2015). One mountain-west school district is facing a similar issue at the local level. According to the Utah Education PACE Report Card, 45% of third-grade students within the district demonstrated reading proficiency on the DIBELS Next assessment at the end of the 2014-2015 school year (USOE, 2015). The Student Assessment of Growth and Excellence (SAGE) scores also indicate that English language arts is an area of weakness district-wide (USOE, 2015). Presented with such data, school districts are compelled to respond.

A large number of empirical studies have claimed considerable evidence regarding effective reading practices and instructional approaches (Adams, 1990; Anderson, Hiebert, Scott, & Wilkinson, 1985; NICHD, 2000; Snow et al., 1998). The NRP was established in 1997 by congressional mandate, and charged with assessing "the status of research-based knowledge, including the effectiveness of various approaches to teaching children to read" (p. 1-1). The findings of the NRP (2000) have been used to direct curricular development in reading since that time. The NRP concluded that quality reading instruction should include the following components: alphabetic (phonemic awareness and phonics), reading fluency, and comprehension (vocabulary and text comprehension). This conclusion is supported by a large body of research (Coley, 2015; Cooke, Kretlow, & Helf, 2009; Good, Simmons, & Kame'enui, 2001; Justice,

Kaderavek, Fan, Sofka, & Hunt, 2009; Kaplan & Walpole, 2005; Roskos, Strickland, Haase, & Malik, 2009; Strickland, 2010; Wilson & Lonigan, 2010).

Additionally, the NRP (2000) report stated:

> The most important grade for teaching phonics is thought to be first grade when formal instruction in reading typically begins in the United States. Children have foundational knowledge and are ready to put it to use in learning to read and write. In contrast, introducing phonics instruction in grades above first means that children who were taught to read in some other way may be required to switch gears in order to incorporate phonics procedures into their reading and writing. (NRP, 2000, pp. 2-105)

The NRP (2000) went on to note that "Correlational studies have identified phonemic awareness and letter knowledge as the two best school-entry predictors of how well children will learn to read during the first two years of instruction" (p. 2-9). These statements by the NRP clearly endorse that early reading instruction should be weighted heavily toward phonics instruction in the early grades (kindergarten, first, and second); thus suggesting that later grades (third and beyond) weighted toward fluency and comprehension (unless a student has failed to adequately master foundational skills).

Though one would conclude that the evidence is clear—quality early reading instruction should include all three elements identified by the NRP; McEwan (2002) found that many teachers *claim* to utilize a balanced approach (inclusion of alphabetics, fluency, and comprehension), but still shied away from direct instruction in phonological and phonetic skills. Bingham and Hall-Kenyon (2013) found most teachers reported the belief that teaching whole-word recognition with minimal focus on phonics was sufficient to develop proficient early readers. This stands in contrast to a large body of evidence that favors a code-based or phonics approach.

The tendency to weight instruction toward comprehension and fluency is evidenced by purported "comprehensive" or "balanced" programs adopted by school districts. For example, the basal program adopted by the local school district mentioned earlier, is the Macmillian/McGraw-Hill program Treasures (Bear & Bear, 2007). Treasures is purportedly a Common Core-aligned comprehensive reading and writing program. The program is structured into daily lessons that contain oral language, phonemic awareness, phonics, comprehension, high-frequency words, and writing components. Each day is divided into mini lessons based on needed skill development. A day's lesson might be broken up into separate skill-based mini lessons each with individual learning objectives and designated materials to be taught throughout the day, or combined during one designated instructional period. The program includes resource materials such as big books, read aloud books, student readers, student activity books, and teaching charts. Though the program claims to be "comprehensive," teachers at the research site believe it is weighted heavily toward fluency, vocabulary/word identification, and comprehension. Thus, given input from teachers and insufficient reading achievement, these districts' officials made the decision to supplement this basal program with a more explicit phonics program.

Purpose of the Study

The purpose of this quantitative study, therefore, was to explore the effects of systematic and structured phonics instruction when paired with basal reading instruction on indicators of beginning reading achievement (oral reading fluency, accuracy, and

nonsense word fluency) of first-grade students. A descriptive, causal-comparative, ex post facto design was used, and involved the collection of quantitative data in the form of DIBELS Next assessment scores for all first-grade students for the 2014-2015 and 2015-2016 school years. This study sought to determine if a difference in specified reading achievement existed between students instructed during the 2014-2015 school year using a basal reading program (without a direct instruction phonics component), and the students from the 2015-2016 school year, who were instructed using the same basal reading program paired with a systematic and structured direct instruction phonics program, Saxon Phonics (Simmons, 2003). This study sought provide additional evidence as to whether systematic and structured instruction in phonics significantly improves first-grade students' reading scores. The organic nature of this study (meaning it took place in authentic school contexts) added ecological validity to the claims that structured and systematic phonics instruction is important to the success of beginning readers.

It is important to note that there are few research studies that have examined specific prepackaged curricula—though many programs claim a "research base." The purpose of this study was to provide evidence whether or not to support the inclusion of systematic and structured phonics instruction, in this case via Saxon Phonics, in early grades. However, due to the authentic context of the proposed research site, and the nature of school curricular development and adoption—an element of program evaluation cannot be avoided. Because the research site has adopted the use of a specific prepackaged phonics program, this study could also be viewed, to a limited extent, as a program validation.

The local school district hosting this study has adopted the use of the Saxon Phonics (Simmons, 2003) to address the perceived weakness in the basal program. Saxon Phonics is a supplemental explicit instruction program, focusing on phonemic awareness, phonics, decoding, spelling, and fluency (Simmons, 2003; Simmons & Calvert, 1996). The program utilizes an incremental and spiraling design with continuous review. Each lesson is broken up into warm-ups (alphabet and phonological/phonemic awareness activities; daily letter, sound, and spelling reviews), new increment instruction (introducing student to the names and written forms of letters, letter clusters, and sight words with writing practice), application (boardwork and whole group practice—both aloud and written), and classroom/fluency practice (intervention or application time though use of games, independent reading, or re-teaching of decoding and fluency skills). Classroom/fluency practice portions of the lessons may be divided out and taught at a different point in the day.

The Buck and Torgesen (2003) have identified the strengths of the Saxon Phonics program to be "A multi-sensory approach and use of manipulatives, built-in assessments to monitor student progress and guide instruction, clearly stated learning objectives for every activity, and clearly linked instruction across components (p. 4). These claims are also supported by a research study authored by Leib (2001), finding that students instructed with the Saxon Phonics program made significant gains in reading achievement over students who received instruction through traditional basal curricula. This study may support these findings, but also speak to overarching idea that structured and systematic phonics instruction is integral to the success of beginning readers.

Research Questions

This study sought to answer the following research questions: (1) Do students taught using a traditional basal program and students taught using both the traditional basal program and a supplemental phonics program (control vs. treatment) differ on selected end-of-year reading achievement scores (i.e., portions of the DIBELS Next assessment—nonsense word fluency [NWF], oral reading fluency [ORF]), and accuracy [ACC])? and (2) Do any possible interactions among selected variables (i.e., instructional program, gender, and beginning-of-year reading level) exist related to performance differences on end-of-year reading achievement scores among students receiving reading instruction with or without a supplemental phonics program?

Significance of the Study

Reading is a major pathway to learning; therefore, it is imperative that students become successful readers early in their academic careers. Thus, schools and district must find effective and efficient methods for accomplishing this task, as evidenced by higher reading achievement scores. Jones (2006) reiterated the need for school personnel to select research-based programs and practices as outlined by congressional mandate. The term "research-based" is used to describe practices and programs that are supported by a well-defined theory or theories, indicating that a practice *should* work. "Evidence-based" refers to practices and programs that are supported by scientific evidence obtained in authentic contexts, indicating that a practice *does* work (Shaywitz, 2014). This clarification seems necessary as often the terms are used interchangeably in school

contexts, but are in fact not synonymous. This confusion, the researcher believes, is the cause for the culture within schools causing educators to become skeptical of products claiming a "research base." The challenge for school administrator, curriculum directors, and those charged with curriculum adoption is to troll through the massive corpus of resources available and to select programs that are supported by research, but also evidenced as has having significant and positive effect on reading achievement.

The results of this study are significant in three ways.

1. This study examined the differences, if any, between students instructed with a basal reading program, serving as a control group, and students who received instruction using the same basal reading program paired with a systematic phonics program. This study adds to the body of scientific evidence, obtained in actual school contexts, of the influence of systematic and structured phonics instruction when paired with a basal program.

2. This study addressed the variance of student performance in reading and explores the effect on student reading scores across reading levels labeled low, medium, and high and based on pretreatment assessment scores. Researchers have learned that one program does not prove best for all students (Sippola, 1985; Snell, 2007). However, educators are limited in time and resources and must make wise and often difficult decisions regarding the most appropriate use of what is available to them, while attempting to meet the needs of all their students.

3. Of the literature reviewed, only three articles focused specifically on Saxon Phonics (Baker, 2010; Leib, 2001; Wicker, 2007), none of which utilized an experimental design. Given the ex post facto research design and the highly contextual research site, this study functioned secondarily as a program evaluation. Though the ultimate goal was not to determine the efficacy of the Saxon Phonics program, conclusions can be drawn as to the success of the program with the identified sample population (or specific subsets of the sample population).

Hypotheses

This research study was grounded on two major assumptions drawn from

educational research by Hattie (2009), Marzano (1998, 2003, 2007), and Wiggins and McTighe (2005), as well as the work of Sippola (1985) and Snell (2007). First, in order for students to achieve at high levels educators need to establish and utilize a guaranteed and viable curriculum that provides opportunity to learn in a rigorous and relevant manner, and utilize instructional methodologies proven to have a large effect size $d = 0.4 – 1.2$, (Hattie, 2009). Second, that students' reading achievement is directly related to their abilities to learn from classroom instruction—that is their ability to engage in learning in a way that is accessible. Not all methods, practices, and/or programs are appropriate to address the needs of all students—some have been designed to address the needs of various populations and are found to apply broadly to many populations, while others are most effective in limited and very specific context and populations. This study targeted both the general outcomes, as well as the interaction between ability level and reading instruction that may play a role in reading achievement scores.

The researcher proposed two hypotheses.

1. There will be a difference in reading scores between the control group, who received basal reading program instruction, and the treatment group, who received basal-based instruction paired with the Saxon Phonics program.
2. The difference between groups could vary depending on pre-instructional reading levels of low, medium, and high.

Delimitations

This study utilized an ex post fact research design, described in chapter three. Due to the nature of this study design, and the pragmatic nature of the data selection, there are several variables which are not be included in the analyses. These variables may have,

however, played a role in predicting students' NWF, ORF, and ACC scores. This study did not examine demographic variables such as ethnicity and socioeconomic status. The school district does collect such data, but it is not paired with student achievement. The pairing of such data to each individual student would be time intensive, and would have required the researcher to access personally identifiable student information. Accessing such information requires written consent under the Family Educational Rights and Privacy Act (FERPA). The goal of the study was to determine the effect of the instructional use of Saxon Phonics on the early reading achievement of first graders. The purpose was not to know which individual students may or may not have benefitted. Using as many scores as possible enhances the statistical power of the analyses. Seeking written consent to identify individual student information would have likely reduced the sample size as parents may not have wished such information to be disclosed.

This study did not control for students who were retained in first-grade or students who received additional reading intervention programs. All students are monitored and provided reading interventions as problems arise in student performance, as identified by on-going progress monitoring. While an important aspect of contemporary literacy instruction, information about special services provided to individual students was not readily available to the researcher. Moreover, not all intervention procedures are the same from school to school, and some may have been more effective than others. The isolation of these procedures class by class, and school by school would have been massive in scope and prohibitive to the completion of the study.

Given the large geographic area of the school district, the variety of the

socioeconomic status across schools, and the number of schools and diversity of intervention programs within in the schools—controlling for these independent variables would have been prohibitive. Though not addressed specifically in this study, these variables could have had an impact on end-of-year reading scores examined in this study.

Limitations

The limitations specific to this study included the following.

1. Maturation of the students may have occurred over the course of the school years between the measures of NWF in the fall to ORF in the spring.
2. The research is limited to one mountain-west school district in northern Utah and, therefore, cannot necessarily be generalized to other populations.
3. The sample was comprised of those students who have beginning-of-year composite, as well as middle- and end-of-year NWF, ORF, and ACC scores. Students who did not have these scores were not present for the full 160 days of instruction, and their scores may be the result of other extraneous factors beyond the scope of this study.
4. DIBELS Next is a formative, screening assessment rather than an achievement test, such as Utah's Student Assessment of Growth and Excellence (SAGE) assessment, because first-grade students in the state of Utah are not administered an achievement test, DIBELS Next scores are currently the only available data.

Definitions

DIBELS: Dynamic Indicators of Basic Early Literacy Skill-a set of six assessments that measure indicators of the essential skills a student must master in order to become a proficient reader (University of Oregon, 2015).

DIBELS Next: The most updated version of the DIBELS assessment system, with new forms and reading passages, adjusted directions for the assessor and students, and

replacement of “sound fluency” with “first sound fluency” (University of Oregon, 2015).

DORF: DIBELS Oral Reading Fluency-a standardized and individually administered test of accuracy and fluency of a student’s reading of text (University of Oregon, 2015). Oral reading fluency is calculated by taking the total number of words read during a 1-minute timing and subtracting errors (words omitted, substituted, and hesitations of more than three seconds).

NWF: Nonsense Word Fluency-a standardized, individually administer test that measures a student’s ability to appropriately apply letter-sound correspondence to the most common sounds and the ability to blend letters into words (Kaminski & Good, 1996).

Preinstruction reading level: During the beginning of year DIBELS Next administration, composite scores that either hit or exceed benchmark (> 113) are identified as proficient or high; scores between the benchmark and the at-risk cut score (112 > 97) are identified as strategic or medium; and scores below the at-risk cut score (96 >) are identified as intensive or low.

CHAPTER II

REVIEW OF LITERATURE

Success in school is heavily dependent up a student's ability to access knowledge, often conveyed through text. Reading difficulties can hinder learning as a student progresses through school (Allington & McGill-Franzen, 2008; Chatterji, 2006). Consequently, early childhood educators are charged with the responsibility of preparing students for reading success by implementing and focusing on educational activities that promote acquisition of reading skills (Coley, 2015; Thompson, 2011). The related research repeatedly identifies first-grade as that time when the most foundational skills of reading are typically taught (NRP, 2000). These critical skills include phonics—the knowledge of letter-sound correspondences that enables rapid decoding of text.

To access the research related to early-grades phonics, a number of search terms and various databases were utilized to find relevant and current studies. The search for current phonics research identified a vast number of studies (N = 9,682). Table 1 describes the search results.

The search was narrowed by reviewing only titles and abstracts of literature published within the last 10 years (with the exception of seminal literature repeatedly cited throughout the literature). Abstracts of titles indicating a focus on evidence-based instructional practices and literature examining specific curricula were read. Additionally, all the abstracts of digital dissertation results were read. Dissertations that utilized an experimental or pre-post investigation method were searched in detail in order to identify relevant studies examining early literacy. Pertinent literature reviews were examined,

Table 1

The Databases, Search Terms, and Results (N Size) of Literature Searched

Database	Search terms	Results
Education Source (EBSCO)	Explicit phonics instruction	102
	Early phonics instruction	260
	Systematic phonics instruction	60
ERIC (EBSCO)	Explicit phonics instruction	133
	Early phonics instruction	525
	Systematic phonics instruction	184
PsycINFO via EBSCOhost	Explicit phonics instruction	90
	Early phonics instruction	194
	Systematic phonics instruction	96
Professional development collection	Explicit phonics instruction	1120
	Early phonics instruction	101
	Systematic phonics instruction	69
JSTOR	Explicit phonics instruction	1293
	Early phonics instruction	3631
	Systematic phonics instruction	1623
Digital Dissertations	Explicit phonics instruction	16
	Early phonics instruction	165
	Systematic phonics instruction	20

which afforded potentially more research literature for consideration. The following literature review represents a synthesis of the knowledge gleaned from these sources. This literature review begins with an overview of early literacy skills, a review of epistemological views of early reading instruction that led to the *Great Debate* (Chall, 1967), a review of the findings of the NRP (2000), the connection between reading and academic success, and instructional models that have developed from reading/literacy research.

Early Literacy Skills

The last few decades have wrought growing consensus among many in the

reading community on the skills that serve as the foundation for reading and English language arts. Bursuck et al. (2004) suggest that in order to make a difference for children, effective reading instruction needs to begin early and focus on phonological awareness, alphabetic understanding, reading fluency, vocabulary acquisition, and understanding of connected text, which mirrors outcomes of the NRP (2000) report. Bursuck et al. further postulate that in order for instructional programing to be effective, instruction needs to be prevention-based, intensive, and employ the use of a research-based curriculum and practices founded on explicit instructional design. The work of Ehri (1979) explains how students transition from recognizing that words consist of sounds to being able to blend and decode words, then to fluent and accurate decoding, which ultimately results in fluent readers. Fluent reading promotes expansion of students' personal reading lexicons, which leads to greater reading comprehension. As already noted, although teachers may teach five key reading components (phonological awareness, alphabetic understanding, reading fluency, vocabulary acquisition, and reading comprehension) and some instructional programing may contain them all, Bursuck and Blanks (2010) suggest that touching on these components is not enough to produce good readers. They concluded that in order for students to become effective and efficient readers, "they need to be taught systematically and explicitly using empirically based instructional design and delivery principles" (p. 424).

Language Development and Phonemic Awareness

A great body of research (Burgess, 2006; Lonigan, 2006, Stahl, 2001; Stockard &

Engelmann, 2010) indicates that phonemic awareness is an essential prerequisite, correlate, and predictor of children's reading achievement. Stahl defined phonemic awareness as "the ability to reflect on units of spoken language smaller than a syllable" (p. 341). Discriminating between these linguistic units (e.g., words, segments, phonemes) strongly correlates with success in reading (NRP, 2000). Typically developing (TD) children demonstrate the ability to discriminate between sounds (phonological awareness) and begin to discriminate individual phonemes (phonemic awareness; Ehri & Roberts, 2006). Evidence suggests that children develop syllabic sensitivity before they develop sensitivity to phonemes, and sensitivity to rhyme before sensitivity to phonemes (Lonigan, 2006; Whitehurst & Lonigan, 1998). Regular activities such as language games and nursery rhymes support this development and reinforce a child's acquisition of these language skills (Maclean, Bryant, & Bradley, 1987). Often language games are built around implicit comparison and contrast of sounds of words and include alliterative phrases. Implicit comparison, as stated by Adams (1990), may not be sufficient to develop these metalinguistic skills because students may not attend to the most important phonemic unit. In addition to recital and play with sound units, children must also develop an explicit understanding that sound units map onto larger units, which are used in language. Reviews and analyses by Dickinson, McCabe, and Sprague (2003) and Scarborough, Neuman, and Dickinson (2001) suggest that phonemic awareness is critical to the development of complex language skills and abilities, which include strands of semantics, syntax, pragmatics, and discourse. Though distinctly different processes, the connection between phonemic awareness and a child's ability to decode has been clearly

established (Ehri, 1998; Lonigan & Shanahan, 2008; Murray, Stahl, & Ivey, 1996; Stahl, 2001).

A child's working knowledge of alphabet letters is a strong predictor of short- and long-term reading success. However, the influence of letter knowledge on later reading is not necessarily about knowing the letter names, per se, but rather the learning of letter names mediates the ability to remember the sounds associated with the letters (Ehri, 1979). There is a reciprocal relationship between these skills: a child's ability to detect and manipulate phonemes is associated with higher levels of letter knowledge and letter knowledge plays an influential role in the development of phonemic awareness. For example, a child who knows the letter "b" is more likely to associate it with the sound /b/. This becomes a phonetic cue as a student develops and is more able to correctly identify the sound /b/ within words, he/she later can recognize initial and final consonant sounds, vowel sounds, and begin blending these sounds together. Murray et al. (1996) found that a student's ability to segment initial phoneme sounds, as demonstrated in the previous example, means a child understands that phoneme sounds are consistent across various combinations. Consequently, letter knowledge reflects a more substantial underlying understanding and familiarity with other literacy related skills such as linguistic units and concepts of print (Murray et al., 1996; Stahl, 2001). As students begin to quickly and automatically recognize letter (and phonemic) groupings, they add to their personal reading lexicons—adding to the many words that they may already be familiar with (e.g., environmental print). This automaticity is reflected in reading fluency (Stahl, 2001).

Adams (2001) argued that reading comprehension follows systematic, explicit

instruction in phonemic awareness and concluded that systematic and explicit instruction in letter-sound correspondences has a positive impact on a student's overall reading development, including reading comprehension and word recognition, regardless of the age, grade-level, or SES of the student. Adams reiterated that a lack of skill in the areas of alphabetic knowledge and phonemic awareness is a leading at-risk predictor for reading failure for students with reading disabilities. Referring to the report from the Committee on the Prevention of Reading Difficulties in Young Children the National Research Council (1998, as cited in Adams, 2001) states:

> Objective, empirical research has proven over and over, using a wide array of methods and instrumentation, that given an alphabetic script, the skillful reader's ability to read with fluency and reflective comprehension depends, integrally and incontrovertibly, on deep, detailed, and ready working knowledge of the spellings and spelling-sound correspondences of the words on the page. (p. 73)

Students who understand the governing alphabetic principle of reading can then more automatically and rapidly decode words. Without the development of the essential phonemic awareness and linguistic unit discrimination (e.g., words, segments, phonemes), improvement in reading achievement cannot be reached (Burgess, 2006; Longian, 2006, Stahl, 2001; Stockard & Engelmann, 2010).

Phonics

As previously noted, in order for children to learn to read, they must understand the relationship between the written symbols (graphemes) and their associated sounds (phonemes), then remember the many letter patterns and sequences that represent various speech patterns (Moats, 2000). To read independently and to construct meaning successfully, beginning readers need to identify words automatically and have an

effective strategy for decoding unknown words (Bos & Vaughn, 2002; NRP, 2000; Snow et al., 1998). This involves awareness of the 44 English speech sounds, and the more than 100 letter combinations that represent them (Blevins, 1998; Bos & Vaughn, 2002). These skills in combination are commonly referred to as "phonics."

Over the years, phonics instruction has been examined and evaluated for its influence on reading success. The seminal publication, *Learning to Read: The Great Debate* by Chall (1967), was written to outline a great body of reading research. Chall's comprehensive review of beginning reading instruction to the mid-1960s found that early and systematic instruction in phonics leads to better achievement in reading over later and less-systematic phonics instruction. Chall's basic findings have been validated repeatedly since that publication (e.g., Adams, 1990; Armbruster, Lehr, & Osborn, 2010; Balmuth, 1982; Bond & Dykstra, 1967; Brady, 2011; Chall, 1990; Dykstra, 1968; Ehri, 2005 Ehri, Nunes, Stahl, & Willows, 2001; NRP, 2000; Shanahan, 2005; Nicholson &Tunmer, 2011).

Effective phonics instruction focuses on understanding the letter-sound relationships, and explicitly teaches the alphabetic principle of English orthography. English is an opaque language with complex phoneme-grapheme correspondences. The complex relationships between symbol patterns and meaning is what makes English spelling different from many other languages (Vaughn & Linan-Thompson, 2004). Instruction that teaches these complexities explicitly, teaching students the complex relationship between English phoneme-grapheme correspondences in various forms is referred to as "phonics." Phonics instruction has been described by Williams, Phillips-

Birdsong, Hufnagel, Hungler, and Lundstrom (2009) as:

> [The] use [of] a variety of hands-on activities, often called *word work*, to help students actively explore these layers of information. When studying the alphabetic layer, students examine the relationship between letters and sounds. They learn to match single letters and pairs of letters (e.g., ch) to specific sounds and, in doing so, to create words. When students study the pattern layer, they look beyond single or paired letter-sounds to search for larger patterns that guide the grouping of letters (e.g., CVCe). (pp. 570-571)

The effectiveness of phonics instruction was examined in a meta-analysis by Hattie (2009). He conducted a study of 14 meta-analyses, involving 12,000 students. Hattie found that systematic phonics instruction had a mean effect size of $d = .60$, larger than the findings of the NRP (NRP, 2000). The Panel's findings indicated that the overall mean effect size at the end of systematic phonics instruction was .41; the effectiveness of this instruction was greatest among typically developing kindergarten and first graders ($d = .55$). Second through sixth-grade students only marginally benefitted from systematic phonics instruction ($d = .27$). Importantly, systematic phonics instruction was of more benefit to at-risk students ($d = .74$) than typically achieving first graders ($d = .48$); and systematic phonics instruction had larger effects for children from low socioeconomic status (SES) backgrounds ($d = .66$).

Research repeatedly suggests that to read independently and construct meaning successfully, beginning readers need to identify words automatically and have an effective strategy for decoding unknown words (Bos & Vaughn, 2002; NRP, 2000; Snow et al., 1998). Thus, understanding the letter-sound relationships and the alphabetic principle of English orthography is foundational to reading efficacy.

Reading Fluency

Phillips and Torgensen (2006) explain that a student's reading fluency and decoding ability impact reading comprehension. In order for a child to comprehend what is being read, he/she must be able to quickly and accurately recognize the words. If a student encounters too many unknown words while reading, fluency is slowed because the reader must apply decoding strategies before moving forward with reading the remainder of the text. Moreover, Stockard and Engelmann (2010) state that students need to decode words quickly and accurately in order to reduce cognitive strain, and to devote mental faculties to comprehension.

In The Nation's Report Card (Perie, Grigg, & Donahue, 2005) noted that oral reading fluency is an important link between word decoding and passage comprehension, and an accurate predictor of overall reading competency. In their 2006 study, Rasinski and Hoffman attempted to measure reading fluency by how accurately and efficiently students were able to decode words in text. Rasinski and Hoffman state that readers have a finite amount of cognitive resources to apply during the reading process; these resources must be utilized to simultaneously decode words and comprehend meaning. Though fluency may deal with the most surface layer of text, negotiating this layer effortlessly (automatically) and accurately will positively affect comprehension (Rasinski & Hoffman, 2006).

As noted earlier, Adams (2001) explained that readers have a limited amount of cognitive resources that can be expended during the reading process—the reader must initially focus on one task at a time. In order for a reader to understand what is being

read, focus must be maintained on the task of constructing meaning and monitoring understanding. Adams states that this can only occur when a reader automatically recognizes words on a page. When words are encountered during reading that the reader does not automatically recognize, comprehension stops and the reader must expend cognitive energy utilizing word recognition and strategies to understand the new word before moving on in the reading process (Adams, 2001).

Reading Comprehension

As stated by the NRP (2000), reading comprehension is the "essence of reading" (p. 13). The NRP expanded the definition of reading comprehension to be the act of constructing meaning from interactions with text through intentional thinking about text and its meaning. Juel (2006) studied the impact of students' early school experience on their ability to read. Her findings suggest that background knowledge and vocabulary are central to reading comprehension. She notes that in order for students to learn to understand what they are reading, teachers must focus on word recognition strategies and vocabulary. Juel also emphasizes the importance of word recognition because it facilitates wide reading, resulting in vocabulary growth.

Stahl (2006) reiterated the work of Juel (2006) and Adams (2001). In his research on understanding the shifts in reading instruction, Stahl discusses reading strategies, and when best to apply them. Stahl advocates reading instruction that includes an explicit and systematic phonics program, which, he suggests, leads to more automatic word recognition. Additionally, Stahl recommends the use of repeated reading as a method for improving word recognition. Students need to practice recognizing words they have

learned in order to make the process more fluent and automatic (Stahl, 2006).

The National Research Council (1998) identified the lower grades as a time when students learn to read, and the upper grades more focused on reading to learn. This assumes that students have learned how to read and can devote cognitive efforts to understanding what is read. Students in second and third grades are said to be transitioning from learning to read, to reading to learn.

In sum, research repeatedly indicates that mastery of phonemic awareness and language skills are prerequisite for mastery of the alphabetic principle (Burgess, 2006; Longian, 2006, Stahl, 2001; Stockard & Engelmann, 2010). Understanding of the letter-sound relationships and the alphabetic principle of English orthography are foundational to decoding effectively (Vaughn & Linan-Thompson, 2004). In order for students to have sufficient cognitive resources available for comprehension, the above-mentioned foundational skills must be well established.

The Great Debate and the National Reading Panel

Reading instruction has always been fraught with epistemological debate that, at times, deeply divided both researchers and instructional programs. The Great Debate (as described by Chall, 1967) revolved around whether beginning reading development was a top-down (whole language) or bottom-up (code-breaking) process. Did children learn to read better via holistic experiences with books and by using their oral language to platform their discovery of text? Or, did most youngsters need a clear explanation of the written code to master the letter-sound relationships in known and new words? To

address the controversy from a scientific perspective, the U. S. government authorized a review of extant research related to reading. Although limited in its scope and very controversial, the NRP published a report (NRP, 2000), which scrutinized the scientific rigor of published studies and refocused reading instruction in American schools. The Panel's charge was to describe narrowly defined research findings and provide educators with quality research-based recommendations so they could proceed with what was best for children, especially those beginning the *learning-to-read* process (Shanahan, 2005).

The NRP (2000) was composed of 14 renowned researchers, selected from a list of more than 300 nominees offered by organizations and agencies such as the International Reading Association and the National Reading Conference. These panelists were prohibited from having financial ties to educational vendors or commercial publishers, in order to be free from outside influences. The Panel did not offer opinions about research findings, but drew conclusions only from the research findings that evidenced instructional elements that cause higher achievement. The Panel reviewed over 100,000 research studies, but included only those studies that utilized a rigorous experimental design. Conclusions were drawn only when research findings evidenced a high degree of certainty. These conclusions were based on a synthesis of a large number of studies, with results that repeated across independent and trustworthy investigations (NRP, 2000; Shanahan, 2005).

The NRP (2000) found 391 studies that met the criteria for inclusion in the meta-analysis. Of these studies, 52 focused on phonemic awareness, 38 on phonics, 51 on oral reading fluency, 45 on vocabulary instruction, and 205 on comprehension strategies. The

Panel ultimately determined that to learn to read effectively, young children should develop three major skillsets: (1) alphabetics (the ability to break apart and manipulate sounds-phonemic awareness, and to understand that sounds are represented by letters that are blended to form words—phonics), (2) oral reading fluency (the ability to efficiently and accurately recognize words and read with prosody), and (3) comprehension (students' understanding of a written passage).

Reading and Academic Success

Many school-aged students in this country are still unable to fluently read grade-level texts (Begeny & Silber, 2006; NAEP, 2015; NRP, 2000). Reading difficulties characterize a major risk factor for long-term academic failure, as well as difficulties later in life (Jarvis, 2016; Lingo, Slaton, & Jolivette, 2006). Juel (2006), in their study on the impact of early school experience on students' initial reading abilities, found that 37% of fourth-grade students nationwide had reading proficiency scores below grade level; the percentage of students from low-SES homes was even greater (56%). Juel noted that students who were below benchmark by the end of first grade typically remained behind their peers throughout their school careers. Juel posited that educators "have a window in school to help children succeed at learning to read before their self-esteem is seriously eroded or they stop even trying to learn, feeling they simply can't do it" (p. 416).

Stockard and Engelmann (2010) point to a number of research studies indicating that students who are performing below grade-level benchmark at the end of first-grade suffer greater academic, social, and emotional problems than their peers. Their research

findings promote the idea that attention to reading achievement in first grade is of paramount importance, substantiating the NRP's (2000) findings, and further suggesting that intervention before students fall significantly behind their peers can keep students on track for greater academic success.

In contemporary American schools, reading instruction is focused on heavily in the younger grades, with the intention of developing the requisite reading skills necessary for comprehension in preparation for content-focused instruction in older grades (i.e., as students advance through schooling, they must read to learn increasingly more complex texts). Teachers and administrators are ever searching for programs that will reach as many students as possible, in the most effective and efficient manner possible, preparing them for future academic success. Though there are always children who learn to read without systematic and explicit instruction (Durkin & Allington, 2004)—there exists a robust body of evidence suggesting that many, if not most, children benefit from systematic and explicit instruction that follows established language development traits (Coley, 2015; Justice et al., 2009; Snow et al., 1998).

Instructional Methods

The best method for instructing children in early literacy is often dependent upon population to be instructed (Blamey & Beauchat, 2016). As previously noted, students who begin their educational journeys with a solid foundation of early literacy skills are better able to appropriately apply these skills, and thus improve academic outcomes (Adams, 1990). The stakes increase dramatically for students from low-SES

backgrounds, who lack preparation for schooling (Blamey & Beauchat, 2016; Dickinson, McCabe, & Essex, 2006; Hart & Risley, 1995; Neuman, 2006). Selection of appropriate instructional strategies to make up ground for these students, as well as extend and deepen the knowledge of adequately prepared students is the goal of all educational institutions.

Basal Readers

Basal readers are "commercially prepared and marketed resource materials that provide classroom reading instruction in elementary and middle school" (Cooter & Reutzel, 2005, p. 162). Basal readers typically include a grade-level student textbook containing selected readings, student workbooks, as well as supplemental reading materials (e.g., audio recordings, big books or flip charts, etc.). Basal readers are considered "comprehensive" in nature, often structuring lessons targeting vocabulary, comprehension, word analysis, and spelling. They are geared toward whole-class instruction, and on occasion, contain elements of re-teaching of struggling readers (Tompkins, 2001).

Basal readers have a long history in the American educational system, beginning with the hornbook and the *New England Primer*—both stemming from the religious freedom movement of the American colonists prevalent at the founding of the U.S. (Cooter & Reutzel, 2005). The "Dick and Jane" readers emerged in the 1940s, and are regarded by many as the quintessential readers, emphasizing the whole word (or *look-say*) method (Shermer, 2003). Today, the basal reader is more comprehensive with reading passages that are systematically structured from simple sentences and storylines

to complex plots with several stories strung together with common themes—an all-inclusive set of instructional materials (Goodman, Shannon, Freeman, & Murphy, 1988). They are designed to be developmental and continuous programs that provide sequential and systematic learning of reading skills.

Limitations of some basal readers, as identified by Cooter and Reutzel (2005), include little skill instruction relating to decoding text or orthographic study, and failing to include reading selections of interest to the student, focusing on reading that is at the instructional level of the student, resulting in low engagement and, ultimately, low comprehension.

Direct Instruction

Direct instruction is an instructional methodology in which teachers break down instruction into small steps so that students are better able to acquire a complicated skill. Adams (2001) describes it as instruction that helps students to focus on the learning relationships that matter most. Ritchey (2011), in her work on teaching reading to students with learning disabilities, defines direct instruction as being teacher-directed with teacher explanations and modeling as well as prompting of students and providing immediate corrective feedback.

According to Ritchey (2011), direct instruction needs to be clear and intensive. Intensity refers to the number of minutes per day of instruction, instructional sessions per week, student-teacher interactions, and the amount of student interactions with text and their use of reading skills. Ritchey further recommends that struggling students need more direct instruction for longer periods of time, and more frequently. She found in her

2011 study that struggling readers are best instructed in small groups, which increase student opportunities to respond and interact with the teacher. In addition, instruction needs to be designed in such a way that students have frequent opportunities to respond (utilizing strategies, such as choral responses, instead of single-student response), engaging students in more practice opportunities.

Popular Phonics Programs

The Saxon Phonics Program is a systematic and explicit phonics program that Scruggs and Mastropieri (2000) describe as an instructional tool that begins with auditory discrimination and sound blending activities and extends to providing meaning to words based on letter combination. It is difficult to provide evidence of the effectiveness of the program on decoding, oral reading fluency, and overall reading skill due to the paucity of research supporting its use. However, the Saxon Publishers Research Department has compiled test-result data and testimonials from schools across the country.

Simmons (1996), the designer of the Saxon Phonics program, claims that for most students to develop a firm foundation in phonics and to become successful readers and spellers' programs should provide a complete curriculum that carefully moves students from letter recognition to letter/sound associations to the reading of increasingly difficult text based on the skills learned cumulatively. Simmons also claims that the Saxon Phonics program includes a structured and predictable sequence of review so that skills learned are maintained and generalized. Students study and apply vowel, spelling, and syllable-division rules to help them decode unfamiliar words. Learning these rules, students no longer need to guess at spelling and sound patterns. Additionally, students

learn to code words with diacritical markings, enabling them to decode unfamiliar words and familiarizing them with dictionary pronunciation and phonetic coding systems.

Simmons' (1996) claims are purportedly supported by Saxon Phonics Results (Simmons & Calvert, 1996), published by Saxon Publishers to record the effectiveness of the program. The report does not describe in detail the effectiveness of the program; however, it does claim that the program positively influences reading achievement over a 3-year period. Results from the *Iowa Test of Basic Skills* of first-grade students were gathered prior to Saxon instruction, as well as two years following Saxon instruction. Simmons concluded that reading scores did increase gradually and steadily over the course of two years, though no assessment, test, or subtests data are noted. No independent research has been done on the Saxon Phonics program, leaving Simmons' claim unsubstantiated. A similar product, Jolly Phonics, utilized extensively in the UK, has been the focus of considerable independent research.

Jolly Phonics is a packaged program developed for direct instruction utilizing structured and systematic organization of phonics skills taught in an explicit manner. One study by Ekpo, Udosen, Afangideh, ekukinam, and Ikorok (2007), using a pre-/posttest experimental design, sought to determine the effectiveness of Jolly Phonics as a fast-track strategy in enhancing first-grade students' reading skills in Nigeria. The data collected were analyzed via means, independent *t* tests, ANCOVA and the Burt Reading Test (Ekpo et al., 2007). The results showed a significant difference between posttest mean scores of the treatment and control groups ($p = 0.001$). There was also a significant difference between the posttest mean score of the treatment group when stratified into

urban and rural schools ($t = 3.40$, $df = 82$), suggesting that Jolly Phonics was more effective for the urban/high-poverty students.

Ekpo et al.'s (2007) results are consistent with findings from other studies. In their comparative study between the Jolly Phonics (JP) program and the Teaching Handwriting, Reading and Spelling Skills (THRASS) program in which first-grade students at different schools were taught one program or the other, Callinan and van der Zee (2010) determined that both word and nonsense word, as well as short-term memory, skills improved with direct instruction in synthetic phonics. Jolly Phonics is a direct instruction program; THRASS is also an explicit phonics-based program that focuses on the instruction of spelling patterns in English though the utilization of pictures. The researchers used a repeated measures ANOVA and post-hoc t-test to determine sustained improvement in the JP schools over the THRASS schools (1.3 points more). More specifically, the JP schools had greater gains in word and nonsense-word reading tasks, as well as short-term memory skills, although no significance or P-values are reported.

Souther (2015) completed a study on the cumulative effectiveness of the phonics-based Letterland program for students in kindergarten through second grade. The study utilized a two-phased, mixed-methods approach wherein teachers were surveyed and observed during instruction. Then, student achievement data were examined to determine the effectiveness of a phonics-based instructional program. The analyses compared students who began using the phonics-based program in 2010 ($N = 258$) against a control group ($N = 286$) to determine if significant growth resulted from implementation of this program.

The results indicated that a significant difference existed between the treatment and control groups, suggesting Letterland (the phonics-based program) is an effective tool to teach students foundational reading skills. Souther (2015) used $\alpha = .05$ with one degree of freedom and a chi-critical value of 3.841, and a chi-statistic of 51.9528. Because the chi-statistic was greater than the chi-critical, the analysis suggests a significant difference in the reading development of students attending schools that began implementing Letterland. These results indicate that students also made significant growth in their comprehension of what they read. Souther surmised that students instructed using the phonics-based Letterland program made significant growth in their reading over those who were not instructed using the program. Souther's finding supports the idea that phonemic awareness and systematic phonics instruction constitute a critical component to the success of early reading.

Overall, these studies would seem to support the claims made by Simmons (1996) that similar programs, such as Jolly Phonics and Letterland, improve students' early reading skills in general, but also that systematic phonics programs work better for at-risk/low skill students who may be less prepared for the learning-to-read process. However, these studies also indicate that highly touted packaged phonics programs may be under-researched and although used globally, there is minimal research of their effectiveness as mainstreamed programs in the U.S.

Conclusion

Juel and Minden-Cupp (1999) state that students are expected to acquire a

significant number of words during their academic careers. Utilizing instruction that helps students to do this is of upmost importance. As stated in the NRP Report (NRP, 2000), first grade is the keystone grade for reading instruction—if students have not acquired foundational reading skills then, it is more difficult for them to make grade-level progress throughout elementary school. Systematic and explicit phonics instruction has been shown to be effective for teaching these foundational reading skills in early grades—kindergarten and first-grade—and marginally effective in grades two through six (Hattie, 2009). Additionally, systematic and explicit phonics instruction has been shown most effective for at-risk student (low-skill readers) and for low SES students (NRP, 2000). Examining the body of related research suggests that there are two major questions that are continually asked in regards to phonics instruction: (1) is phonics instruction more effective under some circumstances and conditions, and (2) is phonics more effective for certain students? It is also important to note that none of the studies reviewed compared supplemental phonics instruction paired with a basal reader to instruction from a basal reader alone.

Effective Circumstances and Conditions.

The work of Brady (2011) and Tunmer and Arrow (2013) tease out some of the conditions and circumstances of when "good" systematic phonics instruction is most effective. Tunmer and Arrow found clear indications that systematic phonics instruction includes intentional direct instruction in orthographic patterns and word identification strategies. Phonics instruction, when delivered directly was more effective than teaching word-analysis skills incidentally for beginning and struggling readers. Brady also

concluded that outcomes increase significantly when phonics instruction is combined with dynamic opportunities for children to practice and receive feedback following instruction as they apply their word analysis skills while actively engaged in the process of reading.

Price (2015) completed a dissertation examining the effects of systematic phonics instruction on kindergarten reading readiness scores ($N = 120$). The quasi-experimental study involved a treatment of 10 minutes of explicit phonics instruction daily, four days a week, for 15 weeks. The control group received instruction through an integrated curricular approach whereby students were taught using a basal program with only incidental phonics instruction. Upon completion of the program, students were evaluated using the STAR Early Literacy Test (a diagnostic assessment of early literacy skills (Renaissance.com, 2017), to determine their mastery of early literacy skills. The treatment group scored higher than the control group ($t = 2.07$, $\alpha = .05$ on a two-tailed t test), suggesting that participating in an explicit phonics program can significantly increase kindergarten reading readiness scores (Price, 2015).

The findings of Price (2015) support the use of systematic phonics instruction in the classroom. However, Price looked at systematic phonics instruction developed by teachers. As previously noted, others have looked at phonics instruction delivered through a prepackaged, supplemental programs. These results beg the question: is there a difference on student reading achievement between typical basal/incidental phonics instruction and structured and systematic supplemental phonics instruction delivered through a packaged program?

Effective for Whom?

To determine for whom is explicit phonics instruction effective, Magnin (2011) conducted a study comparing the progress of typical first-grade students being taught using the code-based intervention program Phonics for Reading (Curriculum Associates, 2011) with students who received instruction through Guided Reading (Fountas & Pinnel, 1996). Progress was measured using several subtests of the DIBELS assessment. The study utilized a pre/posttest experimental design where students were administered a pretest, grouped by reading score level, and assigned instructional method (low readers, n = 4, received Phonics for Reading and medium readers, $n = 4$, received guided reading). Following four months of instruction, students were again assessed using a posttest to determine progress toward proficiency and to see if one instructional practice was making more progress over the other.

The results of Magnin's (2011) study found that students who received Phonics for Reading improved an average of 20 points ($M = 20.50$, $SD = 6.608$) as measured by the Nonsense Word Fluency subtest of the DIBELS Assessment. Students who received instruction through Guided Reading improved an average of 11 points ($M = 10.75$, $SD = 6.602$). Additionally, Magnin found that students who received Phonics for Reading achieved higher DIBELS scores ($M = 31.00$, $SD = 8.165$) over those who received Guided Reading ($M = 29.25$, $SD = 13.623$). These results indicate that the systematic and structured code-based (phonics) program benefited lower-achieving students and early readers, helping them make progress toward reading proficiency in later years—echoing findings of both Hattie (2009) and the NRP (2000) regarding systematic phonics

instruction for low-skill readers and for low SES populations. Because of the small sample size, the ability to generalize the findings is quite limited, but Magnin's finding suggests that a larger study examining whether a supplemental phonics program supports lower-achieving or at-risk students is warranted.

These studies suggest that there may be various ways to provide the recommended phonics-based instruction for beginning readers with possible differential effects. The challenge for classroom teachers and school administrators, as noted by Joshi et al. (2009), is to find the best method to teach phonics that reaches as many students as possible, in the most efficient manner possible. Research is clear in this respect: phonics instruction should be structured systematically with a gradual change of focus from phonics in younger grades to comprehension in older grades, and from beginning of the year to the end in order to reach as many students as possible, or to achieve maximum benefit. It is also clear that systematic and explicit phonics instruction is of more benefit for some students than for others—that is low-skill or low SES (Hattie, 2009).

Students are expected to acquire a vast number of skills and mastery of numerous strategies by the end of first grade in order to read fluently and comprehend what they read as texts become more difficult in later grades (Allington & McGill-Franzen, 2008; Chatterji, 2006). These skills are essential for students to access grade-level content conveyed through text. Research repeatedly identifies first grade as the cornerstone grade—when foundational phonics skills (letter-sound correspondences and English orthography) are taught (NRP, 2000). These skills enable rapid and automatic decoding of text, reducing cognitive demand, and freeing up cognitive resources to utilize in

comprehension (Stockard & Engelman, 2010). This study, therefore, explores the effects of explicit and systematic instruction in phonics when paired with basal-reading instruction on indicators of beginning reading achievement (i.e., oral reading fluency, accuracy, and nonsense word fluency) of first-grade students in order to provide additional evidence as to whether systematic and structured instruction in phonics significantly improves first-grade students' reading scores.

CHAPTER III

METHODOLOGY

The purpose of this quantitative study was to examine the possible effects of a supplemental phonics program—Saxon Phonics—on early reading achievement of first-grade students in one mountain west semi-rural school district. Specifically, this study sought to determine:

1. Do students taught using a traditional basal program and students taught using both the traditional basal program and a supplemental phonics program (control vs. treatment) differ on selected end-of-year reading achievement scores (i.e., portions of the DIBELS Next assessment—ORF, ACC, and NWF?)

2. Do any possible interactions among selected variables (i.e., instructional program, gender, and beginning-of-year reading level) exist related to performance differences on end-of-year reading achievement scores among students receiving reading instruction with or without a supplemental phonics program?

The research design was causal-comparative and ex post facto. Upon approval from the Institutional Review Board, archived DIBELS data were collected for the 2014-15 and 2015-16 school years, formatted, and entered into R studio. These data were analyzed using mixed effects model (or multi-level regression) to determine any significant differences and interactions among variables.

Setting

The study took place in one semi-rural school district servicing a pre-kindergarten through post-high-school—age students, with a total enrollment of 11,341 students as of October 1, 2015 (USOE, 2015). The total student population encompasses the following

demographics: 5,792 (51.07%) males, 5,549 (49.92%) females, with a gender ratio of 1.04:1; 71 (.006%) Native American, 72 (.006%) Asian, 70 (.006%) Black, 1,173 (10.34%) Hispanic/Latino, 49 (.004%) Pacific Islander, 110 (.009%) Mixed Race, and 9,796 (86.3%) White/Caucasian (USOE, 2015). Forty-one percent of the district population is considered economically disadvantaged (USOE, 2015), with 12% being classified as disabled and accessing special education services (including all pre-kindergarten, as well as those receiving speech and language services), and 3% identified as English Language Learners (USOE, 2015).

The instructional model used within the research sites during treatment consisted of 180 minutes of English language arts instruction daily for first-grade students. Of the 180 minutes of daily instruction, the current instructional model entail 65 minutes of word study (10 minutes-phonemic awareness, 55 minutes—phonics), 45 minutes of small skill-based reading groups (15-20 minutes of targeted intervention or extension and 25-30 minutes guided reading), 10 minutes of whole group fluency instruction, 30 minutes of whole group writing instruction, 10 minutes of whole group vocabulary instruction, and 20 minutes of whole group comprehension instruction (see Figure 1). This instructional model was introduced during the 2015-16 school year.

Prior to 2015-16, the instructional model used at the research site included 180 minutes of English language arts instruction, but was much less prescribed. Before 2015-16 phonemic awareness and phonics instruction was taught incidentally through mini-lesson woven through the basal reading program and directly using teacher created materials and curricula when teachers felt additional support was warranted.

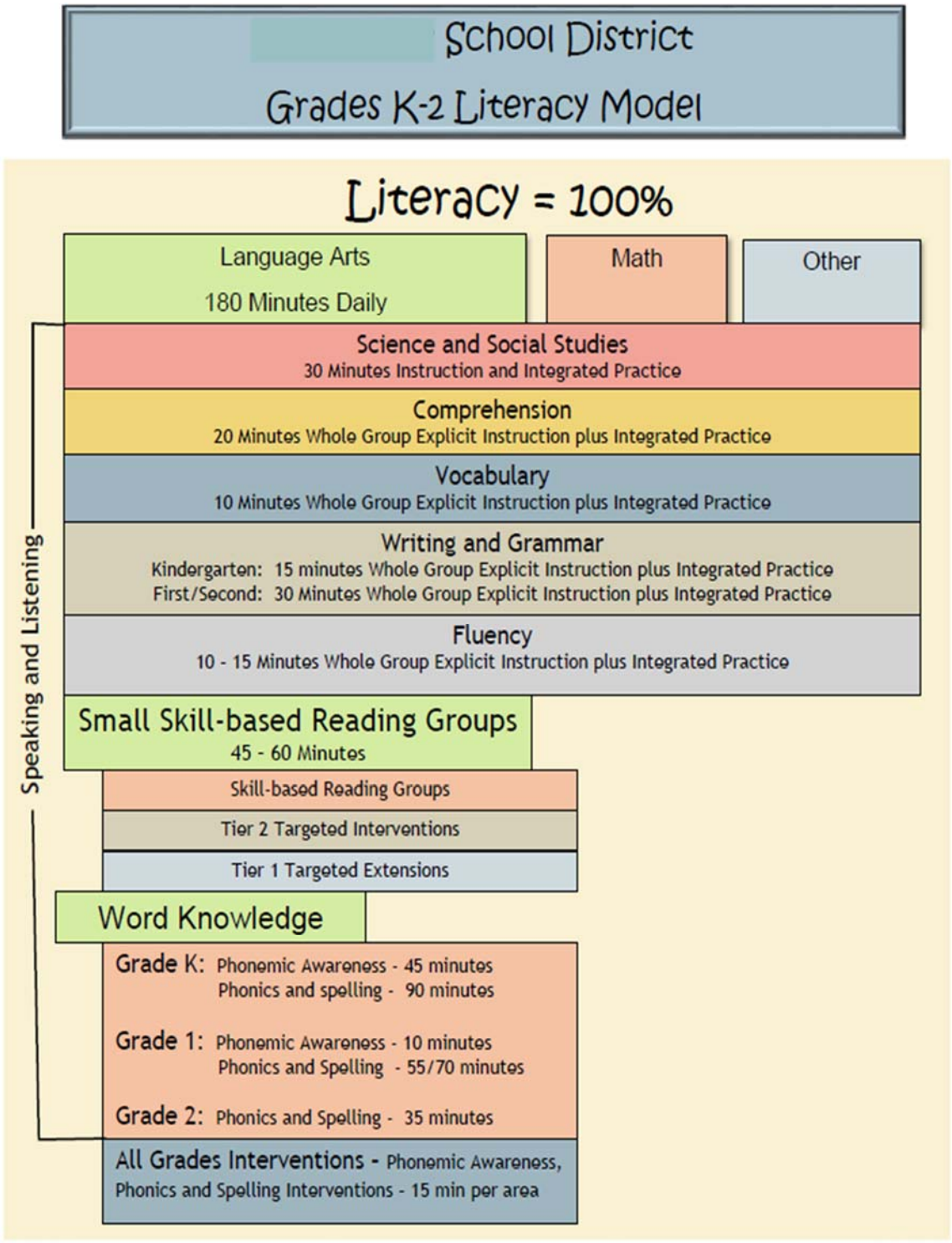

Figure 1. District literacy model for grades kindergarten through fifth-grade illustrating the estimated time per day spent on literacy instruction.

A formalized literacy instructional model was created with the intent of unifying district instructional practices and to provide each student with the same quality of instruction, regardless of the school or classroom.

Sample

The participants of this study included the first-grade students who were assessed using the DIBELS Next assessment in Box Elder School District and who were enrolled in, at least, 160 days of instruction during both the 2014-15 and 2015-2016 school years. The 1,657 participants were grouped by reading level as determined by ORF scores from their kindergarten end-of-year administration. 859 were from the 2014-15 school year and were taught using a traditional/basal reading approach, with no supplemental phonics programing. The other 798 students attended first grade for the 2015-2016 school year and were instructed using a literacy model that includes the traditional/basal reading instruction, as well as instruction using a supplemental phonics program. The participants represent nearly all first-grade students in the school district and Box Elder County; however, one school adopted the Saxon Phonics curriculum for use during the 2014-15 school year—those students are not included in the sample. Given the large sample size, however, and the wide distribution of students and schools, the participants were considered a complete sample.

Design

As stated by Kerlinger and Lee (2000), the cardinal rule of research planning is to

have research questions that dictate a research design. Because this study sought to determine the extent to which a structured phonics curriculum has an effect on oral reading fluency and decoding skills of first-grade students, an ex post facto design was appropriate to answer the proposed research questions and to test the identified hypotheses because the variables of interest have already occurred (Newman, Newman, Brown, & McNeely, 2006). Ex post facto research, by definition investigates "the world as it naturally occurs" and explores phenomena that have since occurred (Johnson & Christensen, 2008, p. 257).

There are three types of ex post facto research design (Newman et al., 2006). The first design, considered the weakest of the three, uses no hypothesis and simply explores a phenomenon. The second design—research with hypotheses—is considered more vigorous. However, the third design, ex post facto research utilizing hypotheses and tests for alternative hypotheses, is considered to be the most robust "in terms of internal validity"; it is considered by some to be a heartier design than certain types of quasi experimental research (p. 101).

Ary, Jacobs, Sorensen, and Walker (2010) noted that an ex post facto research design is most applicable and useful when randomization and manipulation of the dependent and independent variables are not possible. While both experimental and ex post fact research often investigates relationships between variables and test hypotheses, ex post facto research does not provide sufficient or adequate safeguards; as a result, inference of causal relationships is more difficult (Ary et al., 2010). Despite these limitations, and due to its pragmatic nature, ex post facto research is invaluable in both

educational as well as social science practice as a means organically explore phenomena as it typically happens.

Validity

Internal Validity and Alternate Hypotheses

In any research, it is important to note any factors that threaten the internal validity of the study (Campbell & Stanly, 1963). If a factor or factors other than the independent variable affects the dependent variables, it must be noted and, when possible, controlled for (Johnson & Christensen, 2008). In the case of ex post facto research, where independent variables are not manipulated, researchers must take particular care in drawing conclusions about causation because other factors may be at play that may not be controlled for (Ary et al., 2010; Best & Kahn, 2006; Johnson & Christensen, 2008; Kerlinger & Lee, 2000). The post hoc fallacy (drawing the conclusion that causation exists because evidence of a relationship is found) a is of most concern to an ex post facto design (Ary et al., 2010). A valid conclusion of causation meets the following criteria: (1) a statistical relationship between X and Y can be established, (2) X preceded Y in time, and (3) other factors did not determine Y (Ary et al., 2010, p. 333; Campbell & Stanly, 1963). The last criterion is the most difficult for ex post facto research designs (Johnson & Christensen, 2008).

Kerlinger and Lee (2000), as well as Newman et al. (2006), suggest that a researcher can improve the credibility of an ex post facto research design by considering a number of alternate hypotheses or possible interaction. The greater the number of

alternative hypotheses and/or possible extraneous interactions a researcher can eliminate, the more internally valid and powerful a study can be considered (Newman et al, 2006, p.101). Keeping these conditions in mind, this study design will attempt to control for the effects of undetermined possible variables by providing for an alternative conclusion. For this reason, reading level was included as independent variable. A possible extraneous interaction specific to this study may include the difference in scores between the control and treatment groups is related to the difference in the reading level of the students before instruction.

The inclusion of this possible interaction acted as a control to extraneous influences that may account for any effect on reading scores. Differences between groups such as a larger number of "high" readers may have accounted for the differences in scores rather than the actual inclusion of a supplemental phonics program. Additionally, the inclusion of this possible interaction requires grouping of data, which may reveal that Saxon Phonics is more appropriate types of readers.

In addition to the above stated controls, the primary means of data analysis was a mixed effects model, which accounts for extraneous variables and interactions.

External Validity

The major factor that significantly affects the ability of a researcher to generalize is sample size and selection (Campbell & Stanly, 1963). Specifically, in regard to the limits of ex post facto research design external validity, Best and Kahn (2006) state that "subjects cannot be randomly, or otherwise, assigned to treatment groups" (p. 145). The sample examined during this study will consist of first-grade students in on mountain

west, semi-rural, school district. Because there was no random assignment to treatment or control groups, any conclusions regarding the entirety of the student population of the school will be cautious. However, given that the sample comprises the complete population of first-grade students, and that the make-up of that population closely resembles that of the entire county which the school district serves—the results may be generalized to students of first-grade age throughout the county. Care will be taken to limit recommendations for practice of first-grade students in other localities in the state and beyond. It would still be appropriate to make recommendations for future research and implications for the broader student population within the research site and perhaps for similar schools and programs.

Ecological Validity

The great "trade-off" in research is that as a researcher controls for one type of validity, the threats to the other increase; that is, when one implements controls aimed to increase the degree of internal validity, these same measures limit the ability to generalize the findings (Ledford, Hall, Conder, & Lane, 2016). Moreover, the interaction between variable within phenomena are "…often multiple and complex rather than single and simple" (Best & Kahn, 2006, p. 145). Many researchers and practitioners in the social sciences and education have begun calling for research that has a high degree of *feasibility of implementation* (Brooks & Baumeister, 1977; Gast, 2014), meaning the research should closely resemble "real-world" conditions or be implemented in typical contexts (Machalicek, O'Reilly, Beretvas, Sigafoos, & Lancioni, 2007). The concept of using "real-world" or "typical contexts" is referred to as ecological validity (Ledford et

al., 2016). The use of ex post facto data, collected in a practicing school, is a "real-world" setting for use and evaluation of phonics curricula—the very definition of a typical school context. The collection of data in this manner provided for a high degree of ecological validity, and gives weight to the findings of this study, and may apply beyond the context of the schools and district of study.

As stated in Chapter I, the research site adopted the use of Saxon Phonics as component of their literacy instruction model. Thus, this study only measured the possible effects of the use of a specific supplemental phonics program and any conclusion will pertain to its particular use.

Measurement and Instrumentation

DIBELS Next (University of Oregon, 2015) is a standardized, screening assessment based on frequent monitoring of a set of six indicators of the essential skills a student must master in order to become a proficient reader. The early literacy indicators for first grade include letter naming fluency, phoneme segmentation, nonsense word fluency, oral reading fluency, accuracy, and retell (University of Oregon, 2015).

The dependent variables of this study were individual measures within the DIBELS Next assessment. The four measures that measure a student's ability to demonstrate adequate phonics knowledge and apply it generally, are phonemic segmentation fluency (PSF), nonsense word fluency (NWF), oral reading fluency (ORF), and accuracy (ACC). Because PSF is only administered once during first-grade, and only speaks to a student's ability to aurally discriminate phonemes, it was not included as a

dependent variable.

Additionally, although DIBELS Next produces a composite score for each student to help determine the overall reading level of the student, the composite score is determined using scores for LNF and PSF, which are not collected both pre- and post-instruction. Hence the DIBELS Next composite score will not be used in the planned analyses.

The measures that were included as dependent variable were NWF, ORF, and ACC. These measures are described and scored as follows:

1. NWF is administered by presenting the student with various, randomly ordered vowel-consonant (VC) and consonant-vowel-consonant (CVC) nonsense words (e.g. ov, sig, rav) for one minute and counting the number of letter-sounds produced correctly (University of Oregon, 2015).

2. ORF consists of students reading a passage aloud for one minute. The score is reported in correct words per minute (CWPM), and is calculated by counting the total words read and subtracting errors (University of Oregon, 2015).

3. ACC, the companion score of ORF, is calculated as a percentage of the total correct words read (University of Oregon, 2015).

The reliability and validity of the DIBELS Next method of assessment was established by Good et al. (2004) in their study *Technical Adequacy of DIBELS: Results of the Early Childhood Research Institute on Measuring Growth and Development.* The conclusion of the study states: "The DIBELS offers educators brief, valid, reliable and repeated measures to assess students' early literacy skills. Knowing how a child performs on the DIBELS measures in kindergarten and first grade strongly predicts their end of first and second grade reading outcomes" (Good et al., 2004, p. 38). Additionally, the teams' findings corroborate the validity of curriculum-based oral reading fluency measurement (or the use of CWPM) described by Shinn (2002) for the purpose of

identifying proficiency.

Data Collection

Data collection consisted of quantitative information in the form of ex post facto data from the 2014-2015 and 2015-2016 DIBELS Next administration, specifically the NWF, ORF, and accuracy scores from all first-grade students. These data will be exported into a Microsoft Excel spreadsheet prepared, then imported into R studio. Student identification number comprised column one of the spreadsheet. Student names were replaced with a student identification code of 1 – *n*, for the approximate number of participants. Column two contained the pre-/posttest label. Columns three, four, and five contained the dependent variable scores (NWF, ORF, and ACC, respectively). Column six contained the instructional program coded by school year (2014 = basal/traditional programing, 2014 = Saxon Phonics). Column seven contained the beginning-of-year reading level of the student, coded with the number one (1) for low, number two (2) for medium, or number three (3) for high. The names of the column labels were simplified for R recognition.

Data Analysis/Statistical Procedures

Both descriptive and inferential statistical analyses were utilized to analyze the data collected to answer the research question. In order to determine whether to accept or reject the research hypotheses, and to determine the significance of the independent variables (treatment, pre-treatment reading level) effect on the dependent variables

(NWF, ORF, ACC), a mixed effects model (or multi-level regression) was used with the R statistical software package to analyze the growth and performance of first-grade DIBELS from middle-of year to end-of-year. Because measurement made on clusters of related statistical units (selected measures of the DIBELS Next assessment) for each participant, a mixed effects modeling was warranted. Although other analysis methods may have been used, mixed effects modeling was selected because it is dynamic and accounts for both fixed and random effects. For this study, the analysis of selected measures of the end-of-year DIBELS Next assessment was considered to be an interaction between pre-treatment reading ability, the group (treatment vs. control), and other random variables within the classroom. Reading ability was calculated on a continuous scale and used to show patterns during and after analysis. Gender was not included as a covariate.

Summary

This chapter lays out the details regarding the study's methodology and research design. The participants included were first-grade students who were enrolled in at least 160 days of instruction during both the 2014-15 and 2015-2016 school years. This study used a robust ex post facto research design, utilizing hypotheses and tests for alternative interactions and conducted within an actual school context. Data collected included NWF, ORF, and ACC scores of students from the 2014-2015 school year (control group) and the 2015-2016 school year (treatment group). The use of a mixed effects model (or multi-level regression) was utilized to test the research questions: Do students taught

using a traditional basal program and students taught using both the traditional basal program and a supplemental phonics program (control vs. treatment) differ on selected end-of-year reading achievement scores (i.e., portions of the DIBELS Next assessment—ORF, ACC, and NWF?), and Do any possible interactions among selected variables (i.e., instructional programing, gender, and beginning-of-year reading level) exist related to performance differences on end-of-year reading achievement scores among students receiving reading instruction with or without a supplemental phonics program?

CHAPTER IV

RESULTS

The purpose of this quantitative study was to explore the effects of a systematic and structured phonics program when paired with basal reading instruction on three indicators of beginning reading achievement (nonsense word fluency, oral reading fluency, and accuracy) among first-grade students. It is well established that beginning reading should include the development of accurate and fluent oral reading, and nonsense word fluency is a robust measure of decoding skills. Thus, these three literacy skills were measured via pre- and posttests using common school-based assessments. Students' scores were retrieved from archival data held and maintained by the local school district. These scores are from the middle-of-year and end-of-year administrations because not all these tests are administered at the beginning of first grade, and thus any gains represent only four months of instruction. This chapter is organized as follows: a short reorientation of the research questions and hypotheses, summary of the data analyses, results (by group and measures), and a summary of the findings.

Research Questions and Hypotheses

Research Questions

This study sought to answer two research questions: (1) Does the inclusion of a supplemental systematic and structured phonics program affect first-grade students' reading scores? and (2) Do any possible interactions among selected variables (i.e., instructional programing and beginning-of-year reading level) exist related to

performance differences on end-of-year reading achievement scores among first-graders receiving reading instruction with or without a supplemental phonics program? This study utilized a large data set from two populations of beginning readers. This data set represented reading achievement for students in 2014-2015 and 2015-2016 as treatment and control groups. These groups were also subdivided into three additional subgroups (based on beginning-of-year reading levels). Scores by group (based on year) and then by subgroup were used to answer the research questions in a more nuanced manner.

Hypotheses

This study targeted both the general outcomes, as well as any possible interactions between ability level and reading instruction by year that may play a role in reading achievement scores. Several research hypotheses were proposed:

1. There will be a difference in nonsense word fluency (NWF) scores between the control group, who received basal reading program instruction, and the treatment group, who received basal-based instruction paired with the Saxon Phonics program.

2. There will be a difference in oral reading fluency (ORF) scores between the control group, who received basal reading program instruction, and the treatment group, who received basal-based instruction paired with the Saxon Phonics program.

3. There will be a difference in accuracy (ACC) scores between the control group, who received basal reading program instruction, and the treatment group, who received basal-based instruction paired with the Saxon Phonics program.

4. The difference between groups, if any, will vary depending on pre-instructional reading levels of low (1), medium (2), and high (3) among students. These levels were determined by dividing students based on district benchmarks from their beginning-of-year DIBELS composite scores.

Results

The following sections discuss the results in relation to benefit for whole groups as well as identified subgroups, by subtest.

Group Results

Research Question 1 of this study asked whether there would be a difference in posttest outcomes based on the use of a supplemental phonics program. A single-factor analysis of variance (ANOVA) was utilized to identify significant differences between the groups' beginning-of-year reading levels, pretest scores, posttest scores, and gains on the three select measures, where $\alpha = .05$. The ANOVA revealed that there was no significant difference in beginning-of-year reading levels ($p = 0.13$). Table 2 displays the descriptive statistics, including group means, standard deviations, and gains by group (control vs. treatment). The group data indicate no significant difference on any of the explored factors; both groups made measurable gains in reading scores, but the differences in gains on all the selected measures were not significant between the control and treatment groups.

Gains in reading skills were expected as students received literacy instruction and developed their reading skills. Based on the district's established benchmarks and DIBELS benchmark goals (University of Oregon, 2015), first graders are expected to make gains of five words for NWF, 24 words for ORF, and 12 percentage points for Accuracy between the mid- and end-of-year assessment periods. These gains are calculated by subtracting pretest scores from the posttest scores. Given these targeted

Table 2

The Established Means, Standard Deviations, and Gains Scores, by Group for Nonsense Word Fluency, Oral Reading Fluency, and Accuracy

	Control group			Treatment group			
Outcome	%	Mean	*SD*	%	Mean	*SD*	ANOVA *p* values
Nonsense word fluency							
Pretest mean		23.46	14.78		22.85	14.28	0.34
Posttest mean		29.90	14.47		29.96	13.80	0.94
Mean gains		6.47			7.16		0.12
Oral reading fluency							
Pretest mean		41.20	31.19		450.54	30.75	0.62
Posttest mean		64.32	34.03		63.53	33.28	0.64
Mean gains		23.17			23.09		0.90
Accuracy							
Pretest mean	78.09		18.26	79.25		17.84	0.26
Posttest mean	89.01		14.07	89.28		18.86	0.69
Mean gains	11.01			10.23			0.14

Note. Nonsense word fluency is reported in whole words read, oral reading fluency is reported in correct words per minute, and accuracy is reported in percentage of words read correctly.

increases, these two groups of students made the expected gains in NWF, but were shy of district goals on the other two measures. As shown in Figure 2, the treatment group did make slightly larger gains in NWF (from pre- to posttest) than the control group. The treatment group also scored comparable to the control group on the ORF and ACC posttest measures. Regardless of the subtle differences in results, analyses determined there were no statistically significant differences in posttest measures or gain scores between the groups, thus the treatment (basal reading instruction with a supplemental structured/systematic phonics program) seemed to provide no benefit in beginning reading skills over the control group (basal instruction with no supplement).

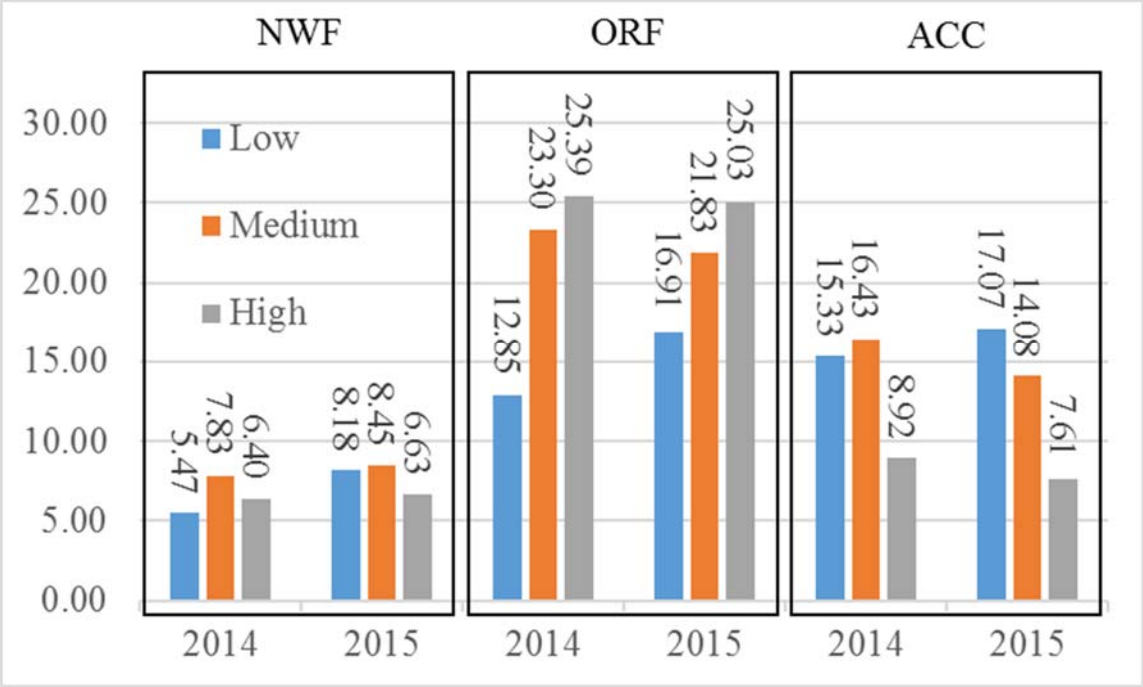

Figure 2. Gains from pre- to posttest for selected measures by control and treatment years, all subgroups.

Results by Measure

As noted in the previous section, there was no significant difference in the students' overall beginning-of-year reading levels. A challenge in reviewing the analyses of this study involves understanding results that may be masked by examination of means and averages. As is typical, classrooms contain a variety of students whose reading skills vary greatly. The subgroups explored consisted of students who ranged from low to high readers. Based on reading scores at the beginning of the respective years, students were categorized as low, medium, or high in order to understand how the basal or basal-plus-phonics programs might differentially support their reading achievement. The distribution of students assigned to these three subgroups is shown in Table 3.

Table 3

Group and Subgroup Distribution

	Control group		Treatment group	
Subgroup	*n*	%	*n*	%
High readers	543	68.1	601	70.0
Medium readers	107	13.4	127	14.7
Low readers	148	18.5	131	15.3
TOTAL	798		859	

The data were divided by year and each subgroup data set was prepared individually (year by year, and subtest by subtest). A single-factor analysis of variance was applied to determine if the gains between years (control vs. treatment) were statistically significant for each of the selected measures. Subgroup means were calculated for each group (control and treatment), each measure, and gains for each subgroup and each measure, for the purpose of understanding the particularities of subgroup performance.

Nonsense word fluency. The data in Table 4 show that students designated as high readers based on beginning scores did not vary by year (treatment and the control groups) on the NWF measure, meaning there was no apparent benefit from the phonics supplemental program for higher-level first-grade readers. Medium-level students' scores indicate that the treatment subgroup did make slightly greater grains in NWF (.61 whole words read) compared to the control group, though the difference was not statistically significant. This outcome suggests that these students did not benefit substantially from the phonics

Table 4

Means, Standard Deviations, Gain Scores, and p Values, by Subgroups for Nonsense Word Fluency

	Control group		Treatment group		
Subgroup	Mean	*SD*	Mean	*SD*	ANOVA *p* values
High readers					
Pretest mean	28.16	13.79	28.07	13.15	0.91
Posttest mean	34.57	12.71	34.71	11.72	0.84
Mean gains	6.40		6.63		0.68
Medium readers					
Pretest mean	16.23	10.99	14.03	8.75	0.14
Posttest mean	24.06	11.83	22.75	10.97	0.38
Mean gains	7.83		8.45		0.63
Low readers					
Pretest mean	8.76	7.96	9.70	9.24	0.37
Posttest mean	14.17	10.55	17.74	12.92	0.01
Mean gains	5.47		8.18		0.01

Note. Nonsense word fluency is reported in whole words read.

treatment. The lower-level readers in the treatment group, on the other hand, experienced a greater gain compared their control-group counterparts on NWF, a difference of an average of 2.70 whole words read. As indicated in Table 4, although the treatment group scored 2.20 words less on the pretest and 1.3 words less on the posttest, they still made more gains in NWF than did the control group (albeit with 10% fewer students in this treatment subgroup). Analysis indicates that gains were significant (at the 0.01 level) in favor of the treatment group. Figure 3 also illustrates the greater improvement of the lower-reader subgroup. Although all subgroups made good progress, the low readers benefited significantly (statistically speaking) from the treatment, more so than the control group as measured by the NWF subtest.

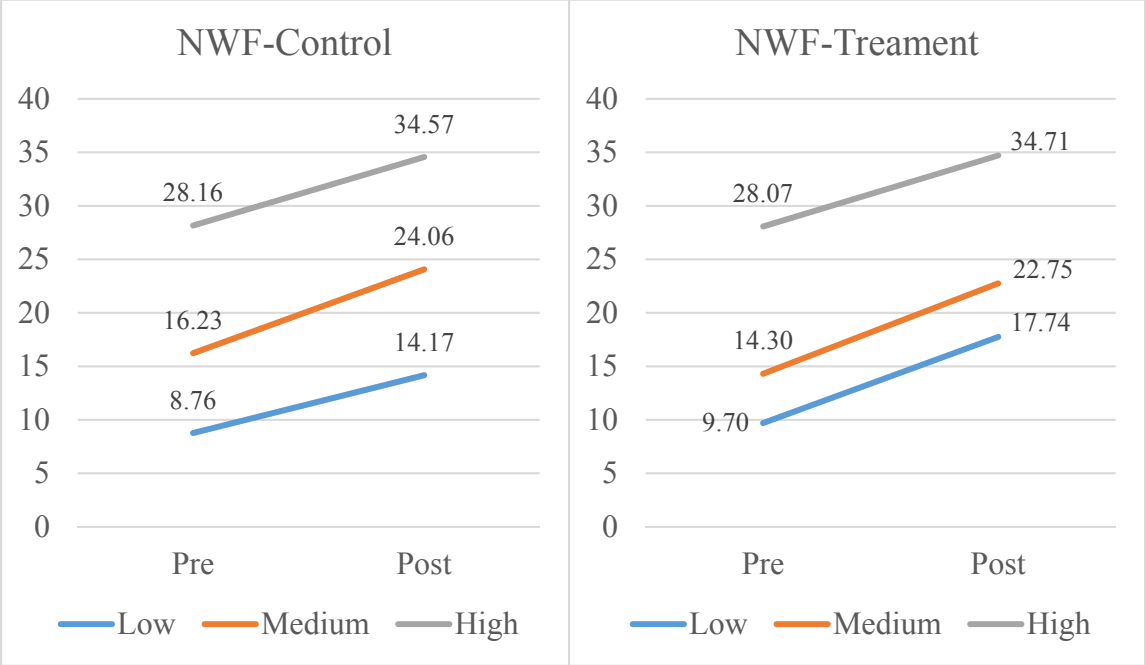

Figure 3. Gains from pre- to posttest for nonsense word fluency by treatment and control years, for all subgroups.

Oral reading fluency. As shown in Table 5, high readers made greater gains on the ORF posttest than other subgroups in both treatment and control groups. The high readers cannot be said to have benefited significantly from the treatment. Medium-level students in the treatment group scored an average of 3.56 correct words lower than the control group on the ORF at posttest, a notable difference, though not statistically significant. Medium readers in the treatment group scored 2.20 whole words less initially and gained 1.47 correct words fewer from pre- to posttest, a nonsignificant difference. The control group did make slightly more gains overall than their treatment group counterparts, though the difference was again not significant. The results for the low readers, however, showed that the treatment group scored 2.01 whole words more initially, but gained 4.06 correct words

Table 5

Means, Standard Deviations, and Gain Scores, by Subgroups for Oral Reading Fluency

	Control group		Treatment group		
Subgroup	Mean	*SD*	Mean	*SD*	ANOVA *p* values
High readers					
Pretest mean	50.99	32.21	51.11	31.13	0.94
Posttest mean	76.37	31.41	76.14	30.01	0.90
Mean gains	25.39		25.03		0.67
Medium readers					
Pretest mean	23.50	12.55	21.41	10.66	0.17
Posttest mean	46.80	20.28	43.24	21.40	0.19
Mean gains	23.30		21.83		0.40
Low readers					
Pretest mean	12.23	8.53	15.24	12.02	0.11
Posttest mean	25.98	17.52	31.95	21.52	0.01
Mean gains	12.85		16.91		0.01

Note. Oral reading fluency is reported in correct words per minute (CWPM) read.

more over all; ending 5.97 whole words higher. The differences in post-mean as well as mean gains were statistically significant.

Figure 4 shows that the trajectory of medium readers in the treatment group was not as steep as the that of the medium reading in the control group, suggesting that these students did not perform as well with a supplemental phonics program paired with a basal reader. The lower readers in the treatment group, in contrast, again fared better on ORF, reading at posttest an average of 5.96 correct words per minute more than the control group. The average gain for the treatment group was 4.06 correct words greater than the control group, a statistically significant difference. The difference was again significant for the low readers in the treatment group, indicating that they also benefited from the

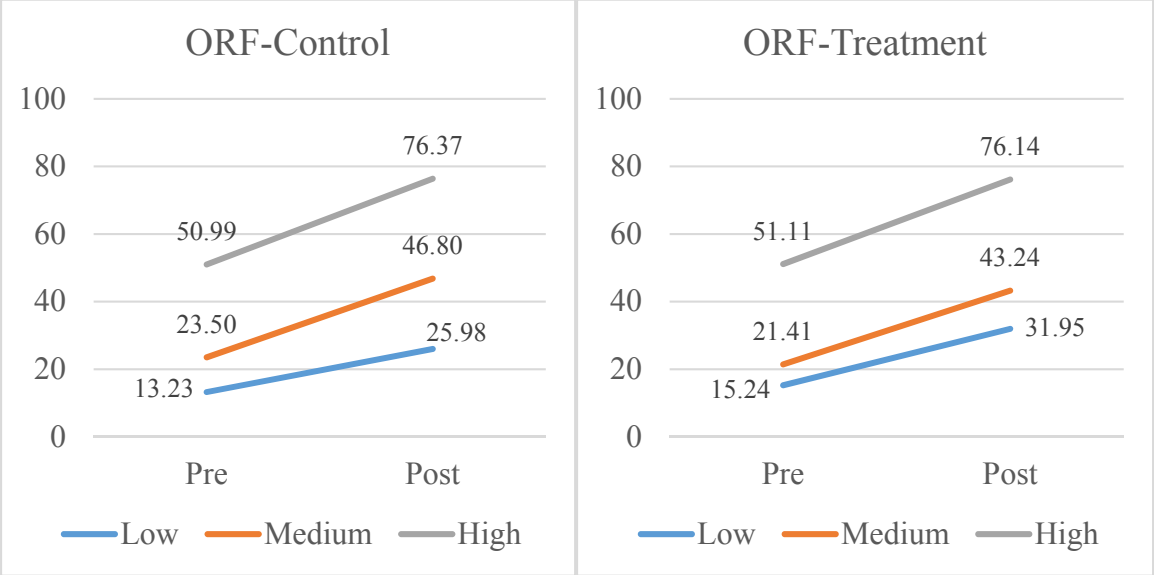

Figure 4. Gains from pre- to posttest for oral reading fluency by treatment and control years, for all subgroups.

supplemental phonics program more than their peers in the treatment group, as measured by the ORF subtest.

Accuracy. As shown in Table 6, higher readers again performed as expected; however, the treatment group did start 1.84% points higher at the pretest (a significant difference), and ended 0.53% points higher at posttest (a nonsignificant difference). The treatment group students' gains were lower than the control group's by 1.31% points, a significant difference, indicating again that the supplemental phonics program did not seem to benefit higher readers. It may have, in fact, been detrimental to high readers. Medium-level students in the treatment group read 2.44% more accurately at the posttest than the control group, a marginal significance of 0.05, suggesting some gains related to the supplemental phonics program.

Figure 5 indicates that, as before, the lower readers from the treatment group were more accurate in their reading at the posttest administration, scoring 5.12% points higher

Table 6

Means, Standard Deviations, and Gain Scores, by Subgroup for Accuracy

	Control group		Treatment group		
Subgroup	Mean	*SD*	Mean	*SD*	ANOVA *p* values
High readers					
Pretest mean	84.72	13.41	86.56	11.17	0.01
Posttest mean	93.64	8.18	94.177	6.91	0.24
Mean gains	8.92		7.61		
Medium readers					
Pretest mean	70.72	14.28	70.63	13.47	0.96
Posttest mean	87.16	10.56	84.71	11.54	0.09
Mean gains	16.43		14.08		0.05
Low readers					
Pretest mean	54.62	19.98	58.38	21.27	0.13
Posttest mean	69.53	20.54	74.66	21.23	0.04
Mean gains	15.33		17.07		0.30

Note. Accuracy is reported in percentage of correct words read.

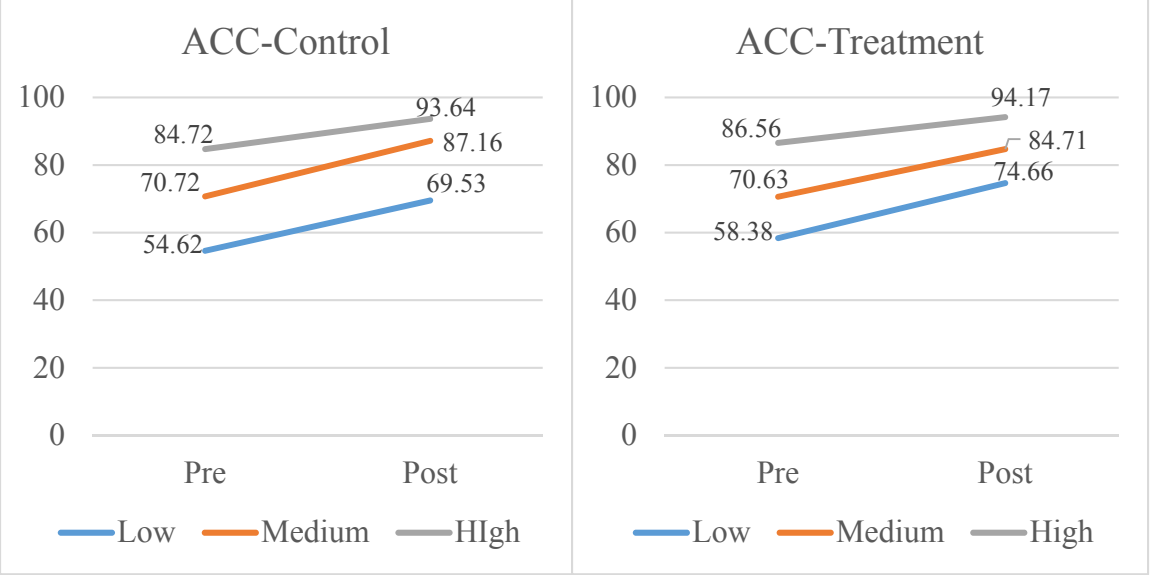

Figure 5. Gains from pre- to posttest for accuracy by treatment and control years, for all subgroups.

than the control group, which is a significant difference in posttest scores. The difference in mean gain score was only 1.73% points, statistically nonsignificant.

Effect size. Table 7 displays the effect size of treatment versus control groups, as reported by Cohen's d and correlation measures of effect size. As previously discussed, low readers were the only students affected by the treatment. Nonsense word fluency and oral reading fluency were again the most impacted areas of treatment. Again, an item of note is the negative effect size for high and medium readers in both oral reading fluency and accuracy, suggesting that perhaps the supplemental phonics program was of little value to those students.

Interaction Results

Research Question 2 asks if possible interactions exist among selected measures

Table 7

The Effect Size of Treatment vs. Control, Reported in Cohen's d and Correlation Measures of Effect Size (r)

Measures	Cohen's *d*	*ES* correlation (*r*)
Nonsense word fluency		
High	0.02455	0.01228
Medium	0.06250	0.03123
Low	0.33307	0.16427
Oral reading fluency		
High	-0.02508	-0.01254
Medium	-0.11073	-0.05528
Low	0.32928	0.16245
Accuracy		
High	-0.15027	-0.07493
Medium	-0.25513	-0.12654
Low	0.12424	0.06200

Note. Accuracy is reported in percentage of correct words read.

relating to performance differences on end-of-year reading achievement scores. Table 8 displays the test of significance for the interaction effects across measures, indicating that the gains by each group depended on the beginning-of year reading level. As Figures 2-5 illustrate, lower readers in the treatment group apparently benefited from the supplemental phonics program across all measures. The corresponding students in the control group did not benefit to the same degree from the basal reading program, at least as reflected by the DIBELS subtests. These findings also indicate that medium-level readers in the control group slightly outperformed their treatment group counterparts, suggesting that these "middle" students did not benefit from this program to the same degree as the low readers. As noted earlier in Tables 3 and 4, high readers experienced no significant difference regardless of group on the NWF and ORF subtest, indicating that these students performed adequately regardless of the instructional programs used. There was a significant difference between groups on the ACC subtest, with some evidence pointing to possible detriment from the supplemental phonics program for high readers. Additionally, the three-way interaction of ORF*reading level*instructional year indicating that there seem to be some interaction of variables in play that affect student performance, which affect students differentially. For example, although only conjecture, perhaps more time spent on skills meant less time on reading that would have supported oral reading fluency development. This would suggest that the time spent on supplemental phonics was of limited value for medium and high readers, who may have benefitted from applying phonics knowledge in reading of real texts.

Table 8

Tests of Significance of the Interaction Effects Between Groups by Treatment/Control and Beginning-Of-Year Reading Levels

Outcome	Interaction effect	*t* value	*p* value
Nonsense word fluency	Test score * reading level	-3.536	0.0004
Oral reading fluency	Test score * reading level	-2.160	0.0309
	Reading level * instructional year	-1.005	0.0339
	Test score * reading level * instructional year	2.157	0.0311
Accuracy	Test score * reading level	3.955	$7.87e^{-05}$

Note. Test score = posttest score, reading level = beginning-of-year reading level, instructional year = treatment vs. control.

Key Findings and Summary

In many ways, the findings of this study are axiomatic and forthright: there were no significant differences between student performances across the three measures from one instructional method to another, when examined as whole, heterogeneous, groups. However, when the data were examined more closely, a more nuanced finding surfaced. Lower readers participating in a first-grade supplemental phonics program paired with a basal reading program made significant progress (scoring significantly higher at posttest, making significant gains, or both—see Tables 3, 4, and 5) on all of the selected measures, when compared to students of similar skill level who received only basal reading instruction. The greatest percentage gains were in ACC; students in the treatment group gain 17.07% points of accuracy (a difference of 1.74% points between treatment and control). The greatest difference in gains was made in NWF and ORF, where students scored significantly higher at posttest (3.57 whole words and 5.97 correct words more, respectively) and made significantly more gains (2.71 whole words and 4.06 correct

words more, respectively). This study supports the claim made by numerous early-literacy research studies that low readers benefit from structured and systematic phonics instruction (Hattie, 2009; NRP, 2000), as measured by selected measures of the DIBELS Next assessment.

Other findings from this study are less straightforward. Research question two targeted possible interactions between the selected variables. The finding of the two-way interactions in the model reinforce the findings previously stated and displayed in Tables 2-4 and in Figures 2-4. The three-way interaction of test scores*reading level* instructional program for ORF suggests that some interaction of variables related to how students performed on the ORF subtest. As previously stated, time spent on this supplemental phonics program may have been of limited value for the majority of students. Time spent practicing oral reading fluency may have been of greater benefit to the more proficient first-grade readers. The three-way interactions clearly reiterate that reading level influences gains made across programs, though a more powerful measure or larger sample could help explain the specific relationship or support the assertion that medium and high readers may have been better served engaged in other reading activities.

What is clear, however, is that for the lower readers in the treatment group, Saxon Phonics as a supplemental instructional program contributed to significant differences and gains in reading scores. From the mixed effects model, it was determined that there were no significant differences or gains for medium and high readers on the NWF and ORF measures, but Saxon Phonics program did benefit low readers the most.

CHAPTER V

DISCUSSION

A great deal of research provides evidence that systematic and structured phonics instruction is the best method to teach early literacy skills for a majority of children (Armbruster et al., 2010; Ehri, 2004; Shanahan, 2005). This study sought to explore the effects of systematic and structured phonics instruction when paired with basal reading instruction on indicators of beginning reading achievement among first-grade students. The sample consisted of 1,657 first-graders attending school in one mountain west, semi-rural, school district. Students included in this study were enrolled for at least 160 days of first-grade instruction. The dependent variables included in this study were three subtests of the DIBELS Next assessment (nonsense word fluency, oral reading fluency, and accuracy). The independent variables were the instructional programming utilized by classroom teachers (identified by instructional year) and the reading levels of low, medium, and high as determined by scores on beginning-of-the-year benchmarks. This chapter will briefly revisit the results and key findings of the data analyses, then discuss the context of the study, implications for research, implications for practice, and concluding remarks. Discussion and implications will be organized by findings (i.e., group, then subgroup).

Key Findings

A mixed effects model was used to explore reading achievement outcomes of a systematic and structured phonics program—Saxon Phonics (Simmons, 2003)—paired

with basal reading instruction—*Treasures* (Bear & Bear, 2007)—on select measures of the DIBELS Next assessment (nonsense word fluency, NWF; oral reading fluency, ORF; and accuracy, ACC) of first-grade students. Comparisons of control to treatment years revealed that group performance from pre- to post-assessments did not differ significantly across any of the selected measures. These findings were unexpected and on the surface seemed to contradict the body of research on phonics instruction and its effect on beginning reading skills (Adams, 1990; Armbruster et al., 2010; Balmuth, 1982; Brady, 2011; Chall, 1990; Dykstra, 1968; Ehri, 2004; Ehri et al., 2001; NRP, 2000; Shanahan, 2005; Tunmer & Nicholson, 2011). Given the great body of research indicating that direct phonics instruction positively impacts reading achievement score in early grades (Bos & Vaughn, 2002; NRP, 2000; Snow et al., 1998), and the mandated use of 70 minutes of instructional time focusing on phonics and spelling in the selected district's literacy model, it was expected that such time and attention would positivity affect decoding efficiency and the related skills of word-reading accuracy and oral reading fluency.

A closer exploration of the data revealed that when students were examined in subgroups (high, medium, and low readers) some significant differences did exist. Lower first-grade readers made significant progress (scoring significantly higher at posttest, made significant gains, or both on all three of the selected measures). There were, however no significant differences or gains for medium and high readers on any of the selected measures. These findings, are in line with current research—that low-skill students benefit more from systematic and explicit phonics instruction (Hattie, 2009;

NRP, 2000).

There were also four two-way interactions and a single three-way interaction among the selected variables that may have affected student performance. These interactions are listed below by measure.

1. NWF-test score*reading level, which indicates that performance on the nonsense word reading task was influenced by beginning-of-the-year reading benchmark scores for both groups of students.

2. ACC-test score*reading level, indicating that word-reading accuracy (i.e., the percentage of words read correctly in a text for one minute) was influenced by beginning of the year reading benchmark scores for both groups of students.

3. ORF-test score*reading level, reading level*instructional year, test score*reading level*instructional year— Oral reading fluency is measured as the number of words in a passage read in one minute. This is a more complicated interaction and not easily interpreted. One explanation is that oral reading performance of first grades is related to the practices promoted by the respective programs and an outcome of the level of reading proficiency.

Interpreting the interactions above is a challenge. The interaction of beginning-of-year reading level and test scores is axiomatic in education (i.e., better readers typically have higher test scores and better readers tend to continue to outscore their lower-performing peers). The two-way interaction reading level*instructional year is less clear. The beginning-of-year reading levels varied by instructional year, but there were no significant differences between mean scores by year or significant differences in the standard deviation of scores between years. Because this analysis used all the same subjects across all three measures, the data were not broken into subgroups, but analyzed by subtest. Given that the mixed effects model compared whole group to whole group, this interaction could have been present in all three subtests, not just ORF. One interpretation from this interaction is that ORF scores varied based on the beginning-of-

year reading levels differently by year. A more powerful measure would be needed to properly explain any relationships among the selected variables and their interactions. However, the two-way interaction of reading level*instruction does indicate that use of this basal program paired with a prepackaged phonics program (without adjustment, adaptation, or differentiation) may not meet the needs of all students, supporting the claim made by numerou literacy research studies that one program does not prove best for all students (Sippola, 1985; Snell, 2007).

Delimitations and Limitations

This study utilized pragmatic research and data collection designs, which may limit this study and its findings. Many variables worthy of consideration were not included in this study (e.g., gender, school demographics, etc.). Several of these limitations and delimitations are important to note. First, extraneous variables such as maturation of the students was not controlled for. The use of a control group and a complete sample was an attempt to reduce the influence of extraneous factors. The examination of many of these variables has merit, but only the identified variables (the selected subtest scores of NWF, ORF, ACC; and the student beginning-of-year reading level of low, medium, and high) were the focus of the study.

Second, although a relatively large and near complete data set, the research was performed using available district data. As previously noted, many variables were not included out of practicality and time constraints. For example, this study did not control for possible differences in school populations or teacher differences, but rather utilized

data from across the district and year-to-year. Differences in student performance between schools as well as differences between students of different SES backgrounds may exist (and most likely do). Much of the research into phonics indicates that explicit phonics instruction benefits lower-achieving students, early readers, as well as low-SES students (Curriculum Associates, 2011, Fountas & Pinnell, 1996; Hattie, 2009; Magnin, 2011) and such factors are worthy of study, albeit not a focus of this study.

Finally, this study included a larger sample size than many of the comparable studies completed recently (Gomez, 2016; Jamaludin, Alias, Mohd Khir, DeWitt, & Kenayathula, 2016; White, 2017). Differences between classrooms and teachers' instruction likely exist but were not controlled for because these differences would be not easily identified given the constraints previously noted (i.e., length of this study, large geographic size of the research site, and number of classrooms). Differences among classrooms nested in schools or differences among schools were not accounted for in this whole group comparison reading achievement from year to year.

Contextualizing the Findings

Despite the noted constraints, details about the instructional conditions might enable a better understanding of the performance results from the group-to-group comparison. The reading curriculum during the control year consisted of basal reading instruction per the norm within the district. No significant changes had been implemented for a number of years. Teachers were familiar with the basal reading program (*Treasures*), and had been using it for the previous four years. According to district

report, teachers had also been allowed to supplement their instruction with found or teacher-designed instructional components paired with the basal program when teachers or grade-level teams identified perceived weaknesses within the curriculum. This cobbling of supplemental components and teachers' expressed desire for stronger phonics instructional components, according to district personnel, drove the adoption of the Saxon Phonics program. But, the same teacher-driven curricular variations might have augmented the control-year instruction in ways that actually supported reading skills measured by DIBELS.

The treatment year consisted of reading instruction using the existing basal reading program (*Treasures*) paired with the new supplemental phonics program (Saxon Phonics). This was the first significant programing change since the adoption of the existing basal curriculum in 2012. Few teachers were familiar with Saxon Phonics, though acquaintances in other districts were familiar with and had used the program. The district reported that the teachers in this study were trained to use the program during one 2-hour session, during which a representative of the publishing company outlined the program, explained all the program components, and demonstrated what each component should look like. As reported by school administrators, entire grade-level teams participated in the training regardless of their level of familiarity with Saxon Phonics, or their belief in its efficacy. Teachers were offered the opportunity to meet with a Saxon Phonics representative of the publishing company six weeks into the treatment year for additional mentoring/coaching—of which first-grade teachers from only three of the nine elementary schools took part. Teachers participating in this additional training were

included because they, the instructional coach, or building administrator requested additional training. These circumstances cast doubt on the fidelity of implementation of the phonics program, which in turn may have affected the assessment outcomes.

As implemented in the district, 70 minutes daily were allocated to Saxon Phonics, a substantial amount of time during an instructional day. The amount of mandated time for *Treasures* was reduced to 60 minutes. In classrooms during the treatment year, *Treasures* was supposed to be utilized for vocabulary (10 minutes), writing and grammar (30 minutes), and fluency (20 minutes) instruction. Thus, how teachers were to supposed to use the basal program changed significantly from the control year to the treatment year. However, there was no documentation of how tightly teachers kept to the mandated schedule. A lack of adherence to the new schedule could be a confounding factor.

The implementation of a new program that most teachers were not familiar with, adjusting and mandating of time allocated to specific elements of reading, and the limited amount of training provided by the district were factors that may have played into the overall effectiveness of the addition of Saxon Phonics. These contextual issues inform the implications of the findings of this study.

Discussion of Findings

Possible Explanation of Group Findings

The following interpretations may explain the lack of significant difference in student performance between groups across the two years of data.

DIBELS next assessments may not measure benefit. According to Good et al.

(2004), DIBELS is a valid and reliable measure of students' early literacy skills. A natural assumption would be that DIBELS assessment is a predictor of students' abilities to generalize skills taught during reading instruction. However, the DIBELS Next assessment is a screening instrument used to determine if students require assistance, not to measure how well a student reads or to measure the upper limit of a student's reading achievement. Perhaps a 1-minute probe was insufficient to show the benefit of the supplemental phonics program because the tasks may be unfamiliar, or because the nature of the task was uncomfortable for some readers. For example, NWF as a test of the alphabetic principle (i.e., letter-sound correspondences and blending letters into words) can confuse children who sense that any combination of letters represents a word that they should know. Students have been trained to read words that hold meaning. The very nature of NWF is for students to read words that have no meaning. Some students attempt to construct meaning from these nonsense words by sounding them out or by substituting known words for unknown nonsense words, lowering their subtest scores. The inclusion of activities to apply or encode phonics, such as oral dictation and spelling, would be a more complete measure of phonics skills gained as a result of instruction and could offset the resistance some students have for reading nonsense words.

It should be noted that of the DIBELS assessments used in this study, only one targeted phonetic decoding (i.e., NWF). The other two subtests measured other reading skills—word-reading accuracy and oral-reading fluency. These two competencies involve more than decoding. Instead, they involve efficient word retrieval and processing speed. As such, they measure decoding indirectly and privilege words that students know more

automatically.

There is also the possibility that student performance on the selected measures reached a threshold where effect can no longer be measured. For example, one can only read aloud so fast, and reading fluency has a ceiling effect, especially for higher-level readers. Such a ceiling effect may result in findings that do not adequately describe the effect of treatment on students' ability to read at the upper range of student performance.

Comments from parents revealed positive reactions, including surprise at the types of words students were able to spell, as well as the level of books students were able to read. These informal reports may support the conclusion that the selected subtests were not accurate measures of phonics skills learned from Saxon Phonics. The results of this study indicate a need for a more accurate measure of student progress by an assessment that is closely aligned to the curriculum being delivered, in order to accurately describe the possible effects of supplemental phonics, or any instructional program.

DIBELS next assessment are not true pre-/post assessment for first-grade. Students who received instruction during the treatment year received 9 months of instruction in the new program. The design of the DIBELS Next assessment is structured such that the use of the selected measures only spanned the last four months of instruction. The selected measures in this study are not administered at the beginning of the year. A beginning-of-year to end-of-year assessment could measure the benefit of an entire year of instruction, perhaps resulting in different findings. Though Saxon Phonics does include weekly diagnostic assessments, they are not designed to measure progress toward overall phonics proficiency, but rather assessment of mastery of skills taught

during the week. Because Saxon Phonics is designed to begin at the first of the year and run until the end of the school year, there is no placement test that could be used as a pre-/post-measure of phonics knowledge. Findings of this study suggest that DIBELS is not an adequate pre-/post- assessment for first-graders' phonics development because it does not include true baseline data.

Teachers' inexperience with the program. As with any new skill or performance task, inexperience is often associated with initial lack of quality of implementation (Darling-Hammond, 2000; Moir, Barlin, Gless, & Miles, 2009). The teachers' inexperience with the phonics program, lack of understanding and a feel for how it could fit within the context of their instructional day may have affected the fidelity of implementations. In addition, not yet knowing how the program could be adjusted to meet the needs of all learners may have led to a lack of quality of instruction that would have benefitted the higher and lower readers. The finding that higher and lower readers did not seemingly benefit from Saxon Phonics combined with the inexperience of the teachers with the program may actually indicate that the program could benefit future students as teachers gain experience—though this is only conjecture.

Though not formally a part of the research design, teachers, instructional coaches, and school administrators indicated a positive view of the program. Several teachers reported that entire classes ended the year reading at higher guided-reading levels than previous years (perhaps this would have been a better measure of the program's effectiveness). Teachers reported that students' classroom behavior during Saxon Phonics instruction was much improved over similar times before the implementation of the

program. School administrators confirmed higher student-engagement rates during classroom observations of Saxon Phonics instruction.

Nevertheless, instructional coaches and school administrators reported that not all classrooms were implementing Saxon Phonics as designed. Some reported a complete lack of the program's utilization altogether. More familiarity with the program and experience with its use may lead to a more complete implementation and better integration into the instructional day improving the overall quality of reading instruction and improving student outcomes. Once again this is only conjecture, though adherence to scripted programs and quality of program delivery is supported by research as an element of success in reading programs (Benner, Stage, Nelson, & Ralston, 2010; Shelton, 2010; Stein et al., 2008). This suggests that a better monitoring system may be needed to confirm productive implementation of the Saxon Phonics program and possible positive effects on measures of reading achievement. Measuring programs before they are fully implemented and before personnel have sufficient experience to adequately deliver instruction complicates the validity of research results. The findings of this study reveal that, contrary to the vast body of research, systematic and structured phonics instruction was generally not effective and may have been due to a lack of experience with the program by teachers at the research site. This indicates that familiarity with a program is essential to its ultimate success.

Possible Explanations of Subgroup Findings

As noted in the key findings, the low-reader subgroup benefited most from the treatment, outperforming the control-group counterparts, and experiencing a statistically

significant difference in gains compared to the other subgroups. These findings were expected and exciting. These findings may be explained by the following explanations.

Direct phonics instruction benefits at-risk students most. As previously noted above and in Chapter II, at-risk students typically benefit most from direct phonics instruction because it explicitly teaches the relationships between English phonemes and graphemes using a variety of hands-on activities beginning with the smallest units and building to blends and more complex combinations (Curriculum Associates, 2011, Fountas & Pinnell, 1996; Hattie, 2009; Magnin, 2011). The publisher advertises Saxon Phonics as a program designed to "captivate all student and ELLs" (Houghton Mifflin Harcourt, n.d.). It could be argued, however, that Saxon Phonics is designed to reach the most at-risk students because direct phonics instruction has been repeatedly shown to be highly effective at reaching these students (Curriculum Associates, 2011, Fountas & Pinnell, 1996; Hattie, 2009; Magnin, 2011). Though there are many definitions of at-risk students, for the purpose of this study "at-risk" equates to "low readers" or those who fell below the designed cut score (a score of 97) on the beginning-of-year DIBELS Next composite. Additionally, 75 or 50.6% of the low readers who received instruction through Saxon Phonic in the treatment year advanced from low to medium readers (compared with 49 or 37.4% of low readers in the control group) by the end of the year. Given that the high and medium readers on average performed as expected, there was no harm to students and inclusion of Saxon Phonics seemingly supported 13.2% more low readers. Ultimately, measuring a treatment that, either intentionally or unintentionally, is designed to meet the needs of a specific population may not yield effect on those for

whom instruction was not targeted. Findings of this study suggest that direct phonics instruction benefits at-risk or low readers the most.

No measure for differentiation. Differentiation has demonstrated a clear effect on student performance in a variety of classroom settings (Brimijoin, 2001; Cabus, Haelermans, & Franken, 2017; Domina et al., 2016). This district's literacy model utilizes skill-based, tier two intervention groups intended to help all students, not just those who are struggling or are not proficient. However, Saxon Phonics is a scripted program, and though options for differentiation are included, they are not readily accessible to teachers and require significant redesign of the prepared and scripted lessons. There was no measure in this study for determining to what extent teachers differentiated their instruction in their classroom to meet the needs of their particular students. In a recent study by McKeown et al. (2016), findings indicated that teachers struggle to differentiate instruction without direct and repeated coaching and mentoring. They also found that students made significantly more progress (relative to ability level) when instruction was differentiated. Inclusion of a qualitative component (such as sustained direct observation) may have helped to explain the possible effects for high and medium readers. Measuring differentiation in tier-one instructional delivery would better describe program implementation and teacher effectiveness. A measure for differentiation could shed light on to how better to tailor instruction (of any instructional program) to meet the needs of diverse students in a classroom. Results from this study indicate that in order to measure program effectiveness applied in a tier-one (or whole group) setting, educators must assess the extent of differentiation to determine program effectiveness.

The possible conditions mentioned here (i.e., DIBELS as an inaccurate measure, a lack of a true pre-/post- measure, a lack of experience with the treatment, a treatment that by design benefits lower readers, and a lack of measurement for differentiation) may explain both the lack of difference in whole group results, as well as the differences in low readers (between the control and treatment groups. Change in these conditions may have painted a different picture of the effect of a supplemental phonics program, or explain why the program did not achieve the expected results. These possible explanations have clear connections in the research as well as implications for future research.

Connections to Current Research

A review of select but similar current studies reveals research that utilized the STAR Early Literacy Test (Renaissance Learning, 2009) as a measure of the effectiveness of phonics instruction on a more frequent basis. In 2017, White conducted a study on the benefits of systematic, explicit phonics instruction in early grades. White utilized the STAR Early Literacy Test to measure the effect on first-grade students' foundational reading skills. Similar studies by Al Otaiba, Allor, Werfel, and Clemens (2016) and Paprzycki et al. (2017) used the STAR Early Literacy Test to measure effectiveness of phonics instruction on first-grade students. In fact, a search of Google Scholar revealed that 10-15 of the most recent phonics studies used the STAR Early Literacy Test, compared to only 1 of 15 utilizing the DIBELS assessment. It is clear that use of the DIBELS Next assessment has a place, as it is widely used and understood by

researchers and practitioners. However, the result of both this and other recent studies would suggest that, though the DIBELS Next assessment may have value, it may not be the best tool for measuring effectiveness of a specific early literacy program or as a pre-/post-assessment for early grades. However, it may also suggest that DIBELS has simply not been validated through experimentation as an accurate measure of all the phonics skills targeted in Saxon Phonics or similar supplemental programs. This study suggests that other assessments might better measure student learning of phonics elements and their contributions to beginning reading achievement.

A review of similar studies exposed that many contain samples of considerably smaller size ($N = 6$ [Gomez, 2016]; $N = 78$ [Jamaludin et al. 2015]; $N = 24$ [Lemons, 2008]; $N = 58$ [Nasrawi & Al-Jamal, 2017]; $N = 925$ [Shapiro & Solity, 2016]; $N = 18$ [White, 2017]). Of the studies cited here, only one, Shapiro and Solity, contained a large number of study participants. Still, the current study contained nearly twice the number of participant as the Shapiro and Solity study. Additionally, of the smaller populations examined, all were specific and targeted groups (such as students with learning disabilities, English language learners, or socioeconomically disadvantaged students.) None of these studies targeted whole populations of students including more diverse students or complete populations. This study breeches this gulf and indicates that phonics programs may not benefit all students equally.

Implications for Research

There is a vast amount of research (cited previously) that addresses the need for

phonics instruction, clearly indicating that systematic phonics instruction is more effective than incidental instruction at improving students' abilities to decode and utilize the alphabetic principle in English (Ehri et al., 2001; NRP, 2000; Ryder, Tunmer, & Greaney, 2008). The group-to-group comparison results presented here stand in contrast to previous findings, suggesting that for the studied population, direct and systematic phonics instruction was no more effective than basal reading instruction on students' abilities to decode words and develop as more proficient readers. The findings of this study have implications for research relevant to stakeholders and researchers, though these implications are more likely to pertain to the research site or similar contexts.

As noted above, anecdotal data from teachers, instructional coaches, and school administrators may provide useful insight that could help explain these results and clarify next steps. Surveys or interviews could be included in future research studies to accurately explore teachers' attitudes, perceptions, or concerns with the implementation or use of such a program. Focus groups of first-grade teams could be held to discuss the effectiveness of the program or instructional delivery. Qualitative pieces could also possibly gage differentiation or methods for differentiation. Inclusion of qualitative analyses may have helped to explain some of the results of this study and could better clarify findings of similar studies in the future.

As stated previously, different informal assessment measures (perhaps curriculum-based measures better aligned with Saxon Phonics, or more basic and familiar phonics application activities such as spelling)—in addition to the select measures of formal assessments (such as the DIBELS Next assessment or STAR Early Literacy

Test)—could be included to better measure phonics skills and progress toward proficiency. Several recent studies exploring similar topics or programs utilized measures that included both formal and informal measures as covariates (Gomez, 2016; Jamaludin et al., 2016; Nasrawi & Al-Jamal, 2017; Shapiro & Solity, 2016; White, 2017). Exploration of these covariates, triangulated with qualitative data (as previously mentioned) would better paint an explanatory picture.

A subsequent study is recommended that examines these same students at the end of second-grade to determine if the trajectory of student gains continues or changes after another year of instruction with Saxon Phonics. Perhaps the effect of the treatment was not fully realized until phonics skills were utilized in more contextual situations (i.e., reading and writing authentic text). Such longitudinal data may clarify the reason(s) for any differences between groups.

This study did not gather data for demographics; therefore, it did not compare demographic variables across or among schools. Some differences in performance between schools may exist and relate to SES, geographic area, or race. Further analyses of this same data set could include comparison of schools by demographics and compare similar school populations to see if some populations are more affected by use of a supplemental phonics program.

Recommendations for Practice

As noted above, adherence to program outlines and quality of program delivery contribute to the success of instructional programs (Benner et al., 2010; Shelton, 2010;

Stein et al., 2008). Anecdotally, fidelity of implementation was a concern raised by instructional coaches and school administrators related to this comparison. A separate study to examine fidelity of program delivery is suggested for the stakeholders. The results of the current study may be cause for concern, as more than 800 students received reading instruction utilizing the supplemental phonics program and the basal curriculum, but only 131 students seemingly benefited significantly from its use. If issues with program delivery that negatively influence student performance exist, exploration of the fidelity of implementation may help stakeholders to improve the quality and increase the benefit of program beyond those for the low readers. Teachers and administrators could also consider ways to better differentiate for different reading levels within classrooms. Exploration of student performance using program fidelity as a covariate may be part of a program evaluation to determine if expenditure of resources on this program is worth continued financial investment.

Stakeholders may also wish to re-evaluate the use of Saxon Phonics as a supplement delivered to all first-grade students. Given the lack of difference between the high and medium reader subgroups, the additional time spent teaching Saxon may be better utilized extending existing reading skills for high readers or targeting strategic/high-leverage skills for medium readers that may improve their reading scores reserving Saxon Phonics for low readers (Bradley, Danielson, & Doolittle, 2005; D. Fuchs & Fuchs, 2006; S. L. Fuchs, Fuchs, & Amita, 2008).

Conclusion

Numerous researchers have discussed and advocated for the use of systematic and structured phonics instruction in early grades (Adams, 1990; Armbruster et al., 2010; Balmuth, 1982; Bond & Dykstra, 1967; Brady, 2011; Chall, 1967, 1990; Dykstra, 1968; Ehri, 2004; Ehri et al., 2001; NRP, 2000; Nicholson & Tunmer, 2011; Shanahan, 2005). This study provides evidence that, within the specified limitations, such instruction was a benefit to low-achieving readers. It also adds to the literature related to phonics instruction by providing an authentic example of research informing practice, and answers questions about the appropriateness of broadly casting the net of programs with the hope of benefitting all students.

In the post-No Child Left Behind (2000)/Reading First era, great weight is given to student achievement tests that are nationally normed and widely used, so that policy and decision makers can compare student achievement from district-to-district, and state-to-state. Stakeholders justify this focus because difficulty reading impacts a student's ability to perform adequately on achievement measures and will hinder a student's progress throughout school (Allington & McGill-Franzen, 2008; Chatterji, 2006). Low-achieving students deserve systematic and explicit instruction, using empirically and evidence based instructional design and delivery, to make significant gains in reading (Bursuck & Blanks, 2010). The stakes are high, as Stockard and Engelmann (2010) point out, students who do not read at a grade-level benchmark level by the end of first grade suffer greater academic, social and emotional issues that their peers. Educators are on the hunt for programs or strategies that will benefit the most students with the greatest

developmental and fiscal effectiveness. However, not all programs meet the needs of all students. The findings of this study indicate that Saxon Phonics, when applied en masse to all students in a first-grade program, benefitted only some students. However, analysis of the contextual details demonstrates a need for a nuanced understanding of the program's benefits and the constraints of measuring them among large samples. The benefit of providing low and at-risk readers with systematic and structured phonics instruction has been demonstrated in this and related studies. Saxon Phonics is just one tool available to educators to meet the needs of beginning readers; additional tools and adaptation are also clearly needed.

REFERENCES

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

Otaiba, S. A., Allor J., Werfel K. L., Clemens N. (2016) Critical components of phonemic awareness instruction and intervention: Recommendations for teacher training and for future research. In R. Schiff & R. Joshi (Eds.), Interventions in learning disabilities. Literacy Studies (Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education) (vol. 13, pp. 9-27). Zurich, Switzerland: Springer.

Allington, R., & McGill-Franzen, A. (2008). Got books? *Educational Leadership, 65*(7), 20-23.

Armbruster, B., Lehr, F. and Osborn, J. Put Reading First: The Research Building Blocks for Teaching Children to Read, K–3. Washington, DC: The Partnership for Reading: National Institute for Literacy; National Institute of Child Health and Human Development; and U.S. Department of Education, 2001. Ary, D., Jacobs, L., Sorensen, C., & Walker, D. (2010). *Introduction to research in education.* New York: Holt, Rinehart & Winston..

Anderson, R. C., Hiebert, E. H., & Scott, J. A., Wilkinson, I. A. G. (1985). Becoming a nation of readers: The report of the commission on reading. Champaign, IL: Center for the Study of Reading

Baker, R. (2010). Data mining for education. *International Encyclopedia of Education, 7*(3), 112-118.

Balmuth, M. (1982). *The roots of phonics. A historical introduction.* New York, NY: McGraw-Hill.

Bear, D. R., & Bear, D. R. (2007). *Treasures: A reading/language arts program*. New York, NY: Macmillan/McGraw-Hill.

Begeny, J. C., & Silber, J. M. (2006). An examination of group-based treatment packages for increasing elementary-aged students' reading fluency. *Psychology in the Schools, 43*(2), 183-195.

Benner, G. J., Stage, S. A., Nelson, J. R., Laederich, M., & Ralston, N. C. (2010). Predicting the cumulative recidivism of juvenile detainees. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention, 2*(1), 51-62.

Best, J. W., & Kahn, J. V. (2006). *Research in education*. Boston, MA: Pearson/Allyn and Bacon.

Bingham, G. E., & Hall-Kenyon, K. M. (2013). Examining teachers' beliefs about and implementation of a balanced literacy framework. *Journal of Research in Reading, 36*(1), 14-28.

Blamey, K. L., & Beauchat, K. A. (2016). *Starting strong: Evidence-based early literacy practices*. Portland, ME: Stenhouse.

Blevins, W. (1998). *Phonics from A to Z: A practical guide*. New York, NY: Scholastic.

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

Bos, C. S., & Vaughn, S. (2002). *Strategies for teaching students with learning and behavior problems*. Boston: Allyn & Bacon.

Bradley, R., Danielson, L., & Doolittle, J. (2005). Response to intervention. *Journal of Learning Disabilities, 38*(6), 485-486.

Brady, S. (2011). Efficacy of phonics teaching for reading outcomes: Indications from post NRP research. In S. Brady, D. Braze, & C. Fowler (Eds.), Explaining individual differences in reading: Theory and evidence. New York, NY: Psychology Press.

Brimijoin, K. (2005). Differentiation and high-stakes testing: An oxymoron? *Theory into Practice, 44*(3), 254-261.

Brooks, P. H., & Baumeister, A. A. (1977). A plea for consideration of ecological validity in the experimental psychology of mental retardation: A guest editorial. *American Journal of Mental Deficiency, 81,* 407-416.

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

Burgess, S. R. (2006). The development of phonological sensitivity. *Handbook of Early Literacy Research, 2,* 90-100.

Bursuck, B., & Blanks, B. (2010). Evidence-based early reading practices within a Response to Intervention system. *Psychology in the Schools, 47*(5), 421-431.

Bursuck, W. D., Smith, T., Munk, D., Damer, M., Mehlig, L., & Perry, J. (2004). Evaluating the impact of a prevention-based model of reading on children who are at risk. *Remedial and Special Education, 25*(5), 303-313.

Cabus, S. J., Haelermans, C., & Franken, S. (2017). SMART in mathematics? Exploring the effects of in-class-level differentiation using SMARTboard on math proficiency. *British Journal of Educational Technology, 48*(1), 145-161.

Callinan, C., & van der Zee, E. (2010). A comparative study of two methods of synthetic phonics instruction for learning how to read: Jolly Phonics and THRASS. *The Psychology of Education Review, 34*(1), 21-31.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.

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

Chall, J. S. (1990). Policy implications of literacy definitions. Toward defining literacy, 63.

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*(3), 489-507.

Coley, R. L. (2015, November). *Early childhood education policy: Preschool program efficacy, composition, and quality improvement*. Paper presented at the 2015 Fall Conference of the APPM: The golden age of evidence-based policy, Miami, FL.

Cooke, N. L., Kretlow, A. G., & Helf, S. (2009). Supplemental reading help for kindergarten students: How early should you start? *Preventing School Failure: Alternative Education for Children and Youth, 54*(3), 137-144.

Cooter, R. B., & Reutzel, D. R. (2005). *The essentials of teaching children to read: The teacher makes the difference*. Chicago, IL: Pearson.

Curriculum Associates. (2011). *Phonics for reading.* Retrieved from https://www.curriculumassociates.com/products/detail.aspx?title=phonicsreading

Darling-Hammond, L. (2000). *Solving the dilemmas of teacher supply, demand, and standards: How we can ensure a competent, caring, and qualified teacher for every child*. Kutztown, PA: Kutztown Distribution Center.

Dickinson, D. K., McCabe, A., & Essex, M. J. (2006). A window of opportunity we must open to all: The case for preschool with high-quality support for language and literacy. *Handbook of Early Literacy Research, 2*, 11-28.

Dickinson, D. K., McCabe, A., & Sprague, K. (2003). Teacher rating of oral language and literacy (TROLL): Individualizing early literacy instruction with a standards-based rating tool. *The Reading Teacher, 56*(6), 554-564.

Domina, T., McEachin, A., Hanselman, P., Agarwal, P., Hwang, N., & Lewis, R. (2016). *Beyond tracking and detracking: The dimensions of organizational differentiation in schools*. Retrieved from https://ssrn.com/abstract=2806727

Durkin, D., & Allington, R. L. (2004). *Teaching them to read*. Boston, MA: Allyn and Bacon.

Dykstra, R. (1968). Summary of the second-grade phase of the Cooperative Research Program in primary reading instruction. *Reading Research Quarterly*, 4(1), 49-70.

Ehri, L. C. (1979). Linguistic insight: Threshold of reading acquisition. Reading research: *Advances in Theory and Practice, 1,* 63-114.

Ehri, L.C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3-40). Mahwah, NJ: Erlbaum

Ehri, L. (2004). Teaching phonemic awareness and phonics. In P. McCardle & V. Chhabra (Eds.), The voice of evidence in reading research (pp. 153–186). Baltimore, MD: Brookes.

Ehri, L. C., & Roberts, T. (2006). The roots of learning to read and write: Acquisition of letters and phonemic awareness. *Handbook of Early Literacy Research, 2,* 113-131.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*(3), 393-447.

Ekpo, C. M., Udosen, A. E., Afangideh, M. E., Ekukinam, T. U., & Ikorok, M. M. (2007, June). *Jolly phonics strategy and the ESL pupils' reading development: A preliminary study*. Paper presented at the first mid-term conference held at the University of Ibadan, Ibadan, Nigeria.

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

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

Fuchs, S. L., Fuchs, D., & Amita, O. R. (2008). Response to intervention. In E. L. Grigorenko (Ed.), *Educating individuals with disabilities: IDEIA 2004 and beyond* (pp. 115-135). New York, NY: Springer.

Gast, D. L. (2014). General factors in measurement and evaluation. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed., pp. 85-104). New York, NY: Routledge.

Gomez, D. (2016). *The impact of explicit phonics instruction on the fluency rate of first graders* (Unpublished master's theses). Goucher College, Baltimore, MD.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*(3), 257-288.

Good, R. H., Kaminski, R. A., Shinn, M., Bratten, J., Shinn, M., Laimon, D., & Flindt, N. (2004). *Technical adequacy of DIBELS: Results of the Early Childhood Research Institute on measuring growth and development*. Eugene, OR: University of Oregon.

Goodman, K. S., Shannon, P., Freeman, Y. S., & Murphy, S. (1988). *Report on basal readers*. Katonah, NY: Richard C. Owen.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Hattie, J. A. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Abingdon, UK: Routledge.

Houghton Mifflin Harcourt. (n.d.). *Saxon® phonics and spelling: Using instrumental instruction for spelling and phonics grades K-3*. Retrieved from http://www.hmhco.com/shop/education-curriculum/literature-and-language-arts/language-arts/saxon-phonics-and-spelling#sthash.qnvb17ZE.dpbs

Jamaludin, K. A., Alias, N., Mohd Khir, R. J., DeWitt, D., & Kenayathula, H. B. (2016). The effectiveness of synthetic phonics in the development of early reading skills among struggling young ESL readers. *School Effectiveness and School Improvement, 27*(3), 455-470.

Jarvis, S. P. (2016). *Physical literacy and south east Wales primary school children: The role of fundamental movement skills* (Doctoral dissertation). University of South Wales.

Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches*. Los Angeles, CA: Sage.

Jones, B. C. (2006). *The effects of a basal reading program on reading achievement in selected Tennessee schools* (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text database. (Publication No. AAT 3222584).

Joshi, R. M., Binks, E., Hougen, M., Dahlgren, M. E., Ocker-Dean, E., & Smith, D. L. (2009). Why elementary teachers might be inadequately prepared to teach reading. *Journal of Learning Disabilities, 42,* 392-402.

Juel, C. (2006). The impact of early school experiences on initial reading. *Handbook of Early Literacy Research, 2,* 410-426.

Juel, C., & Minden-Cupp, C. (1999). Center for the Improvement of Early Reading Achievement: One down and 80,000 to Go: Word Recognition Instruction in the Primary Grades. *The Reading Teacher, 53*(4), 332-335.

Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers' early literacy development through classroom-based teacher-child storybook reading and explicit print referencing. *Language, Speech, and Hearing Services in Schools, 40*(1), 67-85.

Kaminski, R. A. & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*(2), 215-227.

Kaplan, D., & Walpole, S. (2005). A stage-sequential model of reading transitions: Evidence from the Early Childhood Longitudinal Study. *Journal of Educational Psychology, 97*(4), 551.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Northridge, CA: Wadsworth, Thomson Learning.

Ledford, J. R., Hall, E., Conder, E., & Lane, J. D. (2016). Research for young children with autism spectrum disorders: Evidence of social and ecological validity. *Topics in Early Childhood Special Education, 35*(4), 223-233.

Leib, T. (2001). *The relationship between Saxon phonics, kindergarten through second grade, and student achievement scores in reading*. (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text database. (Publication No. AAT 3014216).

Lemons, C. J. (2008). *Exploring the effectiveness of phonics-based instruction for children with Down syndrome* (Unpublished doctoral dissertation). Vanderbilt University, Nashville, TN.

Lingo, A. S., Slaton, D. B., & Jolivette, K. (2006). Effects of corrective reading on the reading abilities and classroom behaviors of middle school students with reading deficits and challenging behavior. *Behavioral Disorders, 31*(3), 265-283.

Lonigan, C. J. (2006). Development, assessment, and promotion of preliteracy skills. *Early Education and Development, 17*(1), 91-114.

Lonigan, C. J., & Shanahan, T. (2008). Executive summary of the report of the National Early Literacy Panel. National Early Literacy Panel—Developing early literacy: Report of the National Early Literacy Panel.

Machalicek, W., O'Reilly, M. F., Beretvas, N., Sigafoos, J., & Lancioni, G. E. (2007). A review of interventions to reduce challenging behavior in school settings for students with autism spectrum disorders. *Research in Autism Spectrum Disorders, 1*, 229-246.

Maclean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33,* 255-281.

Magnin, M. C. (2011). *A comparison of guided reading and phonics for reading* (Unpublished doctoral dissertation). Saint Mary's College of California, Moraga, CA.

Marzano, R. J. (1998). *A theory-based meta-analysis of research on instruction*. Aurora, CO: Mid-continent Research for Education and Learning.

Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexanderia, VA; Association for Supervision and Curriculum Development.

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA; Association for Supervision and Curriculum Development.

McEwan, E. K. (2003). *Seven steps to effective instructional leadership* (2nd ed.). Thousand Oaks, CA: Corwin.

McKeown, D., Brindle, M., Harris, K. R., Graham, S., Collins, A. A., & Brown, M. (2016). Illuminating growth and struggles using mixed methods: Practice-based professional development and coaching for differentiating SRSD instruction in writing. *Reading and Writing, 29*(6), 1105-1140.

Moats, L. C. (2000). *Speech to print: Language essentials for teachers*. Baltimore, MD: Brookes.

Moir, E., Barlin, D., Gless, J., & Miles, J. (2009). *New teacher mentoring: Hopes and promise for improving teacher effectiveness*. Cambridge, MA: Harvard Education Press.

Murray, B. A., Stahl, S. A., & Ivey, M. G. (1996). Developing phoneme awareness through alphabet books. *Reading and Writing, 8*(4), 307-322.

Nasrawi, A., & Al-Jamal, D. (2017). The effect of using jolly phonics on Jordanian first grade pupils' reading. *International Online Journal of Education and Teaching*, *4*(2), 106-119.

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

National Research Council. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academies Press.

Neuman, S. B. (2006). The knowledge gap: Implications for early education. *Handbook of Early Literacy Research, 2*, 29-40.

Newman, I., Newman, C., Brown, R., & McNeely, S. (2006). *Conceptual statistics for beginners* (3rd ed.). New York, NY: University Press of America.

Nicholson, T., & Tunmer, W. E. (2011). Reading: The great debate. In C. M. Rubie-Davies (Ed.), *Educational psychology: Concepts, Research, and Challenges* (pp. 36-50). London, UK: Routledge.

Paprzycki, P., Tuttle, N., Czerniak, C. M., Molitor, S., Kadervaek, J., & Mendenhall, R. (2017). The impact of a framework-aligned science professional development program on literacy and mathematics achievement of K-3 students. *Journal of Research in Science Teaching, 54,* 1174-1196. doi:10.1002/tea.21400

Perie, M., Grigg, W., & Donahue, P. (2005). *The Nation's Report Card.*™ Reading, 2005. Washington, DC: National Center for Education Statistics.

Phillips, B. M., & Torgesen, J. K. (2006). Phonemic awareness and reading: Beyond the growth of initial reading accuracy. *Handbook of Early Literacy Research, 2*, 101-112.

Price, S. M. (2015). *The teaching of explicit phonics effects on kindergarten reading readiness scores* (Unpublished doctoral dissertation). Liberty University, Lynchburg, VA.

Rasinski, T., & Hoffman, J. (2006). *Seeking understanding about reading fluency: The contributions of Steven A. Stahl*. K. S. D. Stahl & M. C. McKenna (Eds.). New York, NY: Guilford.

Renaissance Learning. (2009). *Guide early growth in less time*. Retrieved from http://www.renaissance.com/products/assessment/star-360/star-early-literacy-skills/

Ritchey, K. D. (2011). The first "R": Evidence-based reading instruction for students with learning disabilities. *Theory into Practice, 50*(1), 28-34.

Roskos, K., Strickland, D., Hasse, J., & Malik, S. (2009). *First principles for early grades reading programs in developing countries.* Retrieved from http://www.equip123.net/docs/e1-EarlyGradesToolkit.pdf

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97-110). New York: Guilford.

Scruggs, T. E., & Mastropieri, M. A. (2000). The effectiveness of mnemonic instruction for students with learning and behavior problems: An update and research synthesis. *Journal of Behavioral Education, 10*(2), 163-173.

Shapiro, L. R., & Solity, J. (2016). Differing effects of two synthetic phonics programmes on early reading development. *British Journal of Educational Psychology, 86*(2), 182-203.

Shanahan, T. (2005). But does it really matter? *Phi Delta Kappan, 86*(6), 452-461.

Shaywitz, S. (2014, September 18). *Evidence-based vs. research-based programs for dyslexics* [Interview]. Washington DC: CSPAN.

Shelton, N. R. (2010). Program fidelity in two reading mastery classrooms: A view from the inside. *Literacy Research and Instruction, 49*(4), 315-333.

Shermer, E. T. (2003). *Reading without Dick and Jane: The politics of literacy in 20th century America*. Retrieved from http://www.rarebookschool.org/2005/ exhibitions/dickandjane.shtml

Shinn, M. (2002). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (Vol. 4, pp. 671–697). Silver Spring, MD: National Association of School Psychologists.

Simmons, L. (1996). *Saxon Phonics K-2: An incremental development*. Norman, OK: Saxon.

Simmons, L. (2003). *Saxon phonics and spelling: Grade 1*. Norman, OK: Saxon.

Simmons, L., & Calvert, L. (1996). *Saxon phonics*. Norman, OK: Saxon.

Sippola, A. E. (1985). What to Teach for Reading Readiness: A Research Review and Materials Inventory. *The Reading Teacher, 39*, 162-167.

Snell, C. A. (2007). *The impact of daily writing on kindergarten students' phonemic awareness* (Doctoral dissertation). Retrieved from Dissertations & Theses: Full Text database. (Publication No. AAT 3261224).

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children committee on the prevention of reading difficulties in young children*. Washington, DC: National Research Council.

Stahl, S. A. (2001). Teaching phonics and phonological awareness. *Handbook of Early Literacy Research*, 1, 333-347.

Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., Yen, L., ... Compton, D. L. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis, 30*(4), 368-388.

Stockard, J., & Engelmann, K. (2010). The development of early academic success: The impact of Direct Instruction's Reading Mastery. *Journal of Behavior Assessment and Intervention in Children, 1*(1), 2-24.

Strickland, D. S. (Ed.). (2010). *Essential readings on early literacy*. Newark, NJ: International Reading Association.

Souther, L. C. (2015). *An evaluation of the letterland, phonics-based reading program for kindergarten through second grade students* (Unpublished doctoral dissertation). Wingate University, Wingate, NC.

Thompson, E. P. (2011). *William Morris: Romantic to revolutionary*. New York, NY: PM Press.

Tompkins, G. E. (2001). *Literacy for the twenty-first century*. Upper Saddle River, NJ: Merrill.

Tunmer, W. E., & Arrow, A. W. (2013). Reading: Phonics instruction. In J. Hattie & E. Anderman (Eds.), International guide to student achievement (pp. 316-319). London: Routledge.

University of Oregon (2015). *DIBELS oral reading fluency and retell fluency*. Retrieved from https://dibels.uoregon.edu/assessment/dibels/measures/orf.php

Utah State Office of Education. (2015, October). *Data reports: Assessment results-NAEP (National Assessment of Educational Progress).* Retrieved from https://www.schools.utah.gov/file/5fe97106-8172-4a2d-be76fcf74b9d7f50

Vaughn, S., & Linan-Thompson, S. (2004). Research-*based methods of reading instruction, grades K-3*. Alexandria, VA; Association for Supervision and Curriculum Development.

White, K. (2017). *Benefits of systematic and explicit phonics instruction with early elementary students* (Unpublished doctoral dissertation). Brenau University, Gainsville, FL.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69*(3), 848-872.

Wicker, K. (2007). *The effect of two reading programs on kindergarten student's reading readiness* (Doctoral dissertation). Walden University, Minneapolis, MN.

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. Alexanderia, VA; Association for Supervision and Curriculum Development.

Williams, C., Phillips-Birdsong, C., Hufnagel, K., Hungler, D., & Lundstrom, R. P. (2009). Word study instruction in the K-2 classroom. *The Reading Teacher, 62*(7), 570-578.

Wilson, S. B., & Lonigan, C. J. (2010). Identifying preschool children at risk of later reading difficulties: Evaluation of two emergent literacy screening tools. *Journal of Learning Disabilities, 43*(1), 62-76.

CURRICULUM VITAE

BRYCE B DAY

505 West 2400 South
Perry, UT 84302
801.645.8847

Education
Ph.D. 2017

College of Education and Human Services
Department of Elementary Education
Utah State University, Logan, Utah
Emphasis: Curriculum and Instruction/Early Childhood Literacy
Dissertation: *Exploring the Relationship between the Use of a Selected Phonics Program and the Oral Reading Fluency and Nonsense Word Fluency Scores of First-grade Students*

English as a Second Language Endorsement

Mountain Plains Regional Resource Center
Logan, Utah
(Complete course-work available upon request)

Administrative Endorsement

College of Education
University of Phoenix – Utah Main Campus
(Completed in conjunction with M.A. graduate program)

M.A. 2007

College of Education
University of Phoenix – Utah Main Campus
Emphasis: Curriculum & Instruction
Endorsement: Educational Supervision and Administration
Thesis: *Structured Training and Mentoring for Paraeducators in Special Education settings*
Professor Sharon Prescott, Ed.D., Action Research Advisor

Teacher Licensure

College of Education and Human Services
Department of Special Education
Utah State University, Logan, Utah
Licensure: Special Education—Severe/Profound
Single Subject Research Project: *Power of Communication*
Cindy Myers, M.S., ATP instructor/Research Advisor

B.A. 2004

College of Social and Behavioral Sciences
Department of Political Science and Philosophy
Weber State University, Ogden, Utah
Major: Political Science / Psychology

Complete Teaching /Work Experience

Box Elder School District 2015-Present Tremonton, Utah
Principal
Mountain View Elementary School
Staffing
Budgeting
Facility Supervision/Maintenance
School Safety Plan
School Improvement Plan
Staff Supervision
504 coordination
Student Discipline

Box Elder School District 2014-Present Tremonton, Utah
Assistant Principal
Alice C. Harris Intermediate School
Staffing
School Safety Plan
School Improvement Plan
6th Grade Supervision
504 coordination
Student Discipline

Box Elder School District 2011-2014 Garland, Utah
Functional Skills/Life Skills Classroom Teacher
Bear River Middle School
Self-contained/Functional Skills

Pre-/Academic Skills
Functional Communication
Fine Motor/ Gross Motor Activities
Social Skills
Life/Independent Living Skills
Paraeducator Management/Training

Weber State University 2009-2016 Ogden, Utah
Adjunct Faculty/Instructor
English/Education Department
ENGL 4410 –*Instructional Strategies and Methodologies for English Language Learners*

Developmental English
ENGL 0955 –Developmental English Reading and Writing

Teacher Education
Student Teaching Supervisor

Learning English for Academic Purposes
ESL 0110 –Written Communication I
ESL 0150 – Pronunciation I
ESL 1241 – Grammar Foundations II
ESL 1230 – Interpersonal Communication
ESL 1250 – Pronunciation II
ESL 2310 –Written Communication III
ESL 2320 –Topics for Academic Purposes III
ESL 2330 –Academic Communication I
ESL 2410 –Written Communication IV

Utah State University 2008-Present Logan, Utah
Adjunct Faculty/Instructor
School of Teacher Education and Leadership
Student Teaching Supervisor
ELED 3000 – Socio-cultural/Historic Foundations of Education
ELED 1010 – Intro to Elementary Education
TEAL 4765 –Second Language Acquisition
TEAL 6765 –Second Language Acquisition
(Distance/Broadcast course)

Utah State University 2008-2011 Logan, Utah
Graduate Teaching Assistant

School of Teacher Education and Leadership
ELED/SCED 4760 – ESOL Instructional Strategies
ELED/SCED 6760 – ESOL Instructional Strategies
Taught in conjunction with Dr. Patricio Ortiz
(Distance course)
ELED 3000 –Historical, Social, Cultural Foundations of Education
Taught in conjunction with Dr. Martha Whitaker
ELED/SCED 4730 – Educational Linguistics
ELED/SCED 6730 – Educational Linguistics
Taught in conjunction with Dr. Gary Ockey
(Distance course)
ELED/SCED 4780 – Assessment for Language Learners
ELED/SCED 6780 – Assessment for Language Learners
Taught in conjunction with Dr. Gary Ockey
(Distance course)
Student Teaching Supervisor

Ogden City Schools 2008-2009 Ogden, Utah
Emotionally Disturbance Classroom Teacher Specialist
District Behavior Specialist
Budgeting
Community Outreach
Functional Behavior Analysis
Behavior Intervention Development
Due Process Procedures

Ogden City Schools 2007-2008 Ogden, Utah
Summer Administrator
Washington High School
Budgeting
Staffing
Curriculum Development
Program evaluation
Student Discipline

Ogden City Schools 2006-2007 Ogden, Utah
Administrative Intern at T.O. Smith Elementary
Budgeting
Staffing
School Safety Plan
School Improvement Plan
Community Council Vice-chair

Program evaluation
Student Discipline

Ogden City Schools 2006-2008 Ogden, Utah
Kindergarten /Special Education Teacher
T.O. Smith Elementary
Self-contained/Functional Skills
Pre-Academic Skills
Functional Communication
Fine Motor/ Gross Motor Activities
Social Skills
Paraeducator Management/Training

Ogden City Schools 2004 – 2006 Ogden, Utah
Kindergarten /Special Education Teacher
Gramercy Elementary
Self-contained/Functional Skills
Pre-Academic Skills
Functional Communication
Fine Motor/ Gross Motor Activities
Social Skills
Paraeducator Management/Training

Weber County School District 2003 – 2004 Ogden, Utah
Paraeducator
Weber High School
Life Skills/Functional Skills/Community Living

Conference Presentations

Presenter. Classroom Behavior Management, Box Elder School District, New Teacher Induction Workshop Brigham City, Utah. August 17, 2014; February 20, 2015.

Presenter. *iPad as Assistive Technology*, CEC New Teacher Conference. Salt Lake City, Utah. March 6, 2013.

Co-presenter. Autism ABCs, Box Elder School District Professional Development Workshop. Brigham City, Utah. October 3, 2012; December 5, 2012; February 20, 2013.

Co-presenter. Autism Basics, Box Elder School District EYE series. Brigham City, Utah. November 26, 2012.

Presenter/Program Assistant. *Positive Behavior Support and Reinforcement in the Classroom*, Running Start: An Ongoing Comprehensive Induction Program for New Unlicensed Special Education Teachers. Salt Lake City, Utah. July 30-August 3, 2012.

Presenter. *Ethics and Diversity.* 18th Annual NAME International Conference. National Association for Multicultural Education. New Orleans, Louisiana. November 12-16, 2008.

Presenter. *Positive Behavior Supports in the classroom: a Win-Win Approach to Behavior Management*. 13th Annual Utah Paraprofessional Conference. Utah Paraprofessional Consortium. Ogden, Utah. November 9-10, 2007

Presenter. *Structured Training and Mentoring for Paraeducators*. 26th National Conference on the Training and Employment of Paraprofessionals. National Resource Center for Paraprofessionals, Albuquerque, New Mexico, May 3-5, 2007.
Presenter. *Positive Behavior Supports in the classroom.* 12th Annual Utah Paraprofessional Conference. Utah Paraprofessional Consortium, Ogden, Utah. November 10-11, 2006.

Other Presentations

Presenter. *Autism ABC's: Foundations of Autism*. Box Elder School District. Brigham City, Utah. October 3, 2012.

Presenter. *Managing Paraeducators in the Classroom*. 06-07 Special Education ATP Severe/Profound Orientation. Utah State University Department of Special Education and Rehabilitation, Cindy Myers instructor. Salt Lake City, Utah. August 2, 2006.

Presenter. *Preliminary Behavior Support*. SPED 5010 ABA I: Principles of Assessment and Analysis. Utah State University Department of Special Education and Rehabilitation, Cindy Myers instructor. Salt Lake City, Utah. March, 2006.

Conference Attendance

USOE, Significant Congitive Disabilities Roundtable: A Day with Neal Kingston. October 12, 2012.

USOE, Significant Cognitive Disabilities Roundtable: The DYNAMIC Meeting. Salt Lake City, Utah. September 24, 2012.

USOE, 2012 Significant Cognitive Disabilities State Conference/Conference Ville. Layton, Utah. February 17, 2012.

18th Annual National Association for Multicultural Education International Conference. New Orleans, Louisiana. November 12-16, 2008.

Utah Council on Children with Behavioral Disorders Annual Conference. Salt Lake City, Utah. October 9-10, 2008.

SCERTS Training with Emily Rubin. Ogden, Utah. February 20-21,2007

Autism/Asperger's Conference. Salt Lake City, Utah. September 22, 2006.

Utah Conference on Special Education Law. Mountain Plains Regional Recourse Center. Ogden, Utah. August 2–3, 2006; Salt Lake City, Utah. August 3-4, 2005.

Utah Conference on Effective Practices in Special Education and Rehabilitation. Utah State University Department of Special Education and Rehabilitation. Logan, Utah. June 19-22, 2006; June 20-23, 2005.

Utah Conference on Curriculum. Severe Disabilities and Low-Functioning Autism Institute. Salt Lake City, Utah. March 10-11, 2005.

Assistive Technology with Emily Rubin. Severe Disabilities and Low-Functioning Autism Institute. Layton, Utah. February 9, 2005.

Severe Behavior with Melisa Genaux, Cindy Myers, and Rob O'Neil. Severe Disabilities and Low-Functioning Autism Institute. Salt Lake City, Utah. January 13-14, 2005.

Severe Behavior with Dr. Michael Powers. Severe Disabilities and Low-Functioning Autism Institute. Layton, Utah. November 30- December 1, 2004.

Classroom Organization and Management. Severe Disabilities and Low-Functioning Autism Institute. Layton, Utah. October 19-20, 2004.

Functional Communication with Dr. Andy Bondy. Severe Disabilities and Low-Functioning Autism Institute. Salt Lake City, Utah. September 20-21, 2004.

Research in Progress

Teacher Development and Training.

Professional Activities

Box Elder School District
District Service
Accreditation Coordinator—Bear River Middle School 2013-2014

Coordinate faculty/staff evaluation
Supervise collection of lines of evidence
Organize/Present Accreditation findings

Professional Development Presenter 2012-2014
Secondary Classroom Behavior Management
SpEd Education File Organization and Compliance
Autism ABC's
Autism Basics

Utah Personnel Development Center 2012-2013
Running Start Presenter/Organizer
Classroom Reinforcement Strategies
iPad's in the Classroom—Making the Most of What You Have
Video Modeling in the Classroom

Ogden City Schools
District Service
UBI Committee *2008-2009*
Develop ABC/UBI/PBS plan for district wide implementation
Supervise/Report on implementation of UBI sites
Supervise/Report on implementation of ABC/UBI site

PBS Committee *2008-2009*
Develop PBIS plan for district elementary level implementation
Train school coaches on elements of PBIS
Monitor Implementation of PBIS in all elementary setting district wide

UPIPS Steering Committee *2007-2009*
Develop program improvement plan
Organized Compliance Audits
Trained teacher for file compliance

Autism Team *2004-2009*
Conference Attendance:
SCERTS Training
Utah Autism/Aspberger's Conference
Utah Conference on Curriculum
Assistive Technology with Emily Rubin
Severe Behavior with Melisa Genaux, Cindy Myers, and Rob O'Neil
Severe Behavior with Dr. Michael Powers
Classroom Organization and Management
Functional Communication with Dr. Andy Bondy

Service

Evaluation of students with Autism and management of behavior in regular educational settings 04-05, 05-06, 06-07 school years
Provide training and education to teachers "Elements of Autism and ASD."

Teacher Mentor Training *2006-2009*

"How to be a Teacher Mentor." Paula Kashiwaeda. August 30, 2006
"Teacher Mentoring in Ogden City Schools." District Mentor Team. September 29, 2006

Professional Memberships

National Association of Elementary School Principals
Utah Association of Elementary School Principals
Council for Exceptional Children (CEC)
Box Elder Administrators Association
Box Elder Education Association
Utah Education Association
National Education Association