Spring 5-2010

AN EXAMINATION OF AN EARLY INTERVENTION READING PROGRAM FOCUSING ON THE PROGRESS MONITORING OF LITERACY SKILLS AND THE READING SELF-CONCEPTS OF STRUGGLING READERS

Teresa C. Samuelson
Western Connecticut State University, trsam91@gmail.com

AN EXAMINATION OF AN EARLY INTERVENTION READING PROGRAM FOCUSING ON THE PROGRESS MONITORING OF LITERACY SKILLS AND THE READING SELF-CONCEPTS OF STRUGGLING READERS

Teresa C. Samuelson

MS, Southern Connecticut State University, 1985
BS, Southern Connecticut State University, 1983

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Education in Instructional Leadership

in the

Department of Education and Educational Psychology

at

Western Connecticut State University

2010

AN EXAMINATION OF AN EARLY INTERVENTION READING PROGRAM FOCUSING ON THE PROGRESS MONITORING OF LITERACY SKILLS AND THE READING SELF-CONCEPTS OF STRUGGLING READERS

Teresa C. Samuelson, Ed.D.

Western Connecticut State University

Abstract

The purpose of this study was to examine progress monitoring, reading self-concept, and the literacy skills of first and second grade struggling readers. Progress monitoring is an instructional process used by teachers to assess students' academic performance on a regular basis, typically weekly or monthly. When based on the skill level of the student, the targeted remediation, and the goals of the intervention, progress monitoring may be used with various reading interventions. The use of progress monitoring is central to good decision-making in a Response to Intervention model.

Academic self-concept has become an integral part of education. Connections have been made regarding academic achievement and academic self-concept. Self-concept specifically of reading is vital in the primary years when the main focus of education is learning to read.

This study utilized a quasi-experimental research methodology as well as a correlational design. The sample size of 40 participants consisted of approximately 19 students in the experimental group and 21 students in the comparison group. All students in the experimental group participated in a reading support program with a Language Arts Consultant (LAC). The LAC's participated in training utilizing progress monitoring and

incorporating biweekly follow-up, specific to each individual student's daily interventions. Students met in groups of three, for 30 minutes, four to five times per week. All struggling readers in the comparison groups were seen in small groups for 30 minutes, four to five times a week. They were instructed by an Early Literacy Tutor (ELT) who had not been trained in and did not utilize progress monitoring.

A two-group multivariate analysis of variance (MANOVA) was conducted utilizing core reading words, core writing words, phonemes, and spelling, as the four dependent variables measuring literacy skills. These are four of the nine scores yielded from the *Dominie Reading and Writing Assessment Portfolio* (DeFord, 2004). The independent variable of reading support group consisted of two levels, progress monitoring and no progress monitoring. Results indicated no significant differences in group means of core reading words, core writing words, phonemes, and spelling.

A standard multiple regression procedure was conducted consisting of progress monitoring and reading self-concept as the predictor variables, and literacy skills, as measured by core reading words, as the criterion variable. The *Reading Self-Concept Scale* (Chapman & Tunmer, 1995a) was utilized to measure struggling readers reading self-concept. Results indicated no significance in progress monitoring and reading self-concept as predictors of students' literacy skills.

APPROVAL PAGE

WESTERN CONNECTICUT STATE UNIVERSITY

School of Professional Studies
Department of Education and Educational Psychology
Doctor of Education in Instructional Leadership

Doctor of Education Dissertation

AN EXAMINATION OF AN EARLY INTERVENTION READING PROGRAM FOCUSING ON THE PROGRESS MONITORING OF LITERACY SKILLS AND THE READING SELF-CONCEPTS OF STRUGGLING READERS

Presented by

Teresa Samuelson, Ed.D.

Marcia A. B. Delcourt, PhD	Marcia A. B. Delcourt	March 13, 2010
Primary Advisor	Signature	Date
Gary Rosato, EdD	Gary A. Rosato	March 13, 2010
Secondary Advisor Committee Member	Signature	Date
Dr. Jessica Devine, EdD	Jessica L Devine	March 13, 2010
Secondary Advisor Committee Member	Signature	Date

2010

ACKNOWLEDGEMENTS

This journey would have been impossible without the support and guidance of countless professional educators. Dr. Marcia Delcourt's professionalism and knowledge has been a guiding force for the past five years. Dr. Karen Burkes' patience and wisdom have been an inspiration. Special thanks to Dr. Nancy Heilbronner who's knowledge of statistics and patience guided my statistical analysis. I want to thank my secondary advisors Dr. Gary Rosato and Dr. Jessica Devine for their guidance and timely responses. I would like to thank my reader, Dr. Rima McGeehan who quickly read and responded with valuable insight. A very special thank-you to all the members of Cohort two who's dedication to education, support, laughter and knowledge was a guiding force throughout the doctoral program.

DEDICATION

I dedicate this dissertation to my family. To my parents and grandparents who have always believed in me. Their love and support through the years have been an inspiration for me. To my siblings who through the laughter and tears are always there for me. I dedicate this to the men in my life. My loving husband Rick for the endless driving the boys around, grocery shopping, and cooking he has done in the past five years. Without his support and patience I would not have been able to complete the doctoral program. To my guiding lights, my boys Matt, Erik, and Nick, for always being there filling my life with laughter and joy.

Table of Contents

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

Table of Contents (Continued)

Table of Contents (Continued)

Figures

CHAPTER ONE: INTRODUCTION AND IDENTIFICATION OF THE TOPIC

In 1997, Congress asked the Director of the National Institute of Child Health and Human Development (NICHD) to work in conjunction with the Secretary of Education to assemble a board of nationally recognized reading specialists (United States Department of Education, 2008b, p. 2). In consultation, they created the National Reading Panel (NRP). This panel consisted of 14 members selected from among nearly 300 persons who were nominated by a wide variety of public sources. The members of the NRP included prominent reading researchers, teachers, and child development experts, leaders in elementary and higher education, and parents. According to the NICHD Report of the subgroups (2000b), the panel was requested to "assess the status of research-based knowledge, including the effectiveness of various approaches to teaching children to read" (p. 1-1).

The panel reviewed and analyzed over 100,000 studies on reading. The NRP then identified five components essential to a child's learning to read: (a) phonemic awareness, (b) phonics, (c) vocabulary, (d) fluency, and (e) comprehension. One result of this panel's work was the finding that educators can focus too much on phonemic awareness, a student's proficiency in the ability to notice, think about, and work with the individual sounds in spoken words, and view this behavior as the only essential skill required to becoming a fluent reader. The panel found that "Phonemic awareness does not constitute a complete reading program. Rather, it provides children with essential foundational knowledge in the alphabetic system. It is one necessary instructional component within a complete and integrated reading program" (National Institute of Child Health and Human Development, 2000a, p. 8)

The panel also stated that:

> In implementing systematic phonics instruction, educators must keep the end in mind and ensure that children understand the purpose of learning letter sounds and that they are able to apply these skills accurately and fluently in their daily reading and writing activities. (NICHD, 2000a, p. 10)

Regarding fluency, the NRP found that it was not effective to use independent silent reading as the only type of reading instruction to develop fluency, especially with students who have not yet developed critical alphabetic and word reading skills. The panel found that it was most effective to teach a combination of reading comprehension techniques to children.

In January 2002, former President Bush signed into law the *No Child Left Behind Act* (NCLB) of 2001. NCLB redefines the federal role in K-12 education and is built on four basic principles: (a) stronger accountability for results in educational achievement, (b) increased flexibility and local control of funding, (c) expanded options for informed parental choice, and (d) an emphasis on teaching methods that have been proven to work through scientific research. Reading First was one of the "signature achievements" of NCLB (United States Department of Education, 2006). The Reading First initiative gave states both the funds and the tools they needed to eliminate the reading deficit in young children. The purpose of this subpart of NCLB was to provide assistance in establishing reading programs for students in kindergarten through grade 3 that were founded on scientifically based reading research, to ensure that every student can read at grade level or above not later than the end of grade 3, by the year 2014. "Reading First builds on the influential findings of the National Reading Panel and more than two decades of research which tracked tens of

thousands of students to determine how best children learn to read" (United States Department of Education, 2006, p. 2). The Reading First initiative identified instruction and assessment as being intricately linked. Four types of assessments were recognized that must be conducted at the classroom level to ensure that all students become successful readers. These four assessments are screening, diagnostics, progress monitoring, and outcome assessments (Hosp & Hosp, 2003; United States Department of Education, 2006).

NCLB also requires that all children in grades 3-8 be tested every year to demonstrate increased accountability for student performance and closing the achievement gap. Since the implementation of NCLB, every year thousands of students have taken standardized tests. Many of these students have failed these tests (Valencia & Buly, 2004). It is now up to individual states, school districts, and schools to create programs that support struggling readers in order for them to become successful in their reading ventures as they meet the requirements of NCLB.

As Valencia and Buly (2004) found, teachers need to go beyond looking at students' scores, they also must analyze students' work to identify their specific needs. Valencia and Buly compared a test score to a fever, believing it is a symptom that demands more specific analysis. Iaquinta (2006) stated that the critical element in teaching reading strategies is the skillful teaching that helps young readers learn the effective strategies they need to become independent readers. Those who are struggling readers must be identified and provided with early intervention. This consists of in-depth analysis of the strengths and needs of each struggling reader to guide instruction in the specific skills required for them to become successful readers.

Chapman and Tunmer (1995a) found that children in kindergarten and first grade made self-concept related differentiations within different domains (e.g. physical ability, peer relationships, reading, math) as well as across a range of domains (e.g. math, reading, music, sports). Reading self-concept is the most prominent of the school-related self-perceptions (Chapman & Tunmer, 1995a). The initial focus in education in the early years is learning to read. Reading is often considered the basis for success in these early years. This focus on reading achievement with these young children will certainly have its own impact on their self-concept as readers.

Rationale for Selecting the Topic

The importance of this topic stems from school districts being under pressure to effectively provide early intervention for struggling readers who need to meet with success in reading as early as possible. Allington (2001) stated, "Schools must enhance classroom instruction so that the number of struggling readers is minimized and then put into place an organizational strategy that ensures children who need intensive, expert instruction receive it" (p. 122). Allington also declared that such services will not necessarily be more expensive but cost less in the long run than the current less intensive interventions. Research showed that most children develop adequate reading skills (phonemic segmentation, self-monitoring, fluency decoding, narrative comprehension, etc.) in the daily routine of reading instruction. A small portion of students have failed to acquire these skills from their daily classroom instruction; therefore, school personnel need to ensure that: (a) classroom lessons focus on phonemic segmentation activities; (b) teachers identify struggling readers through an early warning system; and (c) reading specialists have an intervention plan for students by

the middle of first grade, which focuses on targeted skills with expert, intensive instruction (Allington, 2001).

Historically, Curriculum-Based-Measurement (CBM) began as a tool to monitor progress in basic academic skills to assist special education teachers in decision-making. CBM is designed to assist teachers in adapting their classroom and individualizing instruction to meet the needs of students in the classroom (Roehrig, et al., 2008). Olinghouse, Lambert, and Compton (2006) support the use of CBM as a scientifically validated progress monitoring instrument.

Progress monitoring refers to keeping track of children's academic development through regular data collection with valid and reliable measures. The material used for progress monitoring must be representative of the academic competence expected of students. These assessments must be sensitive to small changes in skills over a period of time for teachers to base their educational decisions and improve instructional effectiveness for individual students' instructional needs. The data are interpreted at regular intervals either weekly or monthly. Systematic progress monitoring can be used to raise teacher concern about students' progress and signal the need for additional or different forms of instruction. Changes to instruction are based on the interpretation of the child's progress (Fuchs & Fuchs, n.d.; Olinghouse, Lambert, & Compton, 2006).

With the signing of the Individuals with Disabilities Educational Improvement Act (IDEA) in 2004, the focus of progress monitoring shifted from the special education to the general education arena, as students with varying reading needs were included in all classrooms. According to McCook (2006), since the onset of the IDEA legislation, general education teachers have been required to monitor individualized interventions within their

classrooms. Daly and Murdoch (2000) believed that educators must have a deepened understanding of assessment and interventions to meet the demands of these struggling learners.

Fuchs and Fuchs (2000) found that there was sufficient research supporting the positive effects of CBM on academic achievement. Educators should be assured that the information yielded from CBM is significant. This, along with CBM's preciseness of monitoring progress of all students, will support and guide teachers in developing more successful instructional programs.

As part of IDEA 2004, along with Progress Monitoring, Response to Intervention (RtI) also became required in general education classrooms (McCook, 2006). RtI is a new, alternative method in identifying children with learning disabilities, which districts may use rather than the IQ-achievement discrepancy (Fuchs & Fuchs, 2006). The RtI model consists of three tiers of intervention that involve scientifically-based response to intervention. Tier 1 incorporates quality classroom instruction with research-based strategies targeting students using Benchmark assessments. Students who have not met benchmarks move to Tier 2, consisting of scientifically-based supplemental instruction in small, flexible homogenous groupings in addition to classroom instruction. A small percentage of students who continue to have marked difficulty in acquiring necessary skills shift to Tier 3, which incorporates specific intensive and explicit instruction prescribed by the Intervention team. McCook (2006) identified six critical components when developing and implementing a response to intervention model. These consist of: (a) a universal screening administered to all students three times a year, (b) identification of problem areas in measurable terms, (c) establishment

of baseline data, (d) development of an accountability plan, (e) design of a progress monitoring system, and (e) comparison of baseline data to results.

This researcher's contention is that through the use of progress monitoring, educators will be able to assess specific reading skills on a regular basis. Progress monitoring will assist educators in meeting the needs of all students by regularly probing specific skills to guide instruction and decision-making long before the individual student fails at high stakes testing. Therefore, the reason this research was conducted was to evaluate a specific early intervention reading program's use of an imbedded progress monitoring system.

Statement of the Problem

In September of 2007, the Nation's Report Card on closing the achievement gap in education showed record gains in working toward the goal of every child reading and doing math on grade-level by 2014. Reading scores for fourth graders were found to be higher than they had ever been in the history of the Nation's Report Card (United States Department of Education, 2007). Fourth graders who were previously left behind had demonstrated the highest rate of gains for lower-performing students since 2002 (United States Department of Education, 2007). These were precisely the learners for which NCLB was intended to have the greatest impact (United States Department of Education, 2007). In the time period from 2000 to 2007, the average reading scores for fourth grade students with disabilities improved by 23 points. At the same time, the average reading scores for limited-English proficient fourth graders improved by 21 points (United States Department of Education, 2008a).

The NAEP 2008 Trends in Academic Progress identified an increase in the average reading scores at all three ages, 9, 13 and 17, from 2004 to 2008 ($p < .05$). For 9-year-olds, the approximate age of students in the NAEP long-term trend assessments, the average

reading scores increased 4-12 points since 1971 ($p < .05$). Significant changes ($p < .05$) in scores were identified for 9-year-olds performing at the 10^{th}, 25^{th}, 50^{th}, and 75^{th} percentiles, while those performing at the 90^{th} percentile did not demonstrate a significant change. At age 13, lowest performing students at the 10th and 25th percentiles demonstrated significant gains in reading ($p < .05$). At age 17, students at the 10th, 25th, and 75th percentiles demonstrated significant gains in reading ($p < .05$). These results revealed that students were making more academic progress in reading than ever before. Through the guidance of NCLB setting high standards and holding school personnel accountable for results, the long-term trends show improvements in reading. Therefore, the core principles fueling the significant progress must be maintained for all students (United States Department of Education, 2007).

NCLB has required that states develop assessment plans so they will have data to support whether or not students are acquiring at least grade-level skills (No Child Left Behind, 2004). In 2002, the Statewide Longitudinal Data System Grant Program was authorized and assisted 27 states in developing and implementing improved data systems to review test score data to drive student improvement (United States Department of Education, 2008a).

Safer and Fleischman (2005) believed that research supported the enhancement of teacher decision making and student learning through the use of progress monitoring. Students' awareness of their own performance was evident. They also stated that student progress monitoring was administered quickly. A one-minute reading probe was just one method of gathering data. The results were understandable and easily communicated. This made progress monitoring easy to fit into the routine of the classroom.

While the reading achievement gap is narrowing, educators must continue to monitor the progress of all students. This will assist in identifying struggling readers as well as distinguish which interventions are successful. These struggling readers must be recognized as early as possible in their educational careers and be provided with early intervention, including progress monitoring, to guide their instruction in the specific skills required for them to become flourishing readers.

Potential Benefits of Research

The potential benefits of this research will be to determine if the use of progress monitoring of struggling readers will significantly improve their literacy skills. This will guide school districts in developing programs to assist all struggling readers within a Response to Intervention model (McCook, 2006). Through progress monitoring, it will become evident whether or not an intervention is working successfully. Teams regularly reconvene to assess student outcomes and determine when and if an intervention should continue, be modified, or stopped, or if another intervention should be implemented. McCook (2006) stated, "The lack of a documented progress monitoring process is fatal to an intervention and the support team process" (p. 67).

NCLB has assisted many districts in making their focus the development and implementation of effective programs in the primary grades to assist students in becoming thriving readers. It is imperative that these students be reached before they fail high stakes testing. Districts now must create effective programs that include student progress monitoring. For progress monitoring to be an effective factor for increasing student achievement, it must meet two criteria. First, it must be aligned with the curriculum content being learned. Second, feedback from progress monitoring must occur during the school

year, multiple times throughout the learning process. These influential factors are necessary to provide feedback that can be used to change instruction or performance. As a result, instructional changes will occur more often to meet the needs of students. Therefore, benefits of using progress monitoring are that students make greater gains in achievement, are more motivated to learn, and see themselves as successful learners (Hosp & Hosp, 2003). This research has significance in guiding school district personnel in developing programs to assist struggling readers in the primary grades within a Response to Intervention model.

Definition of Key Terms

1. *Attitudes toward reading* refer to the "affective component of reading self-concept, which is defined in terms of feelings toward and affinity for reading" (Chapman & Tunmer, 1995b, p. 154).
2. *Core Reading Words* are assessed when students are asked to read a list of high-frequency words (DeFord, 2004).
3. *Core Writing Words* are assessed when students are asked to write their names and all the words that they can within a 10-minute period (DeFord, 2004).
4. *Curriculum-Based Measurement* is the direct and continuous evaluation of student progress toward specific instructional objectives for the purposes of determining appropriate instructional objectives (Roehrig et al., 2008).
5. *Diagnostic assessment* provides an in-depth analysis of student's strengths and weaknesses to guide instruction (Hosp & Hosp, 2003).
6. The *Dominie Reading and Writing Assessment Portfolio* is a comprehensive literacy assessment program (DeFord, 2004).

7. The *Developmental Reading Assessment, 2nd Edition* (DRA2) is a criterion-referenced reading assessment for students in kindergarten through Grade 8 (Beaver, 2005).
8. An *Early Literacy Tutor (ELT)* in this study is required to hold a K-6 teaching certification with extensive, demonstrated experience in the area of language arts.
9. *Fluency* is "the ability to read a text accurately and quickly" (Armbruster, Lehr, & Osborn, 2003, p. 22).
10. A *Grapheme* is "the smallest part of written language that represents a phoneme in the spelling of a word" (Armbruster, Lehr, & Osborn, 2003, p. 4).
11. A *Language Arts Consultant (LAC)* in this study is required to hold a Reading and Language Arts Consultant, K-12 certification (097-endorsement in Connecticut).
12. *Leveled Literacy Intervention (LLI)* is a small group, supplementary intervention designed for children who find reading and writing difficult. LLI includes initial and ongoing assessments, progress-monitoring, and record-keeping instruments (Fountas & Pinnell, 2009).
13. *Literacy Skills* in this study will be measured utilizing the *Dominie Reading and Writing Assessment Portfolio* subscales of core reading words, core writing words, sentence writing and spelling (DeFord, 2004).
14. An *Onset* is "the initial consonant(s) sound of a syllable" (Armbruster, Lehr, & Osborn, 2003, p. 4).
15. *Oral Reading and Fluency* are assessed when a student must read accurately and fluently from passages of a new text. Guides are provided for observing and assessing the oral reading behavior of a student (DeFord, 2004).

16. An *Outcome assessment* can determine whether a student achieved grade-level performance or not (Hosp & Hosp, 2003).
17. *Perceptions of competence in reading* refer to "beliefs regarding ability and proficiency in reading tasks" (Chapman & Tunmer, 1995b, p. 154).
18. *Perceptions of difficulty in reading* refer to "beliefs that reading activities are hard or problematic" (Chapman & Tunmer, 1995b, p. 154).
19. *Phoneme* is "the smallest part of spoken language that makes a difference in the meaning of words" (Armbruster, Lehr, & Osborn, 2003, p. 4).
20. *Phonemic awareness* is "the ability to hear, identify, and manipulate the individual sounds-phonemes-in spoken words" (Armbruster, Lehr, & Osborn, 2003, p. 4).
21. *Phonics* is "the understanding that there is a predictable relationship between phonemes (the sounds of spoken language) and graphemes (the letters and spellings that represent those sounds in written language)" (Armbruster, Lehr, & Osborn, 2003, p. 4).
22. *Phonological Awareness* is "a broad term that includes phonemic awareness. In addition to phonemes, phonological awareness activities can involve work with rhymes, words, syllables, and onsets and rimes" (Armbruster, Lehr, & Osborn, 2003, p. 4).
23. *Progress Monitoring* refers to keeping track of children's academic development through regular data collection with valid and reliable measures. The data are interpreted at regular intervals. For the purpose of this study, data will be

collected weekly and reviewed biweekly. Changes to instruction are based on the interpretation of the child's progress (Fuchs & Fuchs, n.d.).

24. A *Reading Support Program* in this study refers to a reading intervention involving reading instruction using a one on one or small group meeting, four to five times per week, in addition to the classroom reading curriculum instruction.
25. *Response to Intervention* is a multi-tiered method of service delivery in which all students are provided an appropriate level of evidence-based instruction focused on their academic needs (Fuchs & Fuchs, 2006).
26. A *Rime* is "the part of a syllable that contains the vowel and all that follows it" (Armbruster, Lehr, & Osborn, 2003, p. 4).
27. *Screening assessment* is a quick measure that focuses on critical reading skills and can be used to distinguish which students are likely to need additional assistance (Hosp & Hosp, 2003).
28. *Segmentation* occurs "when words are broken into their individual phonemes" (Armbruster, Lehr, & Osborn, 2003, p. 6).
29. *Self-Concept* is the idea or mental image one has of oneself and one's strengths and weaknesses (Shavelson, Hubner, & Stanton, 1976).
30. *Sentence Writing and Spelling* is assessed when the examiner reads a sentence to the student and then repeats each word in the sentence one at a time, instructing the student to say the words slowly and write them (DeFord, 2004).
31. A *struggling reader* in this study is identified as a first or second grade student who performs in the bottom 10% of the readers in the grade level of each school, based on the *DRA2* and *Dominie*.

Research Questions and Hypotheses

1. Is there a significant difference in literacy skills (core reading words, core writing words, phonemes, and spelling) between struggling readers who are part of a reading support program with an imbedded progress monitoring system and those who participate in a reading support program without an imbedded progress monitoring system?

   Directional Hypothesis: Students who participate in a reading support program with progress monitoring will demonstrate significantly higher mean scores on literacy skills (core reading words, core writing words, phonemes, and spelling) as compared to those who have participated in a reading support program without progress monitoring.

2. To what degree and in what manner do progress monitoring and reading self-concept predict literacy skills (core reading words)?

   Non-Directional Hypothesis: Literacy skills (core reading words) are significantly impacted by a reading support program (Progress Monitoring or no Progress Monitoring), and reading self-concept (difficulty in reading, competence in reading, attitudes toward reading).

CHAPTER TWO: RELATED LITERATURE

This review of the literature is presented in five sections. The first section is focused on Lev S. Vygotsky's theoretical construct of the *Zone of proximal development.* He believed that individuals followed examples and developed the ability to do certain tasks with or without assistance. The remaining four sections are recent appraisals of literature pertinent to the topic. The initial focus is on studies involving how to best identify and meet the needs of struggling readers. The next section's focal point is a review of literature on how progress monitoring informs literacy instruction. This is followed by a description of the effectiveness of the response to intervention model for students with reading difficulties. The final section consists of a review of literature on the development of reading self-concept in young children.

The Theoretical Foundation of Lev Vygotsky's Zone of Proximal Development

Lev S. Vygotsky studied the development of cognitive processes. He believed that "individual development could not be understood without reference to the social and cultural context within which such development is embedded" (Driscoll, 2005, p.247). The focal point for Vygotsky was that there were developmental stages within the cognitive processes. Vygotsky believed that instruction could precede and contribute to development of the cognitive processes.

Vygotsky (1978) termed a phrase "zone of proximal development," which initially determined that there were at least two developmental levels. He called the difference between what a child can do with help and what he or she can do without guidance the "zone of proximal development." The first level refers to the mental functions that have been established due to already completed cycles determined by independent problem solving.

This "defines functions that have already matured, that is, the end products of development" (Vygotsky, 1978, p. 86).

The second level, potential development, consists of functions that have not yet matured but may be in the process of maturation. The zone of proximal development separates actual from potential development as determined through problem solving under adult guidance or in collaboration with peers (Vygotsky, 1978).

According to Driscoll (2005), "Vygotsky viewed the processes of learning and development to be separate, in that learning is not the same thing as development, but linked, in that learning can set developmental processes in motion" (p. 255). Vygotsky believed that a teacher's goal is to identify learning tasks in the zone of proximal development and gradually reduce his or her assistance with these tasks until the student is capable of independently performing the task, therefore, attaining his or her actual developmental level.

Vygotsky's theory supports that each and every student develops at his or her own pace. As Vygotsky (1978) stated, "Although learning is directly related to the course of child development, the two are never accomplished in equal measure or in parallel" (p. 91). Keeping this in mind, students also learn in different ways. One program of instruction may not assist each student in the class to master the material presented. This also supports the practice of progress monitoring since it addresses individual needs and may occur in a group measuring a variety of skills.

According to Vygotsky (1978), "A crucial aspect of human mastery, beginning in infancy, is the creation and use of auxiliary or "artificial" stimuli: through such stimuli an immediate situation and the reaction linked to it are altered by active human intervention" (p. 123). The use of progress monitoring will assist in identifying how to guide and support

struggling readers through their zone of proximal development to their actual development as successful and confident readers.

Meeting the Needs of Struggling Readers

This research incorporates the first two components identified by the National Reading Panel as essential to a child's learning to read: phonemic awareness and phonics. The following studies support the inclusion of each of these components in a progress monitoring program. Research on early reading instructional approaches, which incorporate phonemic awareness and phonics, when working with struggling readers, is also reviewed.

Explicit Instruction with Struggling Readers

The following four studies review the use of explicit instruction in reading interventions with struggling readers. These studies examine the use of explicit instruction of orthographic units, phonologic connections, and letter/sound relationships. They support the use of progress monitoring in assisting to identify which of these varied instructional models are successful for each individual student.

Levy and Lysynchuk (1997) developed a study with an interest in finding the orthographic units that led to the fastest acquisition, best retention, and best generalization to words with similar orthographic units. Their study, which included two experiments, explored the relative efficiency of several segmentation methods and of a whole word repetition method for different samples of children acquiring initial reading vocabulary. The first experiment utilized nonreaders in kindergarten and first grade, while the second experiment included second grade students who were already reading below grade-level.

In Experiment 1, all children were given the Word Identification subtest of the *Woodcock Reading Mastery Test* (*WRMT*; Woodcock, 1987) and of the *Wide Range*

Achievement Test-Revised (*WRAT-R*; Jastak & Wilkinson, 1984). Students who read fewer than seven words on any of these two screening tests were identified to participate in this study. The sample included 83 children in first grade and 17 children in kindergarten, selected from all students in two schools in Canada. All 100 children (48 boys and 52 girls) were randomly assigned to four training groups and one control group, with each group consisting of 20 children.

In the pretest phase, letter knowledge and phonemic sensitivity were measured. This was followed by the training phase when four groups of children were trained to read a set of words, using four different methods. The post training phase included tests of retention of trained words and generalization to untrained words and nonwords. This pretest-training-posttest design consisted of two further standardized measures including four reading-related subtests of the *Wechsler Intelligence Scale for Children-Revised* (*WISC-R*; Wechsler, 1991; vocabulary, similarities, picture arrangement, block design) and the *Peabody Picture Vocabulary Test* (*PPVT*; Dunn & Dunn, 1981).

The control group received only their regular classroom regime during the training phase. This consisted of a whole language orientation with incidental rather than systematic instruction on segmentation. The four training groups all learned to read the same set of 32 words, as well as participated in the classroom program. Each group varied in how the words were grouped during instruction and in method of instruction of the 32 words: (a) rime, (b) onset, (c) phoneme, or (d) whole word training. All training methods included 25 trials broken into 15 color trials or a set criterion of two consecutive correct readings, then 10 black and white trials or the set criterion of two consecutive correct readings. If a child reads

words successfully two times in succession prior to the prescribed number of trials, he or she moved on to the next set of words.

The words in the rime training condition arranged the 32 words so that the four words of a rime family were shown together. The onset training organized the 32 words in groups of four, which shared the initial consonant(s)-vowel segment. In the phoneme training group, four words were randomly selected from the 32 words, except that no two onset or rime family members occurred in the same group. For the whole word training, four words were randomly selected, with the restriction that no family members occurred in the same group and the same eight random words were read each day of training.

The posttest phase occurred for all children following their sessions. First, they were given a generalization test. Children attempted to read 48 new words then 48 nonwords, printed in block ink. Half of each of these sets began with an onset and the other half with a rime. The 48 words were each randomly arranged for all groups, rather than presented in family groups.

For the short-term retention test, children were seen one week following the end of their training. They again pronounced each of the 32 training words arranged randomly and not presented in family groups. The long-term retention test occurred four to six months after training. At this time, the words were again randomly arranged and printed in black ink for all groups.

To minimize the number of contrasts, the onset and rime groups were combined to form one unit (Onset/Rime) since the training for each consisted of segmenting into larger family units differing only in whether the units occurred in the rime or onset of the word. This larger Onset/Rime group was compared with both the Phoneme group and Whole Word

group for a total of three training/treatment groups. The Onset/Rime demonstrated significantly higher mean scores compared to the Whole Word, $t(76) = 4.3$, $p < .01$, as well as when compared with the Phoneme group, $t(76)$, $= 1.5$, $p < .05$. A third comparison demonstrated that the Phoneme group was significantly different from the Whole Word group, $t(76) = 2.5$, $p < .05$, with the Phoneme group having a higher mean. The three segmentation training conditions (onset, rime, and phoneme) led to faster acquisition of the trained reading vocabulary over 15 trials as compared to students in the Whole Word group.

Data were analyzed for short-term retention ($n = 97$). A comparison of the Onset/Rime and the Phoneme groups showed no significant difference due to the orthographic unit size presented in the training. The three groups (Onset, Rime, & Phoneme) were then combined into one "segmentation" group demonstrating significantly higher means than both the Whole Word, $t(92) = 3.0$, $p < .05$, and the Control, $t(92) = 10.4$, $p < .01$ groups. The Whole Word group was superior to the Control group, $t(92) = 6.3$, $p < .01$. Therefore, all methods lead to good short-term retention, though the best short-term retention followed the segmentation training (Onset, Rime, & Phoneme groups combined).

Long-term data analysis ($n = 87$) occurred after 4 to 6 months. A difference among the five groups was found, $F(4, 82) = 7.17$, $MSE = 80.45$, $p < .001$. Segmentation ($t(82) = 5.2$, $p < .01$) and whole word training ($t(82) = 3.1$, $p < .01$) conditions were better than the control group receiving only their regular classroom instruction. No differences among the four training conditions were indicated utilizing the same four linear contrasts. The researchers stated, "The main result is that over the long term, retention of training was good irrespective of method used during training" (p. 372).

Degree of retention was also compared for the 11 children in the Whole Word group who met the criterion of reading the thirty two trained words, correctly on 2 successive days (learners) in comparison to the 9 children in the same group who did not meet the same criterion (non-learners). Here the importance of complete learning until mastery during the colored training phase prior to changing to the black training phase was evident. In short-term retention (one week following the end of training), students who had met the criterion (learners) showed retention of $M = .89$ in comparison to $M = .43$ ($t(18) = 3.98$, $p < .01$) for those who had not met the criterion (non-learners). In the long-term test (four to six months following the end of training) with $n = 9$ for the learners who met the criterion, and $n = 8$ for non-learners who had not met the criterion, retention was $M = .87$ versus $M = .50$, $t(15) = 3.87$, $p < .01$.

The effect of learning on generalization was also analyzed within the Whole Word group looking at children who met criterion versus those who did not learn the set completely. For the word test, the mean generalization was .44 for learners who had met the criterion and .07 for non-learners who had not met the criterion, $t(14.5) = 4.14$, $p < .01$. For the 48 nonwords test that were read correctly during generalization, mean performance was .61 for the learners and .13 for the non-learners, $t(18) = 4.23$, $p < .01$. Therefore, when training is carried to full learning of the training items, the generalization to new words and nonwords was 40% to 65% mastery of the words on the first encounter with the new items, irrespective of the instruction used in acquisition. The method of instruction differences was observed only when full learning was not evident. In conclusion, when students were given the time to meet a criterion during learning, their retention and generalization were excellent, regardless of the method of instruction used in acquisition. This conclusion supports the

importance of utilizing reading support programs with imbedded progress monitoring to closely monitor and allow a student time to meet a specific criterion.

In Experiment 2, Levy and Lysynchuk (1997) questioned whether children who experienced reading difficulties early in their school careers acquired reading vocabulary more rapidly via whole word repetition, rather than with a method that required them to understand and use subword units. The sample consisted of 125 children in second grade classrooms from 16 different schools in the same school system. Students were identified through the administration of the same screening assessments as those in Experiment 1 and randomly assigned to each of the five training conditions. These five training conditions included: (a) control, (b) rime, (c) onset, (d) phoneme, or (e) whole word training groups. The pretest phase was shortened to contain two letter knowledge tests and two of the phonemic sensitivity tests.

In addition, a spelling test and a continuous digit naming task were added to the pretest battery. Training for the four trained groups remained exactly as in Experiment 1. A new set of thirty two words, appropriate to the Grade 2 level, was created as the training word set to be learned. The posttest phase began as soon as the individual child reached the end of the learning phase and tested for retention after one week. Long-term retention could not be tested due to the large number of schools involved in this study.

Following trials 1 to 5, the larger orthographic units of onset/rime combined demonstrated significance in comparison with whole word training, $t(97) = 4.0$, $p < .01$. Following trials 6 to 10, the onset/rime group also demonstrated significance compared to the whole word training group, $t(97) = 3.2$, $p < .01$. By trials 11-15, the final trials, most of the children in all groups had reached the criterion and there were no significant group

differences. Therefore, those who were taught the larger orthographic units (onset/rime combined) outperformed the whole word learners following trials 1 to 10.

Regarding retention, an ANOVA comparing all five groups indicated a significant effect of condition, $F(4, 12) = 17.73$, $MSE = 47.42$, $p < .001$. All four treatment groups demonstrated significance in comparison to the control group ($p < .01$ in all cases).

An ANOVA showed a significant effect for the conditions, $F(4, 120) = 6.19$, $MSE = 52.05$, $p < .02$. All students in training groups read more new words than those in the control group $p < .01$ for all cases. The superiority of all training groups compared with the control group indicated that the training experiences resulted in knowledge gains that could be generalized to the reading of new words.

The educational implications from this research showed that segmentation leads to more efficient acquisition of reading vocabulary than whole word repetition for normal beginning readers and for young delayed readers. An important conclusion points out that the degree of learning (meeting the learning criterion of mastery for two consecutive days) rather than the method of instruction will ensure strong retention and generalization. The use of a reading support program with imbedded progress monitoring will assist in documenting which segmentation of letters a student has mastered to ensure retention and generalization. It will also assist in making decisions regarding the next instructional step for each student.

Soon after this research was completed, Foorman, Fletcher, Francis, Mehta & Schatschneider (1998) hypothesized that over one school year, children who received explicit instruction in the alphabetic principle with an emphasis on letter-sound correspondence would show greater growth than children who received less explicit instruction on spelling patterns or children receiving implicit instruction in alphabetic principle. According to

Foorman et al., the alphabetic principle means "…letters in a word relate to speech in a conventional and intentional way" (p. 37).

Their study included 285 participants in first and second grade all of whom were eligible under Title 1 funding. Title 1 refers to federal funding provided for economically disadvantaged children with low achievement. The urban district, which included 19 elementary schools, defined low achievement as scores in the bottom quartile of the district's emergent literacy survey. The sample included 61% males. The ethnic composition of the sample was: 60% African American, 20% Hispanic, and 20% White. The instructional groups did not differ in age, gender or ethnicity (Foorman et al., 1998).

Students were instructed in one of three classrooms during a 90-minute daily language arts period. All classrooms incorporated a literature-rich environment. In the three classrooms, instructional methods included: Direct Code (DC), Embedded Code (EC) and Implicit Code (IC), which included use of the district standard curriculum (IC-S) or a research implementation (IC-R) developed to ensure comparability of training across instructional approaches.

Direct Instruction (DC) involved letter-sound correspondences practiced in decodable text. The emphasis was on a balance of phonemic awareness, phonics, and literature activities. Embedded Code (EC) included less direct instruction in systematic spelling patterns (onset rimes) embedded in connected text. In EC, the emphasis was on phonemic awareness and spelling patterns in predictable books. Implicit Code (IC) incorporated indirect, incidental instruction in the alphabetic code embedded in connected text. The school district also emphasized the IC approach to reading instruction. The IC strategy included two options. One option called the district standard curriculum (IC-S) required that

teachers were trained and supervised by district personnel. The IC-R option was research implemented with teachers by the project director, an experienced doctoral –level teacher-trainer. Each model was directed by an advanced graduate student who had been a teacher and who had expertise in professional development and did not include the authors of the study.

Students were assessed four times during the year to measure changes in vocabulary, phonological processing, and word-reading skills. *The Peabody Picture Vocabulary Test-Revised (PPVT-R*; Dunn & Dunn, 1981) was utilized to assess growth in receptive language. A 50-word list was utilized to assess changes in reading skills. This list was representative of a diversity of linguistic features, and spanned the levels of difficulty from first to third grade. Students received scores based on the number of words they read aloud correctly out of the 50. An internal consistency estimate of .90 demonstrated the word list had excellent reliability. The word list also demonstrated high concurrent and predictive validities as evidenced by correlations exceeding .80 with the Letter and Word Attack subtests of the *Woodcock-Johnson Psycho-Educational Battery-Revised* (*WJ-R*; Woodcock & Johnson, 1989). The *Torgesen-Wagner Battery* (as cited in Foorman et al., 1998, p.41), synthesis and analysis tests were administered to measure phonological processing.

At the beginning of the year, a standardized reading assessment was not administered because "tests such as the WJ-R lack a sufficient number of items to discriminate initial reading levels for beginning readers and are not adequately sensitive to change over short time intervals" (p. 41). The *Wechsler Intelligence Scale for Children-Revised* (*WISC-R*; Wechsler, 1974) was individually administered at the end of the year along with standardized reading and spelling tests. The Letter-Word Identification, Word Attack, and Passage

Comprehension subtests of the *WJ-R* were administered to measure decoding and reading comprehension, respectively. A formal reading inventory was used to measure comprehension of narrative and expository text. The Spelling Dictation subtest from the *Kaufman Test of Educational Achievement* (*KTEA*; Kaufman & Kaufman, 1985) was administered to measure spelling.

The researchers acknowledged that in all three instructional groups, children with higher initial scores in phonological processing skills in October (M = .37 to .68 in Grade 1 and M = 1.38 to 1.74 in Grade 2) exhibited growth in word-reading skills. Though, students who had low initial status in phonological processing in the DC group (direct explicit instruction) appeared to show more growth in word-reading skills than similar students in the other instructional groups. Students receiving DC had significantly higher scores in April than those in the EC (t = 2.99, p < .003, f = 1.06) and IC-R groups (t = 4.58, p < .001, f = 1.61). Therefore, the researchers believed that direct explicit instruction (DC group) using phonemic awareness appeared to facilitate word-reading development for children who started the year with low scores. The results of this study revealed that not all instructional approaches had the same impact.

End-of-year achievement after the first year clearly demonstrated differences in the instructional groups. The IC-R group's means at the 29th percentile (decoding) and the 35th percentile (passage comprehension) were well below the national average. The EC group means were also well below the national average with scores at the 27th percentile (decoding) and 33rd percentile (passage comprehension). The DC group's 43rd percentile and 45th percentile values approached the national average in comparison to the IC-R and EC groups. The Direct code group had higher mean decoding scores than the Embedded code group

($F(1, 197) = 9.41$, $p = .003$, $f = 1.17$). The DC group had significantly higher scores than the Implicit code research group ($F(1, 197) = 7.00$, $p = .009$, $f = 1.22$). The DC group also had higher mean passage comprehension scores than both the EC group, ($F(1,197) = 4.76$, $p = .030$, $f = 0.72$) and the IC-R group ($F(1, 197) = 3.68$, $p = .056$, $f = 0.76$), though neither were significant at the Bonferroni adjusted alpha level of .0167.

While this research indicated that early instructional intervention made a difference for the development and outcomes of reading skills in first and second grade children at-risk for reading failure, it also demonstrated that not all instructional approaches had the same impact. This supports the concept that early instruction which incorporates the explicit teaching of the alphabetic principle may prevent reading failure. Monitoring progress within a reading support program will assist in identifying whether or not an instructional approach is successful for a specific student. This may guide the educators involved to identify which instructional approach may be more successful in supporting the student to become a thriving reader.

In the following study, Blachman et al. (2004) evaluated the effectiveness of an intensive reading intervention for second and third grade children who were identified with reading disabilities with poor word-level skills. Their goals included the evaluation of the intervention, and monitoring of student progress for one year after the intervention. They investigated whether gains were maintained and determined which areas of reading and spelling, if any, demonstrated long-term gains.

The sample of participants was selected from 11 schools in four school districts in upstate New York. This included two cohorts of children, one identified in the spring of 1997 and the second in the spring of 1998. The first phase of screening requested first and

second grade teachers in participating schools to identify the lowest 20% of readers in their classrooms, excluding children who were: (a) left-handed, (b) had a hearing loss, (c) severe articulation problems, (d) severe emotional disturbance, (e) autism, (f) mental retardation, (g) neurological problems, or (h) learning English as a second language.

The second phase included the screening of 295 children. Students who obtained a standard score below 90 (25th percentile rank) on either the Word Identification or the Word Attack subtests of the *Woodcock Reading Mastery Tests—Revised* (*WRMT-R*; Woodcock, 1987) and who also obtained a standard score below 90 (25th percentile rank) on the Basic Skills Cluster of the WRMT-R met the reading criteria to be in this study and were given the *Wechsler Intelligence Scale for Children-Third Edition* (*WISC-III*; Wechsler, 1991). To be considered eligible for the study, students had to meet the reading criteria and have a Verbal IQ score of at least 80. Students began the study with word-level skills below the 25th percentile on the standardized test.

Eligible for the study were 100 children (65 male, 35 female). To bring the proportion of girls closer to the proportion of boys, 11 boys were randomly eliminated from the eligibility pool. The final participating sample included 69 students (42 males, 27 female) randomly assigned within schools, grade, and gender to treatment and control groups. The treatment consisted of one-to-one tutoring, five days per week including explicit and systematic instruction of the phonological and orthographic connections in words. The diversity of these students included 80% White, 14% African American, below 1% Hispanic, and below 1% other.

All participating students were given two batteries of tests. One battery consisted of standardized (norm-referenced) testing, which included measures of reading, spelling, and

math. Children completed these three times: (a) as pretests prior to the treatment; (b) as posttests at the end of the treatment; and (c) again, one year after the posttest, at follow-up. Math measures were included to determine whether the intervention actually targeted reading. The second battery was referred to as a skills battery. This was repeated every 10 weeks or eight times throughout the treatment year and follow-up year. These non-standardized measures included phonological processing, word reading accuracy, word reading efficiency, and spelling. The order of administration of tests in both batteries was fixed across all administrations. All educators providing testing, were extensively trained in administration and scoring, retrained before each wave of testing, and had no knowledge of the condition of the children.

Treatment children received 50 minutes of one-to-one tutoring, five days per week for the treatment year, in addition to regular classroom reading instruction. All tutoring was provided by 12 teachers certified in either reading or special education. The tutoring program included explicit and systematic instruction to help children develop an understanding of the phonologic and orthographic connections in words, as well as many opportunities to read texts that were phonetically controlled and texts that were not phonetically controlled (both narrative and expository texts) to develop fluency, build comprehension strategies, and foster reading for pleasure.

Each lesson included five steps though the program was not scripted and each lesson was individualized based on the child's progress along with feedback provided to the tutor following observations. The five steps included: (a) quick-paced review of sound symbol associations learned in previous lessons and the introduction of new sound-symbol correspondences, (b) manipulation of letter cards to practice phoneme analysis and blending,

(c) fluency building activity, (d) oral reading practice of phonetically controlled text, and (e) dictation of words used in earlier steps of the lesson. To assist in fidelity each child was observed an average of nine times during the treatment year.

The control group received whatever remedial reading instruction was provided by the school, in addition to their classroom reading instruction. These students met in small groups, three to five times per week for an average of 41 minutes per session. Students participated in an average of 104 sessions. All teachers in the control groups were certified in either reading or special education and all had a master's degree in one of these areas.

The progress of all children was monitored for one year following the treatment. The skills battery was administered four times during the school year, in September, December, March, and June. In June, the original pretest battery was also administered again. During the follow-up year, all children received regular classroom reading instruction. In some cases, they also received remedial reading instruction outside the classroom. Throughout the follow-up year, reading instruction outside the classroom occurred for 51% of the treatment children and 63% of the control children. Data continued to be collected from the classroom and remedial reading teachers.

There were no differences between treatment and control groups at pretest for any of the standardized measures. Results at posttest of the standardized measures indicated significant differences between the treatment and control groups on the two reading and one spelling measure. Effect sizes ranged from 0.55 for the *Gray Oral Reading Test—Third Edition* (*GORT-3*; Wiederholt & Bryant, 1992) Comprehension to 1.69 for the *WRMT-R* Basic Skills Cluster.

Analysis of nonstandardized measures of pretest scores revealed that the groups did differ significantly on nonword repetition, favoring the treatment group, $t(67) = 2.07$, $p = .0425$. At pretest there were no other significant differences between groups on the other five measures. At the end of the treatment year, there were significant differences on Word Reading, $t(416) = 3.70$, $p \leq .01$; Word Reading Efficiency, $t(416) = 4.87$, $p \leq .01$; and Spelling, $t(416) = 3.79$, $p \leq .01$, favoring the treatment group.

Results for the follow-up year revealed a similar pattern of findings though some of the differences that were significant at posttest were no longer significant at follow-up. At follow-up all of the measures of reading and spelling remained significant except for two subtests of the GORT-3, accuracy and comprehension. At the end of the follow-up year, significant differences between the groups remained on all three measures: Word Reading, $t(416) = 2.72$, $p \leq .01$; Word Reading Efficiency, $t(416) = 4.24$, $p \leq .01$; and Spelling, $t(416) = 2.75$, $p \leq .01$, all favoring the treatment group. While differential growth rates during the treatment year were significant favoring the treatment group, they were not significantly different during the follow-up year. The treatment students maintained their gains, though their rate of growth during the follow-up year did not differentially increase or decrease relative to the control children.

The three phonological processing or skill measures of phonological awareness, rapid naming of letters, and nonword repetition, were significantly different on two of the three measures at the end of the treatment year. There were differences on phonological awareness, $t(416) = 2.22$, $p \leq .05$, and rapid naming of letters, $t(416) = 2.15$, $p \leq .05$, favoring the treatment group. Growth rates during the treatment year were also significantly different for phonological awareness, $t(416) = 2.06$, $p \leq .05$. At the end of the follow-up year, a

significant difference only remained for rapid naming of letters, $t(416) = 2.27$, $p \leq .05$. All these results favored the treatment group. There no longer was a significant difference on phonological awareness. The rates of growth during the follow-up year were not significantly different on any of the three phonological processing measures.

Researchers concluded that the treatment students "who participated in an intensive, systematic, and explicit program that emphasized the phonologic and orthographic connections in words and text-based reading, showed significantly greater gains than the control children on measures of both real word and nonword reading, reading rate, passage reading and spelling, and for the most part, maintained these gains at a 1-year follow-up" (p.454). Reading support programs with imbedded progress monitoring systems are by nature systematic and explicit. Progress on phonologic and orthographic connections in words and text may be monitored for each student.

To further investigate reading interventions, Christensen and Bowey (2005) conducted a study that compared the efficacy of two decoding skill-based programs to a control group exposed to an implicit phonics program. The skill-based programs consisted of one based on explicit orthographic rime (OR) and one on grapheme-phoneme correspondences (GPC). Their study included 116 children in their second year of schooling from 7 classrooms in two elementary schools in Australia. The second year of schooling in Australia was equivalent to the first grade in the United States. The average age was 7 years and 1.5 months at the commencement of the study. It was stated that "they were slightly older than children in their second year of schooling in systems that have a kindergarten year (Grade 1 in Queensland). Thus, they could be considered as advanced beginning readers" (p. 330).

All three programs were administered by research assistants, who were given professional development, a teacher's manual, and a simulation. Programs were implemented for 20 minutes per day for 14 weeks. There were no significant differences at the time of the pretest among the three instructional groups using a multivariate analysis of variance (MANOVA).

All study participants were given a pretest, two interim tests and a posttest. The pretest assessed: (a) letter-sound knowledge, (b) decoding, (c) sight word recognition, (d) phonemic segmentation of words, and (e) spelling. To assess letter-sound knowledge, participants were initially asked to provide the sounds of the 26 letters of the alphabet. Pretest measures also included reading ability utilizing a modified version of Clays' (1979) *Ready-to-Read Words Test* (as cited in Christensen & Bowey, 2005). The modified version omitted the items *a* and *I*, as children can read these on the basis of letter-name knowledge alone. Children were asked to identify each of the phonemes in 18 words. They were also asked to spell 24 words to assess spelling ability.

Interim test 1 included the reading and spelling of consonant-vowel-consonant (CVC) words with orthographic rimes covered in all three intervention programs. Interim test 2 also required students to read and spell words, though the words contained initial and final consonant clusters.

The posttest also assessed reading and spelling. Children were asked individually to read 55 program words (words that had been directly taught) and 55 transfer words (words with similar orthographic structures that were not directly taught) to assess their decoding skills. Spelling was assessed as a whole-class group giving students 20 program words and 20 transfer words. Oral reading was assessed using the *Salford Sentence Reading Test*

(Bookbinder, 2002). Reading comprehension was assessed utilizing a story from the Basic Academic Skills Samples.

When a MANOVA was employed, posttest results indicated that there was a significant effect for group, F (2, 182) = 8.71 $p < .001$. Follow-up comparisons utilizing Tukey's procedure (α=.05) revealed significant differences on all measures. The grapheme-phoneme correspondences (GPC) group in comparison to the implicit phonics group demonstrated statistical significance on all word-level reading and spelling measures and on both the Basic Academic Skill Samples and the *Salford Sentence Reading Test* ($p \leq .01$). The orthographic-rime (OR) group in comparison to the implicit phonics group demonstrated statistical significance on the accuracy with which they read both program and transfer words ($p \leq .001$). The OR group in comparison to the implicit phonics group demonstrated statistical significance when spelling program words with silent e ($p \leq .001$). The implicit phonics group in comparison to the OR group demonstrated statistical significance when spelling program words containing vowel digraphs and transfer words containing silent e ($p \leq .001$). No other between-group contrasts were significant.

Reading extended text is an essential component of a comprehensive literacy program. The researchers believed the data strongly suggested that systematic instruction and practice in decoding words, based on symbol-sound relationships, enhanced children's reading rather than by learning about symbol-sound relationships only within the context of reading extended text. This overall superiority suggested that an important component of any reading instruction program was the incorporation of systematic decoding instruction including extended practice in decoding words.

Christensen and Bowey (2005) found that regardless of the orthographic structure of the word, grapheme-phoneme correspondence was superior to orthographic rime. Their study showed that programs designed to provide explicit practice in the use of letter-sound relationships to decode unfamiliar words can significantly enhance children's performance across a wide range of reading and spelling measures. They also stated that for early readers, a program with focused attention on individual phoneme-grapheme correspondences and encouraged students to analyze every grapheme in a word is superior to one that encouraged them to focus on larger orthographic units, specifically rimes.

Program Components for Struggling Readers

The following three research studies focus on various program components of reading interventions utilized with struggling readers. These program components consist of the use of sentence context when identifying partially decoded words, scripted vs. nonscripted models, as well as grouping conditions for one instructor with one student, three students or ten students in a group.

More recently, Tunmer, and Chapman (2006) completed a longitudinal study to test the relationships among the major learning tasks, learning strategies, and cognitive prerequisites of beginning reading development. The sample comprised of 152 school children in New Zealand with a mean age of 5 years, 1 month, common to the beginning of the Kindergarten year in the United States. Initially, 25 classroom teachers were involved in the study, though this number increased as the study progressed. All participating teachers strongly adhered to the whole-language philosophy of teaching reading.

Reading-related measures were administered at the end of Year 1 (Kindergarten in U.S.), in the middle of Year 2 (first grade in U.S.), and at the end of Year 2. Path analyses

were used to examine these data. These included reading-related measures of: (a) metalinguistic abilities, (b) oral language skills, (c) learning strategies, and (d) reading achievement. Trained research assistants experienced in working with young children individually administered all instruments.

At the end of Year 1, children were presented words containing irregular spelling patterns, in isolation and in context. There was a significant difference, $F(1,140) = 724.82$, $p < .001$, when incorrect pronunciations of target words were presented in isolation compared to when the same mispronunciations were presented in underdetermining contexts (natural occurring situations). This suggests that underdetermining contexts also known as sentence context makes an important contribution to the identification of partially decoded words.

The ability to use letter-sound patterns ($r = .86$) and the ability to use sentence context ($r = .81$) made the strongest independent contributions to variance in early reading achievement ($p < .01$). Tunmer and Chapman (2006) found:

> An important implication of these findings is that when beginning readers encounter unfamiliar words in text, they should be encouraged to use letter-sound patterns first and then to use sentence context, but only to confirm hypotheses about what unfamiliar words might be, based on information from partial decoding attempts. (p. 624)

In conclusion, the researchers suggested that sentence context made an important contribution to the identification of partially decoded words. The reading support program utilized in the current study is *Leveled Literacy Intervention (LLI;* Fountas & Pinnell, 2009), which incorporates the use of daily texts for students to assist in the identification of partially decoded words.

The goal of a study by McIntyre, Rightmyer, and Petrosko (2008) was to examine first-grade struggling readers' phonics and reading achievement. These students received instruction from either a scripted reading model or a non-scripted reading model. Their study included 108 struggling first-grade readers in 37 classrooms across 12 schools. They identified the lowest achieving 20% of students in each class as struggling readers.

Scripted reading programs, also referred to as Direct Instruction, provided teachers with what they were to say verbatim with explicit and systematic instruction. The steps were detailed in action and words. The scripted program included a specific sequenced guide and expected responses from students. Teachers were expected to follow the script regardless of the students' responses. The scripted program utilized in the study by McIntyre et al. was SRA Reading Mastery.

Non-scripted reading programs had a wide-range of pedagogical strategies and time devoted to the teaching of phonics. These programs were dependent on the knowledge of teachers for effective practices. This study utilized, Breakthrough to Literacy, Early Success, Four Blocks and Together We Can as non-scripted reading programs in one group of non-scripted models.

All participating students were pretested in September and given the posttest during the month of May in one-on-one situations for no longer than 30 minutes at a time. One assessment utilized was Clay's (1993) *Hearing Sounds in Words Test* (as cited in McIntyre et al., 2008), a phonics application task that included encoding a sentence. An informal reading inventory involving students reading fiction and nonfiction passages, retelling what they read, and answering a series of comprehension questions was also completed.

A one-way analysis of variance (ANOVA) was completed to compare the phonics achievement as measured by Clay (1993) after students spent one year in either scripted or non-scripted instructional settings. The ANOVA revealed no significant difference among the mean phonics gain scores of these first graders. Therefore, first graders progressed equally in the learning of phonics after receiving scripted or non-scripted instruction.

An ANOVA was also completed comparing reading achievement as measured by an informal reading inventory after two years of instruction in either scripted or non-scripted models. The results of this ANOVA revealed that there was no significant difference between the mean gain score for the scripted model ($M = 2.02$) and the mean score for the non-scripted models ($M = 2.06$).

The authors concluded that neither scripted nor non-scripted approaches to early reading instruction were more effective for phonics or reading achievement of first-grade struggling readers. They explained that students may not have been ready for the kind of instruction they were receiving. Another is the fidelity of implementation of the scripts being followed accurately and similarly. In addition, McIntyre et al. (2008) believed that if instruction met the developmental needs of struggling readers, varied instructional models may have made a difference in achievement. They recognized that limitations of their study included the large number of schools and teachers, which resulted in only four observations per classroom. The current research utilized *Leveled Literacy Intervention* (*LLI*; Fountas & Pinnell, 2009) a scripted model, therefore, this study was valuable in determining that the use of scripted models for reading instruction is as effective as non-scripted models.

A study by Vaughan et al. (2003) involved three grouping conditions 1:1 (one teacher with one student), 1:3 (one teacher with three students), or 1:10 (one teacher with ten

students), the main question addressed was: “How do struggling second-grade readers who are provided the same supplemental reading intervention compare when assigned to one of three grouping conditions: 1:1, 1:3, or 1:10?” (p. 303). The study took place in 10 Title I elementary schools in two neighboring school districts in an urban area of the southwest. The final sample group consisted of 77 students who met the following criteria: (a) identified as a struggling reader and at-risk for referral to special education or failure to read effectively by teacher, (b) failure on the second-grade state-level screening benchmark, and (c) parent and child agreement to participate.

Ethnic diversity included: 74% Hispanic, 22.1% African American, and 3.9% White. Boys consisted of nearly 52% of the group. Students were assigned to one of three group sizes: (a) one teacher to one student (1:1), (b) one teacher to three students (1:3), or (c) one teacher to ten students (1:10). Students in each group received the same 30-minute intervention, five times a week, for a total of 58 sessions over a 13-week period.

Scores on the *Dynamic Indicator of Basic Early Literacy Skills* (*DIBELS*; Good & Kaminski, 1996) phoneme segmentation probe were used to categorize students’ phoneme segmenting skills as high, medium, or low. Following student identification, they were assigned to form student groups of 1, 3, and 10. Students in each of these groups received the same 30-minute intervention, five times a week, for a total of 58 sessions over a 13-week period. A state screening and inventory used to evaluate reading development and comprehension skills in children in kindergarten through second grade was used as a pretest to identify potential participants and as a posttest to determine how many students were able to achieve acceptable ratings from the state accountability system after the intervention.

Word Attack and Passage Comprehension subtests of the *Woodcock Reading Mastery Test-Revised* (*WRMT-R*; Woodcock, 1987) were used as pre- and posttest measures and as a follow-up measure at 4 weeks postintervention. The *Test of Oral Reading Fluency* (*TORF*; Children's Educational Services, 1987) was administered at each of the three assessment points. Three subtests of the *DIBELS*: (a) phoneme segmentation fluency, (b) letter naming, and (c) nonsense words, were used for progress monitoring on a weekly basis. All pretest data were collected prior to the start of the intervention. Posttest data were collected at the end of 58 sessions, and follow-up data were collected after 4 weeks of school following posttest. All data collection was completed by trained research assistants who were blind to the conditions and hypotheses.

Teachers providing instruction included five females with at least one year experience in teaching reading to students with reading difficulties either individually or in small groups. All teachers taught students in groups with ratios of 1:1 and 1:3, in addition two teachers with the most teaching experience taught groups with a ratio of 1 teacher to 10 students. Prior to the beginning of the study teachers were given 22 hours of training on the five instructional elements of the lesson. Each lesson included the following five elements: (a) fluency building, (b) phonological awareness, (c) comprehension strategies during instructional-level reading, (d) word study, and (e) progress monitoring.

Statistically significant differences occurred among the three groups, $F(2, 70) = 3.52$, $p = .035$; $MSE = 31.85$, on passage comprehension. In a follow-up to this analysis, the 1:1 ($M = 17.8$, $SD = 7.1$) and 1:3 ($M = 17.6$, $SD = 6.0$) groups outperformed the 1:10 group ($M = 14.9$, $SD = 5.7$) at the .05 level of significance. There were no statistically significant differences between the 1:1 and 1:3 groups. In phoneme segmentation, the grouping main

effect was statistically significant, $F(2, 70) = 4.13$, $p = .020$; $MSE = 125.97$. When the three groups were compared, the 1:1 group ($M = 5.3$, $SD = 8.6$) significantly outperformed the 1:10 group ($M = 46.4$, $SD = 9.3$). There were no statistically significant differences between the 1:1 and 1:3 groups or the 1:3 and 1:10 groups.

Reading fluency was statistically significant for the grouping main effect, $F(2, 70) = 3.40$, $p = .039$; $MSE = 224.82$. The 1:1 group ($M = 47.4$, $SD = 34.0$) outperformed the 1:10 group ($M = 39.2$, $SD = 14.8$) at the .05 level of significance. The differences between the 1:1 and 1:3 groups and the 1:3 and 1:10 groups were not statistically significant. None of the main and interaction effects on word attack were statistically significant.

In conclusion, for reading comprehension, both the 1:1 and 1:3 groups were superior to the 1:10 group. The 1:1 group was not superior to the 1:3 groups on any outcome measure; therefore, both are highly effective intervention group sizes for supplemental reading. One limitation of this study was the lack of a control group of students with reading difficulties who did not participate in the intervention. To control for teacher and classroom effects, the researchers stated that they sampled students across teachers' classrooms. Despite attempts to have equal group sizes, it was not possible due to students moving away from the district prior to completion of the intervention. In summary, this supports the use of a *Leveled Literacy Intervention* (*LLI*; Fountas & Pinnell, 2009) study, which provides instruction in groups of three students with one interventionist.

These seven research studies reviewed are all uniquely related to the current research study. They support the use of scripted and non-scripted programs that utilize daily text for sentence context to assist with partially decoded words. The use of progress monitoring systems closely monitors and gives students time to meet a specified criterion to full

retention. Not all instructional approaches are successful for all students. The use of progress monitoring will assist in identifying when an approach is not doing well. A systematic and explicit program may focus on phonological and orthographic connections to words and texts, segmentation of words or any specific needs a student may have.

Progress Monitoring to Inform Instruction

The following research review incorporates the relationship of progress monitoring with specific reading skills and teacher feedback for effective instructional impact. The importance of choosing an appropriate progress monitoring system along with barriers and facilitators for using progress monitoring to inform literacy instruction is also examined.

Progress Monitoring with Specific Reading Skills

The following three research studies review the use of progress monitoring with specific reading skills. They investigate the use of word identification and nonsense word fluency for assessing early reading development, progress monitoring with the use of diagnostic feedback, as well as specific reading skills and their function in reading development.

Fuchs, Fuchs, and Compton (2004) designed a study to contrast the concurrent and predictive validity for the two alternative curriculum based measurement (CBM) early reading measures: word identification fluency and non-sense word fluency. The sample of students was screened from an intervention study examining the effects of Peer-Assisted Learning Strategies (PALS) in first grade (McMaster, Fuchs, Fuchs, & Compton, 2005). At least three first-grade teachers in each of eight schools in a southeastern, metropolitan public school system volunteered to participate in the study.

The sample consisted of 151 first grade at-risk students, the lowest third (identified utilizing rapid letter naming), of each classroom. Approximately 50% were in Title I schools; 59% received subsidized lunch; 52% were male. Ethnicity diversity included 38% African American, 35% European American, 24% Hispanic and less than 1% Asian.

Participating students were assessed in the spring and fall using the two CBM measures as well as a set of criterion reading measures. The two CBM measures were used to monitor students for 20 weeks, once weekly for the first 7 weeks and twice weekly for the final 13 weeks. Criterion measures included the Word Attack and Word Identification subtests of the *Woodcock Reading Mastery Test-Revised* (*WRMT-R*; Woodcock, 1987). These instruments were administered in the fall and spring, with fluency and reading comprehension measures administered in the spring only. Progress-monitoring measures included word identification fluency and nonsense word fluency. All data were collected by trained master or doctoral students, all of whom had been trained to 100% accuracy in data-collection procedures prior to beginning any data collection.

Results indicated that at the beginning of first grade a statistically significant correlation existed between word identification fluency and word identification (.77, $p <$ .001), and another correlation existed between nonsense word fluency (.58, $p < .001$) and word identification. This supports this researcher's use of core reading words as a measure of literacy skills.

In conclusion, the researchers found that these results indicated word identification fluency functioned better than nonsense word fluency as a CBM tool for assessing early reading development in first grade. Student progress monitoring with word identification fluency can contribute significantly to the identification of students likely to experience

difficulty in learning to read in the first grade. Word identification fluency over time reflects improved performance on important end-of-year reading outcomes. A limitation of this study was the restricted sample of at-risk students. With this in mind, the large correlations for word identification fluency are even more impressive.

Continuing to look at CBM, Capizzi and Fuchs (2005) conducted a study to assess the effects of curriculum-based measurement (CBM), with and without diagnostic feedback, on general and special educators' instructional planning in reading. Participants included 35 elementary school teachers in 15 schools in a southeastern metropolitan school district. This sample included 19 general education second-grade teachers and 16 special education resource teachers all of whom volunteered and were compensated $150 for participation. Teachers were randomly assigned to three treatments. Treatments included: (a) CBM+D, CBM with diagnostic feedback, (b) CBM, CBM without diagnostic feedback, and (c) Control, no CBM.

Students were drawn from the 35 participating teachers' classrooms. For eligibility, students had to read on at least a first-grade level (as judged by their teacher), be proficient in English (as judged by their teacher), and have a signed parental consent (n = 427). Research staff selected one high-achieving (HA), one average-achieving (AA), and one low-achieving (LA) student from each classroom as target students in order to assess instructional differentiation for students of varying achievement levels. Trained research assistants conducted all sessions collecting CBM data using a CBM computer program. All students, regardless of the students' grade level, read passages at a second-grade level for 1 minute. The number of words that the student read correctly in 1 minute was automatically calculated by the computer. During the third week, students reading below a benchmark score

associated with decoding competence were administered a CBM decoding inventory in order to give more specific feedback on the class diagnostic feedback profile.

Control group teachers completed classwide and individual planning sheets. They then received their CBM+D class reports along with a 40-minute information session on interpreting these documents. CBM teachers were provided CBM class reports without diagnostic profiles and 20-minute training in interpreting the CBM class report. They then completed their classwide and individual planning sheets. CBM-D teachers were given reports and a 40-minute training sessions in interpreting the reports. Following the training, teachers completed classwide planning sheets for one week of reading instruction. CBM-D reports detailed student instructional needs in comprehension, fluency, or phonics. Individual planning sheets were also completed detailing the nature of the individualized instruction for the HA, AA, and LA target students.

There was no statistically significant difference identified for the amount of time allocated for weekly reading instruction for teacher background or CBM condition. Control teachers incorporated a greater number of different objectives than CBM+D teachers did (p = .011, ES = .56). Capizzi and Fuchs (2005) found, disappointing, though not unexpected, that the more focused nature of teachers' planning when given CBM-D did not demonstrate a significant effect of condition or teacher background on the fit between students needs, and planned instruction. In other words for second-grade teachers, the only effect was for AA students when CBM was compared to CBM+D ($F(1, 12) = 5.18$, p = .02), favoring CBM+D.

Special Education teachers who received curriculum-based measurement with diagnostic feedback as compared to teachers receiving CBM without diagnostic feedback, planned instruction with a better fit to student needs of target low-achieving ($F(1, 11) = 7.54$,

p = .01), and average-achieving students (F(1, 11) = 5.24, p = .02). CBM teachers, in turn, planned more responsively than teachers receiving no CBM information. When Special education resource teachers and Grade 2 CBM teachers were combined into one group, there was significance favoring those who received CBM with diagnostic feedback as compared to teachers who received CBM without diagnostic feedback, in planning for all students LA (F(1, 23) = 5.24, p = .02), AA (F(1, 23) = 9.99, p = .00), and HA (F(1, 21) = 6.48, p = .01).

In conclusion, the researchers understood that CBM with diagnostic feedback had the potential to enhance resource teachers' instructional differentiation for students of varying levels. Limitations of this study included minimal training in interpreting and using CBM and diagnostic feedback reports for teachers. The researchers believe future study is warranted with teachers being provided more extensive training and support.

In another study, Hosp and Fuchs (2005) completed a study with the primary purpose of addressing whether the relation between curriculum-based-measurement (CBM) and specific reading skills changed as a function of reading development. A secondary purpose was to identify if CBM cutscores corresponded with benchmark performance on the *Woodcock Reading Mastery Test-Revised* (*WRMT-R*; Woodcock, 1987) measures. Participants included 310 English-speaking students in 16 classrooms from four schools in a southeastern metropolitan area. The 16 classrooms included four at each grade, first through fourth. By grade level a range of 53%-57% were males, 37%-56% were African American, 39%-56% were Caucasian, and 1%-11% were other. The percentage of students receiving free or reduced lunch at each of these four schools was 81.8, 42.4, 41.2, and 34.4.

All participating students were administered two CBM reading passages at students' grade-appropriate reading level. The CBM score was the average of the number of words

read correctly in one minute across the two passages. Participants were also given three subtests of the *Woodcock Reading Mastery Test-Revised* (*WRMT-R*; Woodcock, 1987): (a) Word Identification (word reading), (b) Word Attack (decoding), and (c) Passage Comprehension (comprehension). The Basic Skills score indexing the students' overall word reading ability and the Total Reading-Short (total reading) score determining the students' overall reading ability were also analyzed. Seven trained research assistants administered all assessments over two sessions during the last two months of the regular school year, with between 2 and 18 school days between the two sessions.

Correlation Coefficients between CBM and the *WRMT-R* scores (decoding, word reading, comprehension, basic skills, and total reading) were all statistically significant at $p <$ .01 level at all grade levels ranging from $r = .71$ to $r = .91$. Correlations across grades indicated that the relation between CBM and decoding were generally higher in Grades 2 ($r = .82$) and 3 ($r = .82$) than in Grades 1 ($r = .71$) and 4 ($r = .72$). Correlations between word reading and CBM were stronger at Grades 1 ($r = .91$), 2 ($r = .88$), and 3 ($r = .88$), than in Grade 4 ($r = .73$). The correlation between word reading and CBM supports this researcher's use of core reading words to measure literacy skills.

Correlations between all variables within each grade indicated that at Grades 1 and 3, CBM had a stronger relation with word reading (Grade 1, $r = .91$; Grade 3, $r = .88$) than with decoding (Grade 1, $r = .71$; Grade 3, $r = .82$) or comprehension (Grade 1, $r = .79$; Grade 3, $r = .84$). The relationship between CBM and total reading appeared to be stronger than the relation between CBM and all other skills at Grades 1, 2, 3, and 4 as shown in Table 1.

Table 1

Correlations Between CBM Decoding, Word Reading, Comprehension, Basic Skills, and Total Reading-Short on the WRMT-R

	Grade 1	Grade 2	Grade 3	Grade 4
Decoding	$r = .71$**	$r = .82$**	$r = .82$**	$r = .72$**
Word Reading	$r = .91$**	$r = .88$*	$r = .88$**	$r = .73$**
Comprehension	$r = .79$**	$r = .83$**	$r = .84$**	$r = .82$**
Basic Skills	$r = .86$**	$r = .89$**	$r = .87$**	$r = .78$**
Total Reading	$r = .90$**	$r = .91$**	$r = .91$**	$r = .83$**

**$p < .01$.

In summary, this study supports the appropriateness of using CBM for monitoring specific reading subskills, such as decoding, word reading, and comprehension, as well as for tracking more global reading competence such as basic skills and total reading. The researchers stated, "The CBM cutscores at each grade level may assist in identifying students who need more intensive instruction in reading in general, and students who require diagnostic testing to determine the subskills on which to focus that instruction" (p. 25).

Effective Instructional Impact of Progress Monitoring

The following two studies review the effective instructional impact of progress monitoring. They analyze the importance of choosing an appropriate progress monitoring tool based on the students skill levels, as well as identifying barriers to using progress monitoring.

Utilizing a more specific curriculum based measure, Olinghouse, Lambert, and Compton (2006) developed a study using an Intervention Aligned Word List (IAWL)

specifically designed as a progress-monitoring assessment. Their study evaluated whether the IAWL and Oral Reading Fluency (ORF) measures differentially predicted growth in reading skills. The sample included 40 special education students identified by their teachers as having word-level reading difficulties. These students were from a metropolitan school district in the southeastern United States. Students met the following criteria to be eligible to participate in the study: (a) received resource room services for reading instruction; (b) had individualized education program goals in the area of decoding skill acquisition; (c) had a composite score on the *Test of Word Reading Efficiency* (*TOWRE*; Torgesen, Wagner, & Rashotte, 1997) below the 25th percentile; (d) had an estimated IQ above 70; and (e) had no documented neurological or emotional problems, no uncorrected sensory deficits, and were not English language learners.

Participating students included 18 boys and 22 girls. This group included 20 third graders, 16 fourth graders, and 4 fifth graders. The reading intervention utilized was the *Phonological and Strategy Training Program* (*PHAST*; Lovett, Lacerenza, & Borden, et al., 2000) developed by Lovett and colleagues at the University of Toronto. The *PHAST* program was a systematic and sequential reading program in which students received phonologically based remediation along with word identification strategies. All students received 60 *PHAST* lessons over 70 hours of instruction in groups of three to five. Scripted lessons and instructional materials were followed by graduate research assistants trained from the University of Toronto.

Trained project staff individually administered all pretest and posttest assessments. These assessments included the *Gray Oral Reading Test-3* (*GORT*; Wiederholt & Bryant, 1992) and the *Test of Word Reading Efficiency* (*TOWRE*; Torgesen et al., 1997). The word

identification and word attack subtests of the *Woodcock Reading Mastery Test Revised/Normative Update* (*WRMT-R/NU*; Woodcock, 1987) were also utilized.

Progress monitoring of all participants occurred six times throughout the study with the IAWL and ORF measures. This first administration occurred prior to the reading intervention, while the five others occurred equally spaced throughout the lessons. Linear growth models provided adequate fit for the IAWL and ORF data. The IAWL measures indicated on average, students read 11.2 words correctly at lesson 0 and gained 3.1 words read correctly each assessment wave, at $p < .01$. The ORF measures indicated on average, students read 44.9 words correctly per minute at lesson 0 and gained 2.6 words per minute each assessment wave, also at $p < .01$. This showed a general trend of improvement for students on both the IAWL and ORF assessments.

The two parameters of IAWL and ORF were significantly correlated ($r = .67$, $p < .000$). The IAWL slope parameter demonstrated significant results in unique variance for word identification ($r = .29$, $p < .01$), word attack ($r = .16$, $p < .01$), word efficiency ($r = .18$, $p < .05$), and passage reading accuracy ($r = .16$, $p < .05$). The ORF slope parameter demonstrated significant results for passage reading fluency ($r = .19$, $p < .01$).

In summary, the researchers felt the results partially supported their hypothesis that the IAWL and ORF measures would differentially predict growth on standardized tests of reading. While the Intervention Aligned Work List (IAWL) appeared to monitor intervention specific goals, the Oral Reading Fluency (ORF) assessment appeared to measure generalized effects. Olinghouse et al. stated, “These results address the importance of choosing an appropriate progress-monitoring assessment based on the skill level of the

student, the targeted remediation, and the goals of the intervention" (p. 100). One limitation of the study was that the IAWL was specific to the *PHAST* intervention.

Recently, Roehrig, Duggar, Moats, Glover and Mincey (2008) explored how teachers in the Reading First context are more or less effective in trying to use student progress monitoring data to inform their literacy instruction. The goal of their qualitative study was to explore what teachers perceived to be the barriers or facilitators to using progress monitoring data to inform literacy instruction.

Participants included 10 kindergarten and first-grade teachers and four reading coaches, working at four Florida Reading First schools in one midsized school district. In Florida, Reading First schools assessment teams administer progress monitoring assessments to all students, at least four times a year. Online progress monitoring data reports are available to teachers to review. The school reading coaches facilitated professional development within their schools.

In the spring semester, written survey data were collected from teacher volunteers (n = 30). Data were analyzed to select schools with varying involvement and success in using data to inform instruction. From the four schools, 10 teachers were selected based on three criteria: (a) teachers' openness or resistance to using data (chief criterion); (b) grade level taught; and (c) group assignment in a concurrent professional development experiment. The latter two criteria were evenly distributed across the range of teachers' levels of success, skills, and attitudes about using data to inform instruction.

Semi-structured interviews based on a framework of 28 open-ended questions were conducted during the last two weeks of school and into the summer. Teachers were specifically asked to address their experiences in: (a) teacher training in reading, (b)

knowledge of teaching reading and the reading program they use, (c) use of assessment data and their challenges, and (d) supports they received in trying to use data to inform instruction. Data were open coded, with relationships between categories being identified. Axial coding was conducted using the categories and subcategories that emerged. Categories were identified that Roehrig et al. (2008) agreed were central to the phenomenon itself. Triangulation of data for all 30 teachers confirmed that no conflicting or new themes emerged.

Surveying of the participating reading coaches was completed to establish credibility of interpretations drawn from the data (Lincoln & Guba, 1985). All of the coaches' responses confirmed how coaches and their teachers used assessment data to inform instruction. Some of the barriers teachers described when attempting to use assessment data to inform instruction included: (a) coach availability and quality of support received, (b) breakdown between receiving assessment results and what to do with children, (c) teacher knowledge, and (d) willingness of teachers to examine the effectiveness of their practices using student assessment. The key factor for greater success in using the data was the reading coach, who was responsible for helping teachers.

Roehrig et al. (2008) concluded that coaches should not be placed in schools to serve as the lone instructional leaders. They stated that few states have enough qualified individuals to fill Literacy Coaches positions. They suggest that patience may be needed to implement school level reform this large. One weakness of the study was that prior to the study all participating teachers and reading coaches did not believe in using assessment data to inform instruction.

The studies reviewed in this section all support the use of progress monitoring in a variety of ways. They supported the use of progress monitoring for specific reading subskills. Diagnostic feedback along with progress monitoring is valuable for the success of the students. It is important that a progress monitoring assessment is based on the skill level of the student. Through progress monitoring word identification fluency was identified as reflecting improved performance supporting the use of core reading words as a measure of literacy skills in this study.

Response to Intervention Improving Student Learning

All states are federally mandated to employ a Response to Intervention (RtI) model in their schools. Various aspects of this model must be reviewed for states to successfully implement RtI. The following studies review the levels of RtI along with how it impacts higher and lower responders. The use of RtI for the identification of a Reading or Learning Disability and the fidelity of implementing this type of program are also analyzed.

Simmons et al. (2008) examined the alterability and stability of reading performance among children identified as at-risk of reading difficulty in kindergarten. They also focused on profile patterns of response to intervention from kindergarten through third grade. All kindergarten students (n = 464) from seven elementary schools in the Pacific Northwest were screened on Letter Naming Fluency (LNF) and Initial Sound Fluency (ISF) on the *Dynamic Indicators of Basic Early Literacy Skills* (*DIBELS*; Good & Kaminski, 1996). Students were identified as at-risk according to the following criteria: (a) scored at or below the 30th percentile on national norms on LNF, (b) performance was confirmed by their kindergarten teachers as being at-risk of reading difficulty, and (c) scored in the bottom quartile of their local kindergarten cohort on ISF. Students who had severe hearing or visual acuity problems

or who had limited English proficiency were excluded from the study. In the fall of their year in kindergarten, 117 students were identified to participate in the study.

All participating schools received Title I funding and 32%-63% received reduced-cost lunch services. Due to the transient population, student participation decreased to 88 students by the end of kindergarten, 71 students at the end of first grade, 60 students at the end of second grade, and 41 at the end of third grade. All analyses were conducted on the cohort of 41 students who participated all four years, kindergarten through Grade 3. The sample included 27 males (66%). Ethnic diversity included 84% European Americans, 14% Hispanic/Latino students, and 1 student who was African American.

Students were all assessed in the fall of each academic year to evaluate the need for an intervention. All students who fell below the 30th percentile on criterion measures administered in the fall continued to receive supplemental intervention. Students who met the benchmark in the fall were discontinued from intervention but were assessed in the fall and spring. In first grade, student progress was also evaluated in late January (midyear). Students who met the midyear benchmarks were discontinued from intervention, while those who did not meet these benchmarks continued intervention.

Interventions across the four-year period focused on a developmental sequence of skills and strategies in beginning reading. On average, students received supplemental interventions for 30-45 minutes, 5 days per week from November through May, in small groups of 3 to 5 students. Intervention was provided by certified teachers or highly trained paraprofessionals. Measures of (a) phonological awareness, (b) phonemic decoding and word reading, (c) oral reading fluency, and (d) reading comprehension were administered at developmentally appropriate measurement points utilizing subtests of the *Dynamic*

Indicators of Basic Early Literacy Skills (*DIBELS*; Good & Kaminski, 1996) and *Woodcock Reading Mastery Test-Revised* (*WRMT-T*; Woodcock, 1987).

The kindergarten interventions compared three programs: (a) highly explicit code-emphasis (HE-C), (b) highly explicit code- and comprehension-emphasis (HE-CC), and (c) moderately explicit code-emphasis (ME-C). Each of these interventions occurred daily for 30 minutes over 21 weeks. The highly explicit code (HE-C) intervention emphasized strategic and systematic instruction of phonemic awareness, alphabetic understanding, letter writing, and spelling. The highly explicit code (HE-C) intervention also focused on high priority phonological and alphabetic skills as well as: (a) receptive and expressive knowledge of vocabulary and (b) expanded knowledge and development of story structure and story retell.

The highly explicit code- and comprehension-emphasis (HE-CC) intervention included two 15-minute components. The first 15 minutes focused on high-priority phonological and alphabetic skills and included the same lessons as the HE-C intervention. The second 15 minutes had two focuses: (a) receptive and expressive knowledge of vocabulary that appeared in story-books and (b) expanded knowledge and development of story structure and story retell. The third intervention, moderately-explicit code-emphasis (ME-C), focused on phonological, alphabetic, and orthographic activities.

Results indicated for the *WRMT-R* Word Attack at the beginning of first grade, the probability of being classified as out-of-risk (90.2%) was statistically higher ($p < .05$) than the probability of being classified as at-risk (9.7%). By the end of third grade, the same probability of being classified as out-of-risk increased to 95.1%. In terms of performance by the kindergarten intervention condition, a higher percentage of students in the HE-C group

attained out-of-risk status earlier, though there was no statistically significant difference between the HE-C and the other two interventions (HE-CC and ME-C) on the change of risk status at the beginning of first grade and the end of third grade.

A similar pattern was found at the beginning of first grade, for *WRMT-R* Word Identification, probability of being classified as out-of-risk (85.4%) was statistically higher ($p < .05$) than the probability of being classified as at-risk (14.6%). By the end of third grade, the probability of being classified as out-of-risk increased to 92.7%. Response by kindergarten intervention revealed that the majority of students across all interventions responded to kindergarten intervention in a similar way, with all but three groups of students performing above the 30th percentile by the end of third grade.

For *WRMT-R* Passage Comprehension, the probability of being classified as, out-of-risk (58.5%), was higher than that of being classified as, at-risk (41.5%), at the beginning of first grade, this difference was not statistically significant. By the end of third grade, the probability of being classified as, out-of -risk (92.7%) was statistically higher ($p < .05$) than the probability of being classified as at-risk (7.3%). There were no statistically significant differences between the kindergarten interventions on the change or risk status based on WRMT-R Passage Comprehension at the beginning of first grade and the end of third grade.

The Oral Reading Fluency (ORF) was not measured in the fall of first grade. A combination of *DIBELS*, Phonemic Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF) were utilized to identify students at-risk. At the beginning of first grade the probability of being classified as out-of-risk (53.7%) was slightly higher than the probability of being classified as at-risk (46.3%), though not statistically significant. Likewise, at the end of third grade, the probability of being classified as out-of-risk (49%) was not

statistically different from the probability of being classified as at risk (51%). This indicated that many students failed to attain the 30th percentile score of 99 correct words per minute. Considering, response by kindergarten intervention, at the beginning of first grade, the odds of students' being classified as at-risk based on PSF and NWF was statistically lower (ß = -1.50, $p < .05$) in the HE-C group than in other intervention groups.

The researchers stated, "In summary, absolute performance levels, on average across a range of reading measures attained at the end of kindergarten positioned students for trajectories of reading performance that exceeded the 50th percentile on the majority of measures" (p. 169). In this study, students not only moved out-of-risk but stayed out-of-risk during the kindergarten-through-third grade period. The researchers concluded that students in all three kindergarten intervention programs (highly explicit code emphasis; highly explicit code and comprehension emphasis; and moderately explicit code emphasis) responded to the kindergarten interventions in a similar way. The need to replicate and extend their research was identified by the researchers.

In another study, Vaughn et al. (2009) examined whether students who demonstrated minimal response to previous, less intensive interventions in first grade would profit from continued intervention after the level of intensity was altered (more time each day and smaller group size) and the intervention was more extensive (provided for a longer period of time). Participants included two cohorts of students from all seven elementary schools in a small school district near a large city in the southwest.

Students were participating in a multiyear longitudinal research project on the effectiveness of a three-tier intervention model. Students meeting at-risk criteria for reading difficulties (Cohort 1: $n = 153$; Cohort 2: $n = 121$) were randomly assigned to treatment and

comparison groups prior to the first grade. Students in the treatment group received intervention from the research team. Students in the comparison group received typical school services. The sample for this study was composed of students who were assigned to the treatment group and remained in the district throughout the 2-year period (first and second grade). Students in the treatment group were further classified as Higher Responders and Lower Responders. Researchers identified oral reading fluency (ORF) scores below 27 in the fall of second grade as the cut point for higher and lower responders. Students with scores below 27 were considered lower responders since they showed some response to instruction though they were unlikely to make adequate progress toward grade-level reading skills. Likewise, students with oral reading fluency (ORF) scores above 27 were considered higher responders.

Throughout first grade, the higher responders received either 13 or 26 weeks of secondary intervention for 30 minutes daily. This occurred in groups of four to six students with one tutor who was hired and trained by the research team. This intervention was in addition to the students' regular primary classroom reading instruction, which focused on phonics and word recognition, fluency, passage reading, and comprehension. These students (n = 34) met the identified benchmark in the beginning of second grade (ORF above 27), therefore did not require secondary or tertiary interventions in second grade. These students continued to receive primary classroom reading instruction.

All lower responders received the same 26 weeks of secondary intervention provided to the higher responders in first grade. These students did not reach the identified benchmark in the beginning of second grade (ORF below 27). Therefore, these students (n = 14) were provided a tertiary intervention. This intervention was more intensive and occurred in small

groups of two to four students for 50 minutes daily with a tutor trained and supervised by the research team for approximately 26 weeks. These sessions included the topics of sound review, phonics and word recognition, vocabulary, fluency, passage reading, and comprehension.

Measures utilized were subtests of the *Woodcock Reading Mastery Tests-Revised* (*WRMT-R*; Woodcock, 1987), *Dynamic Indicators of Basic Early Literacy Skills* (*DIBELS*; Good & Kaminski, 2002), and the *Peabody Picture Vocabulary Test-III* (*PPVT-III*; Dunn & Dunn, 1997). A social skills instrument designed to assess behavioral, academic, and social competence were completed by all classroom teachers for each student. All assessments were administered individually by trained graduate and research associates who were unaware of treatment and comparison group assignment. The *WRMT* and the *DIBELS* ORF subtest were administered three times per year, while the *PPVT-III* was administered only in May of first grade.

A regression-discontinuity (RD) research design was utilized to assess the effectiveness of the tertiary intervention. The researchers stated, "It is appropriate to use RD when the group receiving intervention (lower responders) and the comparison group (higher responders) are purposely selected to differ in ability as assessed by a quantitative criterion prior to the introduction of the intervention" (p. 174).

There were no statistically significant effects for the tertiary intervention on the spring ORF score or on the Word Attack score. There was a statistically significant interaction effect for intervention on the Passage Comprehension score (ß = 0.67, $p < .001$). This demonstrated that the program was differentially effective depending on students' fall ORF scores. The researchers felt that students in the lower responder group with the highest

ORF scores in the fall of Grade 2 benefited significantly from the tertiary intervention. There also was a statistically significant interaction effect for the intervention on the Word Identification score (ß = 0.57, $p < .001$). This finding indicated that the program was differentially effective depending on students' fall ORF scores. Again the researchers believed that students in the lower responder group with the highest ORF scores in fall of Grade 2 benefited significantly from the tertiary program, just as for Passage Comprehension.

Vaughn et al. (2009) felt their findings yielded several important outcomes and raised several critical questions. They felt it was valuable to note that over time the majority of students responded well to early reading interventions and made appropriate progress. While lower responders made statistically significant progress on important outcomes (i.e., reading for meaning and reading words correctly), there were also areas (oral reading fluency) where progress was less evident. Progress on outcomes of reading words correctly; support the use of core reading words as a measure of literacy skills in the current research. The researchers, Vaughn et al., also identified possible limitations of their study: (a) that they did not use both growth and benchmark measures to identify students as lower and higher responders, as well as (b) a small sample size.

A different focus of the response to intervention model for identifying students with reading/learning disabilities was examined by Vaughn, Linan-Thompson, and Hickman (2003). The sample included second-grade students ($n = 45$) identified as at-risk for reading disabilities using a two-tiered identification process. Criteria identified for participation included: (a) reading below-grade level in English, (b) scoring at the second-quartile or

below in reading ability, and (c) being at-risk or having failed a second-grade state reading inventory.

A total of 45 students from three Title I schools received the supplemental interventions. The participants were 78% Hispanic/Latino, 13% White, and 9% African American. All measures were administered individually by trained research associates immediately prior to intervention and then 30 weeks later. These measurements included subtests of the: (a) *Woodcock Reading Mastery Test-Revised* (*WRMT-R*; Woodcock, 1987), and (b) *Comprehensive Test of Phonological Processing* (*CTOPP*; Wagner, Torgesen, & Rashotte, 1999). The *Test of Oral Reading Fluency*, (*TORF*; Children's Educational Services, 1987) was administered four times, prior to treatment, then after each of three 10-week intervals.

The intervention was provided by four female tutors experienced in teaching reading to students with reading difficulties. The focus of the intervention was on five elements of reading development identified as essential for beginning readers: (a) phonemic awareness, (b) phonics with special attention to systematic mastery of sound-letter relationships as well as word families, (c) word and text fluency, (d) instructional level reading and comprehension, and (e) spelling. Each group of students received approximately 35 minutes of daily supplemental reading instruction throughout three 10-week periods.

Following each 10-week period, students who met exit criteria were discontinued from participation in the treatment, but continued participation in all testing periods. Exit criteria included: (a) a passing score on the screening; (b) a median-score above 55 CWPM (correct words per minute) on a second grade level passage, with fewer than five errors on the *TORF*; and (c) a score of 50 CWPM on a second grade fluency progress-monitoring

session for at least three consecutive weeks. The instructional format for lessons during the third 10 weeks of intervention was modified to adjust for individual student skill level given the differential rate of progress of the students (n = 21) who had not met exit criteria. Three modifications of instruction occurred: (a) assessment of basic skills was increased, (b) word study and fluency instruction was intensified, and (c) phonological awareness instruction was reduced.

Effect sizes for all groups during the intervention were large (a) early exit: 10 weeks (n = 10), 2.74; (b) mid exit: 20 weeks (n = 14), 3.23; (c) late exit: 30 weeks, (n = 10), 6.06 and; (d) no exit after 30 weeks (n = 11), 2.66. Students who never met criteria to exit from supplemental instruction (n = 11) represented fewer than 25% of the struggling readers in the study.

A priori contrast of early, mid, and late exit groups established that students who met exit criteria demonstrated significantly higher scores at pretest than students who did not meet exit criteria on several indicators: (a) fluency (t = 4.40, p < .01), (b) passage comprehension (t = 2.73, p < .01), and (c) rapid naming (t = 3.70, p < .01). The researchers concluded that in their study, fluency, passage comprehension, and rapid naming were the significant predictors of students who would not meet exit criteria, with rapid naming being the best predictor, therefore, supporting the use of core reading words as a measure of literacy skills in the current study. They believed that there were two lessons from their study with respect to identification and treatment: (a) establishing a priori criteria for success and a maximum amount of time for supplemental instruction can assist in identifying a distinct cohort of students who require intensive and explicit instruction possibly, special education:

(b) students who did not exit the program differed from other students in rapid naming, fluency, and word attack.

Concerned with an important aspect of response to intervention, Kovaleski, Gickling, Morrow, and Swank (1999) examined the academic performance of students participating in Instructional Support Teams (IST) as contrasted with other at-risk students who did not have access to ISTs. Participants consisted of 492 students in Grades 1 through 4 who were referred to ISTs because of academic or behavior problems that affected their academic performance. Educators who had not begun the IST process were requested by their principal to identify students who were "academically at risk" to participate in the study as a control group, designated as the non-IST research sample (n = 237). For each IST and non-IST student one or two comparison students from the same classroom who were not at-risk were selected for observation. These students constituted the third group.

Students were observed at three different intervals: (a) pretest- at the point of identification, (b) posttest- approximately 45 days after the initial observation, and (c) follow-up- at least 80 days after the initial observation. Observations focused on reading, math, or, to a lesser degree, inappropriate classroom behavior. Academic learning time (ALT) was the method used for measuring the impact of the IST process on student performance. The three observable behaviors used as measures of ALT were: (a) time-on-task, (b) task completion, and (c) task comprehension.

The researchers included two phases of participants. Phase one included schools sampled in the 1991-1992 school year that were in their 2nd year of operation for utilizing the IST model (n = 232). Phase two schools were sampled in the 1992-1993 school year, also in their 2nd year of operation for utilizing the IST model (n = 260). A three-person validation

team composed of IST trainers and field practitioners (e.g., principals, support teachers) from different areas of the state visited the schools to evaluate the extent to which the required elements of IST were in place. This validation process occurred at the end of a school's second year of the IST process. This process was used to identify high-implementation and low-implementation schools in each of the two phases.

The analysis compared six groups: (a) a combined average group, (b) the non-IST group, (c) the four IST groups (Phase 1, high implementation; Phase 1, low implementation; Phase 2, high implementation; Phase 2, low implementation). For each dependent variable, a repeated-measures ANOVA was conducted with between groups factors. These between-groups factors included Phase 1 IST, high implementation; Phase 1 IST, low implementation; Phase 2 IST, high implementation; Phase 2 IST, low implementation and non-IST. The repeated factors were the points at which the data were collected (pretest, posttest, follow-up).

There were significant differences between the average group and the treatment groups for comprehension from pretest to posttest. The high-implementation IST groups (Phase 1 and 2) displayed significantly higher gains in comprehension than either low-implementation IST groups (Phase 1 and 2) or the non-IST groups from posttest to follow-up stages, $F(1, 1{,}627) = 7.48$, $p = .006$. The non-IST group displayed a greater gain in comprehension than the low-implementation IST groups, $F(1, 1{,}627) = 4.10$, $p = .043$.

Regarding task completion, the high-implementation IST groups demonstrated significantly higher mean scores in comparison to the low-implementation IST and non-IST groups, $F(1, 1{,}626) = 14.41$, $p < .001$. This continued with time-on-task where the high-implementation IST groups displayed significantly greater gains than the low-

implementation IST and non-IST groups, $F(1, 1{,}628) = 13.11$, $p < .001$. The researchers found that these results implied "that half-hearted efforts at IST implementation are no better for at-risk students than what is traditionally practiced in non-IST schools" (p. 180). Apparently, for certain outcomes, a lack of IST involvement is even better than low implementation. They felt this demonstrated that ISTs must have overall high implementation of all features for improved student performance.

These research studies further supported the current research. The use of a reading support program with an imbedded progress monitoring system will assist in high implementation of the response to intervention model. High implementation is imperative for the RtI model to be effective in each school. Early intervention when providing a reading support program is vital for students to move out-of-risk by the third grade, just in time for federally mandated testing. The use of reading words correctly further supported the use of core reading words as a measure of literacy skills.

Reading Self-Concept

Research focused on reading self-concept is reviewed to examine how this construct develops in children. Subcomponents of a student's reading self-concept are investigated. Research in the development of a suitable scale to measure a student's reading self-concept is reviewed.

Chapman and Tunmer (1995a) conducted four experiments examining the development of reading self-concept in young children. In Experiment 1, their aim was to confirm if young children responded differently to negative items than they did to positive items about reading. Participants included 520 children attending four primary schools in a

large New Zealand provincial city. Students were in Year 1 (Kindergarten in U.S.), Year 2 (first Grade in U.S.), and in Year 3 (second Grade in U.S.) of their educational program.

The original version of the *Reading Self-Concept Scale* (*RSCS*; 1992; as cited by Chapman & Tunmer, 1995a) was administered individually to children by one of four research assistants, taking approximately 25 minutes, per child. This version of the *RSCS* consisted of 50 items designed to assess a range of reading-related self-perceptions. Positive and negative items were dispersed throughout the scale to prevent students from responding to all the items with the same response. The results of a 3 X 2 ANOVA showed significant main effects for year in school, $F(2, 517) = 5.32$, $p < .01$, item type, $F(1, 517) = 300.17$, $p < .01$, and a significant Year X Item Type interaction, $F(2, 517) = 12.21$, $p < .001$. Scheffé tests for individual differences between means indicated that the interaction effect was due to Year 1 students obtaining lower scores (less positive self-concepts) on the negative items than students in both Year 2, $F(2, 517) = 12.95$, $p < .01$, and Year 3, $F(2, 517) = 32.66$, $p < .01$. The results demonstrated a negative item response factor with younger children with a measure of self-concept that specifically focused on reading and reading-related self-perceptions.

Experiment 2, focused on examining whether the use of declarative statements with the referential pronoun *I,* in comparison to interrogative statements with the referential pronoun *you* change the negative item response characteristic of young children. The sample consisted of 267 children in Year 1 (Kindergarten in U.S.), Year 2 (first Grade in U.S.) and Year 3 (second Grade in U.S.) enrolled in two primary schools.

The *Reading Self-Concept Scale: Question Form* (*RSCS-Q*) was developed by changing each item of the *RSCS* into a question, and replacing the pronoun *I* with the

pronoun *you*. It was also administered individually by one of four research assistants. A 2 (experiment) X 2 (item type) X 3 (year) ANOVA found significant interaction effects for Experiment X Item Type, $F(1, 781) = 49.83$, $p < .01$ revealing a smaller mean difference between positive and negative items in Experiment 2 than it did in Experiment 1. Significant interaction effects were also found for Year X Item Type, $F(2, 781) = 15.80$, $p < .01$, mirroring the interaction effect observe d in Experiment 1. The mean scores of positive and negative scales in Experiment 2 with the correlational data indicated that children responded with more consistency to positive and negative items when questions were worded in interrogative statements with the referential pronoun *you*.

In Experiment 3, Chapman and Tunmer (1995a) examined whether the measurement of reading self-concept was both conceptually and psychometrically meaningful if items were grouped in terms of three factors: (a) perceptions of competence in reading, (b) perceptions of difficulty with reading, and (c) attitudes or feelings toward reading. This sample included 444 students enrolled in four primary schools, similar to the students and schools that participated in the other experiments.

For Experiment 3, the researchers developed *The Reading Self-Concept Scale: 30-Item* Version (*RSCS*-30) from the *RSCS-Q* that was used in Experiment 2. The *RSCS*-30 included three subscales, each containing 10 items. Questions representing each of the three subscales were distributed throughout the scale. This scale was administered in the same manner as in Experiment 1 and 2. Confirmatory factor analyses (CFA) were performed with five CFA models tested. Factor loading coefficients for this new model of the *RSCS* indicated that the three factors were well defined. The new model of the *RSCS* proposed three factors: (a) perceptions of difficulty, (b) perceptions of competence, and (c) attitudes.

Coefficients for the competence factor mean = 0.59, for the difficulty factor median = 0.61, and for the attitude factor median = 0.56. Together, the CFA data supported the three-factor structure of reading self-concept in young children.

Experiment 4, examined the relations between the subcomponents and measures of reading and reading-related performance and the developmental trends in scores on the subcomponents. This sample was comprised of 771 children in Years 1 (Kindergarten), 2 (first grade), 3 (second grade), 4 (third grade), and 5 (fourth grade) from 16 schools in a large New Zealand city.

The *RSCS*-30 was administered to all Year 1 to Year 3 students individually by one of the four research assistants. The *RSCS*-30 was administered in groups to all students in Year 4 and Year 5. Suitable reading measures were only available for students in Year 1 and for students in Year 4 and Year 5 in 10 of the 16 schools. For those in Year 1, three reading measures and one spelling task were utilized. Reading performance was assessed through a reading comprehension assessment standardized for use in New Zealand for all students at Year 4 and Year 5 levels. A correlation between measures of reading and reading-related performance at Year 1 demonstrated that only the perceptions of difficulty in reading subscale showed a positive relationship with reading and reading-related tasks ($p < .01$). It was not until year 4 that Competence in reading was significantly associated with reading performance ($p < .01$). Attitude toward reading did not become significantly associated with reading ($p < .01$) until Year 5.

A 5 X 3 ANOVA illustrated there was a significant interaction effect between year level and subscales, $F(8, 1532) = 8.44$, $p < .001$. The Scheffé test discovered a highly significant difference, $F(8, 1532) = 38.90$, $p < .01$, in favor of attitude mean scores for the

Year 1 to Year 3 groups compared to the mean for the Year 4 and Year 5 groups. In conclusion, during the first 3 years of school, children's attitudes toward reading are very positive. Following approximately the 4th year, children indicated less optimism.

In 1997, Chapman and Tunmer focused a longitudinal study on the examination of the emerging causal interplay between reading self-concept and beginning reading performance. The sample consisted of 118 Year 1 (Kindergarten) children. These students were in 22 schools in New Zealand. Measures of pre-reading abilities administered to all students included a phoneme deletion task, a sound matching task, and a letter identification task, that were overseen within the first four weeks of school. All students were administered the *Reading Self-Concept Scale* (*RSCS*, 1993; as cited by Chapman & Tunmer, 1997). This assessment was first administered after a student's first six weeks at school, then again at the end of his or her second and third years of school. Measures of reading performance were selected according to developmentally appropriate activities such as word recognition and reading comprehension. Oral vocabulary was assessed through the use of the *Peabody Pictures Vocabulary Test*—Form M (*PPVT*; Dunn & Dunn, 1981). These assessments were given at the end of a student's second and third year in school.

The relationship between reading achievement and reading self-concept appeared to emerge by the middle of the second year (first Grade) of schooling ($r = .39$, $p < .01$). Reading self-concept scores in mid-Year 2 ($r = .39$, $p < .01$) and in mid-Year 3 ($r = .10$, no significance) in comparison to reading scores in mid-Year 2 ($r = .19$, $p < .05$) and in mid-Year 3 ($r = .49$, $p < .01$) "suggests that reading self-concept is more a consequence than a cause of reading performance" (p. 287). The researchers believed that intervention strategies

for struggling readers should include attention to the child's developing achievement-related self-concepts.

In following years, Chapman, Tunmer and Prochnow (2000) identified children who had developed positive, negative, or typical Academic Self-Concepts (ASC) after two years in school, then examined if these were preceded and followed by differences in reading-related performances and reading self-concept. Participating students were selected from an original sample of 152 Year 1 (Kindergarten) children from 22 schools in New Zealand.

The *Perception of Ability Scale for Students* (*PASS*; Boersma & Chapman, 1992) was administered to measure academic self-concept. The *Reading Self-Concept Scale* (*RSCS*, 1999; as cited by Chapman, Tunmer & Prochnow, 2000) was administered individually on three occasions to measure reading self-concept.

Phonological sensitivity and letter-name knowledge were included as measures of prereading skills. Reading performance was assessed through the use of a variety of subtests on standardized reading assessments: (a) toward the end of Year 1 (Kindergarten), (b) during the middle of Year 3 (second grade) utilizing a word recognition measure, (c) at the end of Year 1 (kindergarten) utilizing reading comprehension, and (d) during the middle of Year 3 (second grade) utilizing a dissimilar reading comprehension assessment. Teachers assessed all students on their reading book level. All tasks were administered individually to each child by research assistants experienced in working with children.

At the beginning of their schooling, students with negative ASCs had significantly ($p < .025$) poorer letter identification ($t = 4.88$) and phonological sensitivity skills ($t = 3.63$) than children with positive ASCs. By the middle of Year 3 (second grade), students with negative ASCs, were inferior ($p < .025$) to students with positive ASCs in word recognition (t

= 5.19), reading comprehension (t = 4.38), and reading book level (t = 3.34). There were also differences with students with negative ASCs and those with typical ASCs though not all significant. These included word recognition (t = 2.19, p < .05), and reading comprehension (t = 2.97, p < .025). In conclusion, these findings demonstrated how rapidly children's achievement-related self-perceptions formed in relation to early learning experiences.

To further support the association between reading self-concept and reading achievement in Year 3 (second grade) children, Rider and Colmar (n.d.) specifically explored the relationship between the components of reading self-concept (difficulty, competency, and attitude) and reading achievement (accuracy, comprehension, rate). The researchers believed that this area required more extensive research due to the recent development of instruments aimed at specifically measuring reading self-concept.

The sample of participants included 80, Year 3 (second Grade) children at three primary schools in the western and southwestern metropolitan regions of Sydney, Australia. The schools drew from a lower range of socio-economic backgrounds. As reported by participants, 52% spoke English at home while 47% spoke a language other than English. Of these participants, 59% were female and 41% were male.

Measures included an Australian developed standardized test of reading ability and the *Reading Self-Concept Scale* (*RSCS*, 1999; as cited by Rider & Colmar, n.d.). Participants were assessed individually in a room free from distractions for a total of approximately 30 minutes. A correlational analysis demonstrated correlations at significant levels between reading achievement and reading self-concept subscales. The strongest relationship was for perceptions of difficulty and accuracy in reading (r = .52, n = 80, p < .001). The relationship

was weakest between attitudes towards reading and reading comprehension ($r = .26$, $n = 80$, $p < .02$), though still significant.

These results led the researchers to formulate a qualitative research question regarding the relative contributions of each aspect of reading achievement to reading self-concept. A standard multiple regression procedure showed that 29 percent of the variance in reading self-concept was explained by the reading achievement subtests ($p < .001$). All three predictor variables made a contribution, though rate (ß = .285, $p = .013$) demonstrated the only statistically significant contribution. Comprehension (ß = .284, $p = .079$) aspects made strong contributions to explaining reading self-concept, while accuracy (ß = .077, $p = .646$) also made a contribution, though not as strong.

In conclusion, researchers understood that this study supported the concept that reading self-concept (competence, difficulty, and attitude) consisted of clearly differentiated sub-components that were all influenced by different skills comprising reading achievement (accuracy, comprehension, and rate). This study also suggested that by 8 years of age, there were firmly established patterns of perceptions of difficulty and competence, negative attitudes, and poor performance. The researchers stated a limitation of their study as the use of only one measure of reading achievement.

All reading self-concept studies in this section support the use of a reading self-concept scale for students in the first and second grade. The use of the *Reading Self-Concept Scale* (*RSCS*; Chapman & Tunmer, 1995b) was supported in the use of positive and negative items when questions are worded in interrogative statements with the referential pronoun you. Children's attitudes toward reading were very positive in the first 3 years of school, beginning to decrease during the 4th year. This supports the use of reading support programs

in the early grades, when students still have positive attitudes about reading, hopefully preventing them from attaining negative attitudes about reading.

CHAPTER THREE: METHODOLOGY

This review of methodology begins with a description of the setting, subjects and the sampling procedure. The research questions and hypotheses along with the research design are confirmed. This is followed by a detailed explanation of the treatment, comparison group procedures, and training of the staff members. All three instruments are explained in detail along with their reliability and validity. The description and justification of the analyses are stated, followed by the limitations of the study. An ethics statement concludes the methodology chapter.

Description of the Setting, Subjects and the Sampling Procedure

This research was conducted in a suburban town in the northeast of the United States, with a total town population in 2008, of 56,852 (Connecticut Economic Resource Center, 2008). On October 1, 2008, there was a student population of 9,953. The district consisted of 11 elementary schools, 3 middle schools, and 2 high schools (Strategic School Profile, 2008-2009). Demographically, the socioeconomic background of the student population was classified as middle to upper class with a median home income of $106,184 and a median house price of $618,000 (Connecticut Economic Resource Center, 2009). Seven percent of the student population was eligible for free or reduced priced meals. The racial/ethnic diversity in the school district included 86.2% white, 6.0% Hispanic, 2.5% Black, 5.2% Asian American and 0.1 % American Indian students, with a total student minority population of 13.8% (Strategic School Profile, 2008-2009).

In the total K-12 population for whom the district was financially responsible, 10.5% of the students were identified with disabilities. According to the 2007-2008 strategic school profile, there were approximately 746 full-time teachers in the district. The average teaching

experience across the district was 14 years and 88.7% of the teaching staff had obtained a Master's degree or higher (Strategic School Profile 2008-2009).

Students and educators from a total of 7 of the 11 elementary schools participated in this study. Refer to Table 2 for the frequency of students at each school as well as those qualified to participate in the study and those who actually participated. The experimental group consisted of students from three elementary schools, while the comparison group included those from four elementary schools.

School principals and Language Arts Consultants (LAC) were asked by the District's K-6 Language Arts Curriculum Leader to participate in the study. Each participating school has 1 or 2 Language Arts Consultants. A LAC in this study is required to hold a Reading and Language Arts Consultant, K-12 certification (097-endorsement in Connecticut). These schools also have 1 to 3 Early Literacy Tutors (ELTs) per building. ELTs, in this study, were required to hold a K-6 teaching certification with extensive, demonstrated experience in the area of language arts. Refer to Table 3 for the frequency of Language Arts Consultants and Early Literacy Tutors who participated in this study.

Table 2

Frequency of Students from Participating Schools

School	Total Enrollment	Population		Qualified		Sample	
Experimental		First Grade	Second Grade	First Grade	Second Grade	First Grade	Second Grade
A	526	85	89	13	7	5	4
B	479	68	82	8	3	6	2
C	440	58	82	6	6	2	0
Total	1445	211	253	27	16	13	6
Comparison							
D	323	59	54	9	6	0	3
E	350	59	65	11	11	4	5
F	502	78	96	5	7	3	2
G	356	59	65	3	7	3	1
Total	1531	255	280	28	31	10	11
Grand Total	2976	466	533	55	47	23	17

Table 3

Frequency of Language Arts Consultants and Early Literacy Tutors from Participating Schools

School	Language Arts Consultant (LAC)		Early Literacy Tutor (ELT)	
Experimental	First Grade	Second Grade	First Grade	Second Grade
A	1	1	0	0
B	1 LAC for First and Second Grade		0	0
C	1	0	0	0
Total	3	1	0	0
Comparison				
D	0	0	0	1
E	0	0	1	1
F	0	0	1 LAC for First and Second Grade	
G	0	0	1 LAC for First and Second Grade	
Total	0	0	3	2
Grand Total	3	1	3	2

The same process to identify the population of struggling readers was used in all schools. The school district utilized scores from the Developmental Reading Assessment, Second Edition (DRA2; Beaver, 2006) as reading benchmarks. In first and second grade the DRA2 was administered three times a year, September, January, and May. Reviewing the January DRA2 results, the lowest performing 20% of all first grade readers and the lowest performing 20% of all second grade readers were identified in each of the participating schools.

To qualify for participation, the 20% who were the lowest performing first and second grade readers were administered the three subscales of the *Dominie Reading and Writing Assessment Portfolio* (*Dominie*; DeFord, 2004). The *Dominie* is a new assessment being employed by the school district. Students with the most significant needs were then identified as struggling readers and targeted for participation in this study. The total number of students identified as struggling readers following the administration of the *Dominie* was 102 students, including 58 first graders and 44 second graders. These struggling readers had not been identified for special education under any categorization. All parents or guardians of these identified struggling readers received documentation describing the plan for this research study along with a permission slip to participate in the study. All parents were aware that participation in the study did not impact their child's participation in the reading support program.

Permission was received from 40 participants who were struggling readers across the seven schools. Prior to the beginning of the study, three schools were identified by district administration to utilize *Leveled Literacy Intervention* (*LLI*; Fountas & Pinnell, 2009) as their reading support program. Therefore, the subjects in these schools were identified as the

experimental group (n = 19). Four elementary schools were identified by district administration as the comparison group (n = 21). These schools continued to employ their current reading support program. Table 4 provides frequency and percentage of the gender for the experimental and comparison group participants.

Table 4

Frequency and Percentage of Gender for the Experimental and Comparison Group Participants

	Qualified	Sample	Experimental	Comparison
	102	40(39%)	19(48%)	21(53%)
First Grade	55(54%)	23(42%)	13(68%)	10(48%)
Males	35(64%)	15(65%)	9(69%)	6(60%)
Females	20(36%)	8(35%)	4(31%)	4(40%)
Second Grade	47(47%)	17(43%)	6(32%)	11(52%)
Males	35(74%)	12(71%)	4(67%)	8(73%)
Females	12(26%)	5(29%)	2(33%)	3(27%)
Total	102	40	19	21
Males	70(69%)	26(65%)	13(68%)	14(67%)
Females	32(31%)	14(35%)	6(32%)	7(33%)

Research Questions and Hypotheses

1. Is there a significant difference in literacy skills (core reading words, core writing words, phonemes, and spelling) between struggling readers who are part of a reading support program with an imbedded progress monitoring system and those who participate in a reading support program without an imbedded progress monitoring system?

   Directional Hypothesis: Students who participate in a reading support program with progress monitoring will demonstrate significantly higher mean scores on literacy skills (core reading words, core writing words, phonemes, and spelling) as compared to those who have participated in a reading support program without progress monitoring.

2. To what degree and in what manner do progress monitoring and reading self-concept predict literacy skills (core reading words)?

   Non-Directional Hypothesis: Literacy skills (core reading words) are significantly impacted by the reading support program (Progress Monitoring or no Progress Monitoring) and reading self-concept (difficulty in reading, competence in reading, attitudes toward reading).

Research Design

The quasi-experimental design employed to address Research Question 1 involved a nonrandomized control-group pretest-posttest design (Isaac & Michael, 1997). The design (see Figure 1) was used to examine the impact of the two levels of the independent variable, reading support program (Progress Monitoring and no Progress Monitoring), on the

dependent variable, literacy skills (core reading words, core writing words, sentence writing, and spelling) of struggling readers in first and second grades.

	Pretest	Treatment	Posttest
Experimental Group	O_1	X	O_2
Comparison Group	O_1		O_2

Figure 1. Research Question 1

A correlational design was employed in this research to respond to question two. This model was used to examine the effects of the reading support program (Progress Monitoring and no Progress Monitoring) and reading self-concept (difficulty with reading, competence in reading, attitudes toward reading) on students' literacy skills (core reading words). The *Readers' Self-Concept Scale* (Chapman & Tunmer, 1995b) was not given as a pretest for self-concept due to the long-term effects of this construct. In other words, self-concept was not expected to change over the time period for this study. To measure growth in readers' self-concept, the survey's authors stated that a minimum of six months was required. Unfortunately, this study was limited to a 14-week time period.

Description of the Experimental and Comparison Group Procedures

Prior to the beginning of the study, Language Arts Consultants (LACs) in all seven participating elementary schools were trained on the administration of all the subtests being utilized in the *Dominie.* LACs in this study were required to hold a Reading and Language Arts Consultant, K-12 certification (097-endorsement in Connecticut). LACs in both the experimental and comparison group schools received this training. Training was provided by the District K-6 Language Arts Curriculum Leader during two sessions for a total of four hours.

Once permission was obtained for struggling readers to participate in the study, they were administered the *Reader's Self-Concept Scale* (Chapman & Tunmer, 1995b). The scale was individually administered by the District K-6 Language Arts Curriculum Leader to students in the experimental and comparison groups.

The experimental group consisted of four Language Arts Consultants (LACs) who provided the treatment reading support program. The LACs were all certified teachers with a minimum of a Masters Degree in Reading. They had an average of 20 years experience in education. The LACs in the experimental group were trained in a reading support program with an imbedded progress monitoring system. The program utilized was *Leveled Literacy Intervention* (*LLI*; Fountas & Pinnell, 2009). The Language Arts Consultants for each of the three experimental schools were initially trained in the use of this program and the data collection involved. They met for a total of eight training sessions for two to three hours at a time, with the District K-6 Language Arts Curriculum Leader for continued training. During each biweekly training session, a model lesson, discussion, and data discussion occurred (See Appendix A).

Each session focused on one of the agenda topics (see Appendix A). The training occurred at one of the identified experimental schools. The coordinator provided a model lesson with a small group of participants from this identified school. The LAC working with these identified participants supplied the coordinator with student performance data prior to the lesson to provide for planning of an effective lesson (See Appendix B). All participating educators observed the lesson. Dialogue followed each lesson discussing the teaching point of the lesson including questions and comments. Student and teacher responses were discussed. This researcher was not involved in any of these meetings so as not to impact any

dialogues or results**.** LACs also brought progress monitoring data regarding weekly text levels of their students to each biweekly meeting (See Appendix C). This focused the discussion on each specific student's growth and instructional needs.

The comparison group included five Early Literacy Tutors (ELTs) who provided the current reading support program. The ELTs all held an elementary education teaching certification. One ELT was certified in business education with a background in finance and marketing. Another ELT had a background in counseling as well as certification as a reading specialist. The five ELTs had an average of 10 years experience in education.

Early Literacy Tutors implemented the reading intervention program currently being utilized in their schools either one-on-one or in small groups of two to three students. Groups met for 30 minutes, four times per week. These comparison groups were instructed utilizing the current reading support program in their schools. The programs varied for each of the schools dependent on the needs of their students. ELTs, in this study, were required to hold a K-6 teaching certification with extensive, demonstrated experience in the area of language arts. The educators in the comparison group did not receive training or utilize progress monitoring of specific literacy skills. Upon completion of the 14-week reading intervention program, all participants were re-administered the three subscales of the *Dominie.*

Instrumentation

There were three instruments utilized with each of the participants in this study. The district utilized the *Developmental Reading Assessment, Second Edition (DRA2;* Beaver, 2006) as the reading benchmark. The lowest 20% of students on the grade-level benchmark were then administered the *Dominie Reading and Writing Assessment Portfolio (Dominie;*

DeFord, 2004). Three of the eight subscales were administered to measure literacy skills and identify struggling readers with the most significant needs. For use in this study, the *Reading Self-Concept Scale (RSCS*; Chapman & Tunmer, 1995b) was utilized to measure the reading self-concept of each of the participants.

Developmental Reading Assessment, Second Edition (DRA2)

The DRA2, utilized in this study, is an individual reading assessment designed to assess students' reading performance in Grades K-3. This assessment provides teachers with information that helps them determine each student's independent reading level. With an identified independent reading level teachers are then able to diagnose students' instructional needs and plan for intervention or instruction as needed. Teachers are also able to observe their students' reading behaviors to gather data to inform their reading instruction.

Students are given a passage to read orally and are requested to retell the story with minimal prompting. Teachers are able to: (a) assess a student's oral reading skills, (b) analyze the strategies that a student uses to read unfamiliar text, (c) determine fluency and phrasing demonstrated during oral reading, and (d) monitor a student's ability to retell a story with minimal prompting. A variety of authentic fiction and non-fiction texts at each reading level are provided and have no time limits. The assessment gathers data on: (a) fluency, (b) vocabulary, (c) comprehension, (d) print concepts, and (d) reading engagement. Teachers complete a teacher observation guide with each reading of text a student completes. Responses are scored using a rubric that utilizes a continuum focusing on reading engagement, oral reading fluency, and comprehension. The guides present directions, questions, and prompts for each book, including story overviews. Responses are only scored

in terms of content. Structural and mechanical errors are ignored. These subscales are then used to place a student in an intervention, instructional, independent or advanced level.

Reliability and Validity of the Developmental Reading Assessment, Second Edition. The inter-rater agreement between the first two raters was .80. The inter-rater reliability was found to be (.74). Internal consistency data indicated high levels of consistency for the five items across all three raters (Cronbach's alpha = .98) and for DRA texts (.97).

To assess criterion-related validity, scores on the DRA were correlated with the *Iowa Test of Basic Skills* Subscales (ITBS) of: Vocabulary, Reading Comprehension, and Total Reading, in one suburban/urban school district. Utilizing Spearman's Rho formula, all correlations between the DRA Instructional Level and the ITBS subscales were significant at the $p < .01$ (2-tailed) level: (a) DRA with ITBS Total Reading $r = .71$; (b) DRA with ITBS Vocabulary $r = .68$; and (c) DRA with ITBS Reading Comprehension $r = .68$. With the use of Spearman's Rho, rank order correlation, however, the highest and most meaningful correlation was between the DRA and the Total Reading score ($r = .71$, $p < .01$).

Dominie Reading and Writing Assessment Portfolio (*Dominie*)

The *Dominie Reading and Writing Assessment Portfolio (Dominie;* DeFord, 2004), is an individually administered assessment tool. The *Dominie* has three alternate forms called A, B, and C. This study utilized form B for the pretest and form C for the posttest. The *Dominie* consists of eight subscales: (a) show me the book, (b) oral reading and comprehension, (c) phonemic awareness, (d) inventory of onsets and rimes, (e) core reading words from lists, (f) letter knowledge, (g) core writing words, and (h) sentence writing and spelling. Three subscales of the *Dominie* were used in the present study to measure literacy

skills of first and second grade struggling readers. The three subscales included core reading words, core writing words, and sentence writing and spelling, which took a total of 20 minutes to administer. The remaining subscales of the *Dominie* were not administered: Show Me the Book, phonemic awareness, inventory of onsets and rimes, oral reading and comprehension, along with letter knowledge.

These three subscales of the *Dominie* were utilized for a variety of reasons. Two of the subscales are to be administered to Kindergarteners; therefore, they are not administered to first and second grade students. These included Show Me the Book and Letter Knowledge. Two of the omitted subscales must be given one-on-one, requiring a great amount of time. Utilizing those subscales would have increased the time needed for administering assessments and decreased the time allotted for the study to occur. These include phonemic awareness (20 minutes) and inventory of onsets and rimes (10 minutes) totaling a minimum of 30 minutes for each participating student.

It was this researcher's intention to include the oral reading and comprehension subtest, where students must read accurately and fluently from passages of text. These data were collected for all students in the experimental group. The oral reading and comprehension subtest scores were only collected for a few of the students in the comparison group therefore; this variable was dropped from the MANOVA. Some of the educators felt that collecting fluency scores from first and second grade students was not appropriate. They believed that fluency should not be gathered until students were in third grade. Therefore, the oral reading and comprehension subtest scores were not utilized in this study.

While three subscales were administered, four scores were attained. To measure literacy skills involving reading, in the core reading words subscale students read from

graded word lists. To measure writing skills, two subscales were utilized: (a) core writing words and (b) sentence writing and spelling. In the core writing words subscale students are asked to write all words they know in 10 minutes, while the teacher may prompt categories of words students may know. For the sentence writing and spelling subscale students must write sentences from dictation. The dictation resulted in two scores one for phonemes and one for spelling, which were analyzed separately. All subscales resulted in raw scores also converted to stanines rankings.

Reliability and Validity of the *Dominie*. Reliability was assessed using alternate forms and time sampling. Reliability of sentence writing (phonemes) and spelling yielded reliability across Kindergarten through grade three. Second-grade correlations of fall and spring administrations of the *Dominie* for sentence writing (phonemes) and spelling, phonics (onset and rime), and text reading materials were all significant ($p < .01$) ranging from .59-.76 as a measure of reliability.

A two-year correlational study of all *Dominie* tasks showed high correlations ($p < .05$ or .01) to the Palmetto Achievement Challenge Test (PACT), the state of South Carolina proficiency test with high coefficient alphas as seen in Table 5. The study utilized a sample of 710 students matched as to reading achievement levels (high, middle, low). Intercorrelations between text reading and other measures (phonemes, spelling, onsets and rimes) ranged from .53 to .71 at a $p < .01$ level. The *Dominie* tasks showed high correlations (.57) to the PACT standardized test at $p < .01$ level, providing evidence of reliability and validity of the *Dominie*.

Table 5

Dominie Correlations to the Palmetto Achievement Challenge Test (PACT), and the State of South Carolina Proficiency Test

Subtasks	r
Phonemic Awareness	.76 - .86*
Phonics	.65 - .78*
Vocabulary	.92 - .93*
Comprehension	.78

*****Dependent on grade and form

Reading Self-Concept Scale (*RSCS*)

The *Reading Self-Concept Scale* (*RSCS*) by Chapman and Tunmer (1995b) was utilized to measure students' self-concepts as readers. This scale was specifically aimed at measuring the reading sub-component of academic self-concept. There are 30 items worded in simple question format. Questions were read aloud to each student individually, with administration of the entire survey typically taking 10 to 15 minutes. Sample questions included: Do you feel good when you do reading work? Are you good at correcting mistakes in reading? Children responded to each item along a 5-point scale, where 1 = no, never; 2 = no, not usually; 3 = undecided or unsure; 4 = yes, usually; and 5 = yes, always. Children were initially taught the response requirements of the scale through four examples and 10 practice items.

Considerable time was taken with the examples and practice items to ensure that the response requirements were fully understood. When the administrator was confident that the child understood the response requirements, the 30 *RSCS* items were then read aloud. Initially, a "yes" or "no" response to each item was sought. Probing of these responses then took place to elicit a response qualifier of "always" or "usually." Question number 20 was rewritten with author permission to make the vocabulary more meaningful to Connecticut students. Question 20 stated, "Do you feel stupid when you are asked to read?" The author agreed that the word "stupid" was no longer appropriate. Through email discussion he shared that it was written as a term children used to describe their feelings or to describe messages they received from significant others, such as teachers or other students. The author stated that synonyms such as unintelligent, dim-witted, dense, or thick were not helpful. He suggested any word that conveyed deep negativity about the students' ability would be "ok" to substitute. I suggested the use of sad, which the author agreed was the best age-appropriate term. Therefore, question 20 was changed to "Do you feel sad when you are asked to read?"

The author categorized the items into three aspects of reading self-concept: (a) attitudes towards reading, (b) perceptions of difficulty with reading, and (c) perceptions of competence in reading. The survey incorporated 10 questions for each of these subscales for a total of 30 questions on the survey. Each response is scored from a one (low reading self-concept) to a five (high reading self-concept). The sum of all responses was then calculated for each subscale. Each was then divided by the number of items (10) to obtain a mean subscale score for each of the three subscales. To produce a reading self-concept total score sum responses for all three subscales are calculated then divided by the number of total items

(30) to receive a total mean score. Each student receives a mean scale score ranging from 1 through 5 for each subscale and for a total reading self-concept.

Attitudes toward reading refer to the affective component of reading self-concept, which is defined in terms of having an affinity for reading (Chapman & Tunmer, 1995a). In this subscale, a low score means that a student does not feel good about reading, while a high score is interpreted to mean that the child enjoys and feels good about reading. Difficulty in reading refers to beliefs that reading activities are hard or problematic. A low score on this subscale means that a student perceives reading to be a difficult task, while a high score is interpreted to mean that the child believes reading is an easy activity. Competence in reading refers to beliefs regarding ability and proficiency in reading tasks. In this subscale, a low score means that a student perceives that he or she has low abilities in reading, while a high score is interpreted to mean that the child believes he or she has high abilities in reading.

Validity and Reliability of the *RSCS*. A Cronbach alpha for the entire instrument obtained a positive value above .80 at each age level with regard to the full scale internal consistency reliability. Internal consistency reliability estimates for each of the subscales ranged from .73 to .79 for attitudes toward reading, .72 to .78 for perceptions of difficulty with reading, and .69 to .75 for perceptions of competence in reading.

Full Scale and Subscale *RSCS* inter-correlations demonstrated the steady increase for competence (M = 3.60, 3.82, 3.75, 3.77, 3.79) and difficulty (M = 3.31, 3.34, 3.36, 3.16, 3.31) from Year 1 to Year 5. Correlations of full scale scores with the subscales of difficulty and attitude were weaker, though still significant ($p < .01$) ranging from .61 to .82. Competence (.82) demonstrated the most significant correlation with Full Scale *RSCS*, then

difficulty (.73) and attitude (.72) all at p< .01. Correlations between measures of reading and reading related performance are shown in Table 6.

Table 6

Reading Self-Concept Scale: Correlations with Reading and Spelling

Task	RSCS-30 Scales			
	Competence	Difficulty	Attitude	Full Scale
Year 1 (*n* = 143)				
Letter identification	.05	.22**	.10	.17*
Word identification	.08	.26**	.12	.22**
Pseudoword naming	.12	.23**	.14	.22**
Spelling	.01	.28**	.06	.17*
Year 4 (*n* = 99)				
Reading comprehension	.40**	.53**	.17	.47**
Year 5 (*n* = 103)				
Reading comprehension	.43**	.65**	.40**	.60**

* $p < .05$. ** $p < .01$. (Chapman & Tunmer, 1995b)

Description and Justification of the Analyses

Descriptive data for each dependent variable (core reading words, core writing words, phonemes, and spelling) are interpreted through means and standard deviations. This research used quantitative analyses to investigate each research question examining the same data. To avoid an inflated Type 1 error rate, a Bonferroni correction was used to adjust the alpha level of .05, which was divided by 2, resulting in an alpha level of .025 (Huck, 2008; Meyers, Gamst, & Guarino, 2006). Question One was analyzed using a Multivariate

Analysis of Variance ($p \leq .025$) employing the *Statistical Package for the Social Sciences* (SPSS, 2006). This statistical procedure determined if there were significant differences between group means of the experimental and the comparison group for the set of dependent variables.

The correlation design of question two was analyzed through a multiple regression analysis at $p \leq .025$. The *Statistical Package for the Social Sciences* (SPSS, 2006) 15.0 was used to analyze the data. The predictor variables of progress monitoring and reader's self-concept were analyzed for their impact on predicting the criterion variable of literacy skills as assessed by core reading words.

Data Collection Procedures and Timeline

The following procedures were followed according to the proposed timeline.

1. Approval was received from the Superintendent and building principals to conduct quasi-experimental research in selected elementary schools in the district (fall, 2008).
2. Approval was received from Western Connecticut State University's Institutional Review Board.
3. The Language Arts Consultants and the District K-6 Language Arts Curriculum Leader identified struggling readers in first and second grades in the seven elementary schools. Students were identified in the winter of 2009 utilizing the district benchmark assessment, *Developmental Reading Assessment 2* (Beaver, 2006).
4. The *Dominie* (DeFord, 2004) was administered to the identified struggling readers by the Language Arts Consultants. Readers with the most significant needs in the

experimental group were identified to participate in the *Leveled Literacy Intervention* (*LLI;* Fountas & Pinnell, 2009). Readers with the most significant needs in the comparison groups were identified to participate in each school's current reading intervention program with no progress monitoring imbedded into the program (winter, 2009).

5. Consent forms were distributed to the parents of the identified struggling readers in the experimental and comparison groups and then collected (winter, 2009).
6. The researcher collaborated with the District K-6 Language Arts Curriculum Leader to plan initial and biweekly training in progress monitoring (winter & spring 2009). See Appendix A for a complete schedule.
7. The researcher gathered pretest data using the *Dominie* (DeFord, 2004; winter, 2009).
8. The researcher gathered school and participating students' demographic information for the study (winter, 2009). These data included age, gender, ethnicity and free/reduced lunch status.
9. The District K-6 Language Arts Curriculum Leader administered the *Reading Self-Concept Scale* (*RSCS)* to all but three participants in the study. One LAC administered the *Reading Self-Concept Scale* (*RSCS)* to these three participants. The researcher gathered these data from all study participants (winter, 2009).
10. Trained Language Arts Consultants implemented *Leveled Literacy Intervention* (*LLI;* Fountas & Pinnell, 2009), a reading support program with an imbedded progress monitoring system in groups of one LAC to three students, for 30

minutes a day, 4 to 5 times a week for 14 weeks (February through early June, 2009).

11. Language Arts Consultants and Early Literacy Tutors implemented each school's reading support program without progress monitoring working in small groups for 30 minutes a day, 4 to 5 times a week for 14 weeks (February through early June, 2009; See Appendix B for complete agendas).
12. The researcher gathered the posttest data of the *Dominie* (DeFord, 2004) from all students in the study (spring, 2009).

Limitations of the Study

According to Campbell and Stanley (1963), a nonequivalent control group design often naturally controls the main effects for six of the eight internal validity sources: history, maturation, testing, instrumentation, selection and mortality. Historically, to this researcher's knowledge, there were no events, which occurred between the pre- and post- testing during the 14-week study which could have impacted the results of the post- testing. To control for maturation, the length of the study was held to 14 weeks within a single school year minimizing the possible biological and psychological changes within the subjects.

The use of the *Dominie's* (DeFord, 2004) three alternate forms (A, B, & C) per subscale supports the internal validity for testing since pre and posttest forms were different. Form A was not utilized for this study. Form B was used for all pretesting, while form C was utilized for all post- testing. Internal validity for instrumentation of the *Reader's Self-Concept Scale* (Chapman & Tunmer, 1995b) was controlled by having one administrator for all but three of the participating students. One LAC administered the *Reader's Self-Concept*

Scale to these three participating students. All Language Arts Consultants were trained simultaneously as a group. Follow-up discussions were held to clarify any questions.

Selecting students from the lower 20th percentile assisted in controlling for selection but may have restricted the variance for the data analysis. Specific and consistent data collection by the District K-6 Language Arts Curriculum Leader and the researcher assisted in controlling for low mortality. Statistical regression is of concern when using extreme scores, which often automatically go higher on the second test (posttest). The use of a comparison group helped to control for this limitation.

One identified limitation of this study is the reliability of treatment implementation. This relates to the necessity of the treatment being implemented in the same way at each site (Isaacs & Michael, 1997). To minimize differences in the effects of treatment among specialists and across sites and time, all participating Language Arts Consultants in the experimental group were trained to follow a specific reading intervention program, *Leveled Literacy Intervention (LLI;* Fountas & Pinnell, 2009), with an imbedded progress monitoring system (See Appendix B). They also met bi-weekly with the trainer, the District K-6 Language Arts Curriculum Leader, to discuss implementation of the model along with specific questions and concerns, as well as the next steps in the process.

The implementers of the progress monitoring model were all Language Arts Consultants (LAC) in the district who held educational certification in this specific field of study. All instructional groups consisted of three students working with the LAC. This assisted in controlling for changes in instrumentation, raters, and observers since there were multiple educators working with the students.

The comparison groups received a reading support program, which did not include a progress monitoring model. These students (n = 21) were administered the reading support program currently employed at their school either by LACs or Early Literacy Tutors (ELT), holding K-6 teaching certifications with extensive, demonstrated experience in the area of language arts. Regarding the comparison group, 38% of the students received instruction in a one-on-one model, while 62% received instruction in groups with two to three students.

Demoralization of respondents, which occurs when members of a group not receiving treatment perceive they are inferior and give up, is a threat to a study employing a treatment. To control for this outcome, all educators were given an overview of the *Leveled Literacy Intervention* (*LLI;* Fountas & Pinnell, 2009) program, and it is planned that all instructors in the comparison group will be trained in the *Leveled Literacy Intervention* (*LLI;* Fountas & Pinnell, 2009) program with progress monitoring following the completion of the study for implementation in the following school year.

Ethics Statement

Permission to participate in this research was sought from the district's superintendent and each school principal. Permission was sought from all parents of participating students identified as struggling readers. The District K-6 Language Arts Curriculum Leader completed human subjects training certification. To assure confidentiality, each participant was assigned a confidential identification number. All data were stored in a locked filing cabinet in the researcher's home or office and will be maintained there until the findings have been published, accessible only to other researchers for whom the data will prove useful in further comparative analyses and who are professors

related to or students enrolled in Western Connecticut State University's Doctor of Education in Instructional Leadership Program.

CHAPTER FOUR: ANALYSIS OF DATA AND FINDINGS

This study investigated the use of progress monitoring on students' literacy skills. In addition, the study examined the relationship between progress monitoring, reading self-concept, and students' posttest core reading words scores. The two research questions addressed in this study were:

1. Is there a significant difference in literacy skills (core reading words, core writing words, phonemes and spelling) between struggling readers who are part of a reading support program with an imbedded progress monitoring system and those who participate in a reading support program without an imbedded progress monitoring system?

   Directional Hypothesis: Students who participate in a reading support program with progress monitoring will demonstrate significantly higher mean scores on literacy skills (core reading words, core writing words, phonemes and spelling) as compared to those who have participated in a reading support program without progress monitoring.

2. To what degree and in what manner do progress monitoring and reading self-concepts predict literacy skills (core reading words)?

   Non-Directional Hypothesis: Literacy skills (core reading words) are significantly impacted by reading support program (Progress Monitoring or no Progress Monitoring) and reading self-concept (difficulty in reading, competence in reading, attitudes toward reading).

This chapter presents the results of this study. A section on the description of the data will be followed by two sections organized by research question. Research Question 1

details pretest and posttest data preparation, which identified three outliers. These outliers were removed and detailed in sections titled pretest data with outliers removed and posttest data with outliers removed. This is followed by data analysis of Research Question 1. Results related to Research Question 2 include data preparation and data analysis.

Description of the Data

This study utilized interval-level data from the *Dominie Reading and Writing Assessment Portfolio (Dominie)* and the *Reading Self-Concept Scale (RSCS).* The *Dominie* consists of eight subtests: (a) show me the book, (b) oral reading and comprehension, (c) phonemic awareness, (d) inventory of onsets and rimes, (e) reading words from lists, (f) letter knowledge, (g) core writing words, and (h) sentence writing, and spelling. The eighth subscale, sentence writing and spelling, yields two separate sets of raw and stanines scores, one set for phonemes and another set for spelling. Therefore, the *Dominie* yields a total of nine separate raw and stanines scores. The current study utilized the raw scores of the following subtests: (a) reading words from lists (core reading words), (b) core writing words, (c) phonemes, and (d) spelling.

The *RSCS* utilized a 5-point Likert scale to identify a total Reading Self-Concept Score as well as scores for each of the three subscales: (a) Perceptions of Difficulty with reading, (b) Attitudes toward reading, and (c) Perceptions of Competence in reading. For the purpose of this study, the total Reading Self-Concept Score was utilized.

Each of the two research questions examined the same data. To avoid an inflated Type 1 error rate, a Bonferroni correction was used to adjust the alpha level of .05, which was divided by 2, resulting in an alpha level of .025 (Huck, 2008; Meyers et al., 2006). The alpha level of .025 was employed for all statistical tests.

Research Question 1

Pretest Data Preparation. Research Question 1 examined the impact of literacy skills (core reading words, core writing words, phonemes, and spelling) for students who participated in a reading support program with imbedded progress monitoring and those who participated in a reading support program without an imbedded progress monitoring system. Pretest data were collected to examine differences between the experimental and comparison groups previous to the intervention.

The *Statistical Package for the Social Sciences* (*SPSS*; Pallant, 2007) was used to analyze descriptive statistics, stem and leaf graphs and histograms. An evaluation of the data from the total sample (n = 40) was conducted. Missing data were identified resulting in varying sample sizes for each variable: (a) Core reading words(n = 36), (b) Core writing words (n = 34), (c) Phonemes (n = 39), and (d) Spelling (n = 39). Pallant (2007) and Meyers et al. (2006), support SPSS in the use of excluding cases pairwise, which accommodates for excluding a case in each variable only if it was missing the data required for the specific analysis. Cases are included in any of the analyses for which they have the necessary information or data.

Descriptive Statistics. Descriptive statistics were utilized to examine the pretest raw scores for core reading, core writing words, phonemes, and spelling assuming equal variances between the groups. Descriptive statistics for the total sample are presented in Table 7.

Table 7

Descriptive Statistics for Core Reading Words, Core Writing Words, Phonemes, and Spelling Pretest Scores

	Core Reading Words	Core Writing Words	Phonemes	Spelling
n				
Experiment	19	19	19	19
Comparison	17	15	20	20
Mean				
Experiment	15.53	26.68	50.26	8.84
Comparison	18.12	22.33	52.30	9.95
Standard Deviation				
Experiment	5.36	9.63	9.90	2.22
Comparison	4.55	9.37	9.50	1.66
Skewness				
Experiment	0.01	0.13	0.73	-0.15
Comparison	-0.23	-0.09	0.06	0.40
Kurtosis				
Experiment	-0.14	0.85	-0.83	-0.42
Comparison	-0.29	1.22	-1.92	-0.17

Outliers and Data Normality. An evaluation of univariate and multivariate outliers were conducted for further assessment of data normality. The normality of the distribution of raw scores for core reading words, core writing words, phonemes and spelling for the experimental and comparison groups were examined. According to Tabachnick and Fidell (1989), normality is assessed through the use of graphical (stem-and-leaf plots, and histograms) or statistical (skewness and kurtosis values) methods. Stevens (2002) states "with small or moderate sample sizes, it is difficult to tell whether the nonnormality is real or apparent, because of considerable sampling error" (p. 264). Therefore, he recommends the statistical test.

This researcher initially inspected stem-and-leaf plots and histograms for all four variables. There were no extreme values found in the experimental or comparison groups. Confirming the stem-and-leaf findings, as recommended by Meyers et al. (2006), box plots for all variables for the experimental and comparison groups did not show any evidence of possible outliers.

As recommended by Stevens (2002) and Meyers et al. (2006), statistical tests of normality were then analyzed. Statistical tests examined were skewness and kurtosis coefficients followed by the Shapiro-Wilk test. All skewness values were within the ± 1 range of normality. Two kurtosis values, core writing words and phonemes for the comparison group, exceeded the ± 1 range of normality.

To further examine the out-of-range kurtosis values Stevens, (2002) recommends that "sample sizes ranging from 10 to 50 that the combination of skewness and kurtosis coefficients and the Shapiro-Wilk test were the most powerful in detecting departures from normality" (p. 264). In other words, the Shapiro-Wilk test produces one value that combines

the skewness and kurtosis coefficients. Stevens (2002) also stated that with sample sizes less than 20, the Shapiro-Wilk procedure can detect extreme non-normality. To support Stevens' contention, Tabachnick and Fidell (1989) stated that ".01 or .001 alpha levels are used to evaluate the significance of value for [the combined value of] skewness and kurtosis with small to moderate samples" (p. 73). Table 8 displays the significance of the Shapiro-Wilk test, indicating that all results are adequate at the $p < .001$ level.

Table 8

Shapiro-Wilk Test of Normality for the Pretest Scores

	Significance
Core Reading Words	
Experiment	.707
Comparison	.784
Core Writing Words	
Experiment	.727
Comparison	.334
Phonemes	
Experiment	.017
Comparison	.003
Spelling	
Experiment	.425
Comparison	.563

Mahalanobis distances were calculated to test for multivariate normality. The Mahalanobis distance was computed using a chi-square criterion of 4 degrees of freedom at a $p < .001$ confidence level, resulting in a critical value of 18.47 (Meyers et al., 2006). The maximum Mahalanobis distance value was 8.335, well below the critical value.

Correlations. Pearson product-moment correlation coefficients were analyzed to determine the relationship between the variables. Refer to Table 9 for the Pearson product-moment correlations between literacy skills variables of pretest data. Pallant (2007) states "Correlations up around .8 or .9 are reason for concern" (p. 282). Since none of the coefficients were higher than .66, data analysis was executed.

Table 9

Pearson Product-Moment Correlations between Literacy Skills Variables of Pretest Data

	Core Reading Words	Core Writing Words	Phonemes	Spelling
Core Reading Words	-	.025	.658**	.507**
Core Writing Words		-	.113	-.068
Phonemes			-	.562**
Spelling				-

**p < .01 (2-tailed)

Homogeneity of Variance. A Levene's test was analyzed to check the assumption of equal variance across both experiment and comparison groups. The Levene's statistic was not significant indicating equal variance in the groups. Refer to Table 10 for the Levene's Test of Equality of Error Variances of Pretest Scores. A multivariate analysis was utilized

for this study; therefore, Box's Test of Equality of Covariance Matrices examined the covariance between the dependent variables. These results were not significant indicating that the assumptions of homogeneity were met and the matrices were equal. Refer to table 11 for the Box's Test of Equality of Covariance Matrices of Pretest Scores.

Table 10

Levene's Test of Equality of Error Variances of Pretest Scores

	Levene Statistic	df1	df2	Significance
Core Reading Words	.039	1	32	.844
Core Writing Words	.022	1	32	.998
Phonemes	.730	1	32	.889
Spelling	.275	1	32	.882

Table 11

Box's Test of Equality of Covariance Matrices of Pretest Scores

Box's M	10.963
F	.943
df1	10.000
df2	4275.539
Sig.	.492

Data Analysis. For the purpose of this study, the Wilks' Lambda was reported to compare the means of the experimental and comparison groups. Results indicated no significant difference between the mean scores for the dependent variables: (a) core reading words, (b) core writing words, (c) phonemes, and (d) spelling. Refer to Table 12 for the results of a multivariate analysis of variance comparing experimental and comparison groups pretest.

Table 12

Results for a Multivariate Analysis of Variance Test Comparing Experimental and Comparison Groups Pretest Scores

Multivariate Test	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilks' Lambda	.743	2.507[a]	4.000	29.000	.064	.257

a. Exact statistic

Posttest Data Preparation. Research Question 1 focused on the impact of progress monitoring for struggling readers who participated in a reading support program with an imbedded progress monitoring system and those who participated in a reading support program without a progress monitoring system, on students' literacy skills. Posttest data were collected after the treatment period from both the experimental and comparison groups.

Data were reviewed for accuracy and missing information. Missing data were identified on a number of students. This impacted the group size for each variable, the total sample contained 40 subjects with data collected for: core reading words (n = 34), core writing words (n = 34), phonemes (n = 37), and spelling (n = 37). Descriptive statistics, stem-and-leaf plots and histograms were analyzed to screen all data from the sample (n =

40). Data were screened for missing values, outliers, and violations of statistical assumptions.

Descriptive Statistics. Descriptive statistics were utilized to examine the posttest raw scores for core reading words, core writing words, phonemes, and spelling, assuming equal variances between the groups. Descriptive statistics for the total sample are presented in Table 13.

Table 13

Descriptive Statistics for Core Reading Words, Core Writing Words, Phonemes, and Spelling Posttest Scores

	Core Reading Words	Core Writing Words	Phonemes	Spelling
n				
Experimental	19	19	19	19
Comparison	15	15	18	18
Mean				
Experimental	21.63	38.79	52.58	12.26
Comparison	22.73	40.33	55.22	11.89
Standard Deviation				
Experimental	3.50	14.90	10.61	8.84
Comparison	2.96	12.40	11.11	2.08
Skewness				
Experimental	-1.47	1.14	0.94	3.70
Comparison	-1.23	0.10	-0.07	0.52
Kurtosis				
Experimental	3.88	0.26	-0.52	15.01
Comparison	2.48	-1.32	-1.13	1.32

Outliers and Data Normality. An evaluation of univariate and multivariate outliers was conducted for further assessment of data normality. The normality of the distribution of raw scores for core reading words, core writing words, phonemes, and spelling for the experimental and comparison groups were examined. Stem-and-leaf plots and histograms were inspected for all four variables. There were no extreme values found in the experimental or comparison groups in phonemes. Core writing words had no extreme values in the comparison group, though there were three outliers identified for the experimental group. Core reading words had one outlier identified for the experimental group. Spelling had one outlier identified for the experimental and one for the comparison group.

As a follow-up, skewness and kurtosis levels were analyzed with some significant results. As it is appropriate to use the Shapiro-Wilk test to locate values significant at the .001 level or higher, Table 14 displays all the values within range except for the variable of spelling for the experimental group.

Table 14

Shapiro-Wilk Test of Normality of Posttest Scores

	Significance
Core Reading Words	
Experiment	.023
Comparison	.094
Core Writing Words	
Experiment	.007
Comparison	.148
Phonemes	
Experiment	.003
Comparison	.011
Spelling	
Experiment	.000
Comparison	.399

Mahalanobis distances were calculated to test for multivariate normality. The Mahalanobis distance was computed using a chi-square criterion of 4 degrees of freedom at a $p < .001$ confidence level, resulting in a critical value of 18.47 (Meyers et al., 2006). The maximum Mahalanobis distance value was 28.729, above the critical value, suggesting multivariate outliers.

To improve skewness and kurtosis, the use of box plots were analyzed to identify extreme scores or outliers, as recommended by Meyers et al. (2006). Outliers were identified

for the variables of core reading words (1), spelling (2), and core writing words (3). In the experimental group, one student demonstrated high posttest scores across all four variables, especially in the area of spelling. For the experimental group, in spelling M = 12.26 with SD = 8.837, the student's score was a 47, which accounts for the 15.01 kurtosis level for the experimental group (Refer to Table 13). This student was also identified as having an extreme score in the core writing words (65 raw score), M = 38.79, SD=14.902; therefore, this student was removed from future analyses. One student in the comparison group was identified who demonstrated slightly higher scores for three of the variables with a raw score of 17 in spelling, M= 11.89, SD = 2.083. This student was also removed from future analyses. One student was identified for having a low score in core reading words in the comparison group (11 raw score), M = 22.73, SD = 2.915; therefore, this student was removed from future analyses. All data from these three students were removed from the data set. All pretest and posttest data were reanalyzed with the outliers removed. Two students identified as outliers in core writing words from the experimental group were not removed to maintain comparable group sizes.

Pretest Data with outliers removed. As previously stated, outliers were identified consequently; all data from three students were removed from the data set to improve skewness and kurtosis values of normality. Data were reviewed for accuracy and missing values. Missing data were identified resulting in varying group size for each variable, the total sample (n = 37), core reading words (n = 33), core writing words (n = 31), phonemes (n = 36), and spelling (n = 36). Pallant (2007) and Meyers et al. (2006) support *SPSS* in the use of excluding cases pairwise, which accommodates for excluding a case in each variable only

if it was missing the data required for the specific analysis. Cases are included in any of the analyses for which they have the necessary information or data.

Descriptive Statistics with outliers removed. Descriptive statistics were examined for the pretest raw scores of core reading, core writing words, phonemes, and spelling assuming equal variances between the groups, once outliers were removed. Descriptive statistics for the total sample with outliers removed are presented in Table 15.

Table 15

Descriptive Statistics for Core Reading Words, Core Writing Words, Phonemes, and Spelling Pretest Scores with outliers removed

	Core Reading Words	Core Writing Words	Phonemes	Spelling
n				
Experimental	17	17	17	17
Comparison	16	14	19	19
Mean				
Experimental	16.35	27.06	51.12	8.88
Comparison	17.75	22.43	51.63	9.74
Standard Deviation				
Experimental	4.96	9.69	9.99	2.29
Comparison	4.44	9.72	9.26	1.66
Skewness				
Experimental	0.00	-0.16	0.65	-0.19
Comparison	-0.16	-0.12	0.16	0.14
Kurtosis				
Experimental	0.27	-0.72	-1.11	-0.42
Comparison	-0.05	-1.38	-1.91	-0.46

Outliers and Data Normality with outliers removed. An evaluation of univariate and multivariate outliers was conducted for further assessment of data normality. The normality of the distribution of raw scores for core reading words, core writing words, phonemes, and spelling for the experimental and comparison groups were examined. Stem-and-leaf plots and histograms were inspected for all four variables. There were no extreme values found in the experimental or comparison groups. Box plots for all variables in the experimental and comparison groups showed no evidence of possible outliers, confirming the stem-and-leaf findings.

Skewness and kurtosis levels were analyzed with some significant results. Refer to Table 14 for skewness and kurtosis levels of pretest scores with outliers removed. All skewness values were within the ±1 range of normality. Three kurtosis values: (a) core writing words for comparison group, (b) phonemes for comparison group, and (c) phonemes for experimental group, exceeded the ±1 range of normality. Follow-up analysis utilizing the Shapiro-Wilk with all results above the .001 alpha levels can be seen in Table 16.

Table 16

Shapiro-Wilk Test of Normality of Pretest Scores with outliers removed

	Significance
Core Reading Words	
Experimental	.834
Comparison	.875
Core Writing Words	
Experimental	.727
Comparison	.334
Phonemes	
Experimental	.018
Comparison	.003
Spelling	
Experimental	.385
Comparison	.591

Mahalanobis distances were calculated to test for multivariate normality. The Mahalanobis distance was computed using a chi-square criterion of 4 degrees of freedom at a $p < .001$ confidence level, resulting in a critical value of 18.47 (Meyers et al., 2006). The maximum Mahalanobis distance value is 7.991, below the critical value of 18.47, assuming there are no multivariate outliers.

Correlations with outliers removed. Pearson product-moment correlation coefficients were analyzed to ensure no violation of linearity between the dependent variables, refer to Table 17.

Table 17

Pearson Product-Moment Correlations between Core Reading Words, Core Writing Words, Phonemes, and Spelling of Pretest Data with outliers removed

	Core Reading Words	Core Writing Words	Phonemes	Spelling
Core Reading Words	-	.003	.615**	.451**
Core Writing Words		-	.108	-.092
Phonemes			-	.510**
Spelling				-

**$p < .01$ (2-tailed)

Significant relationships ($p < .01$) are identified between core reading words and phonemes as well as with spelling. An additional significant relationship ($p < .01$) between phonemes and spelling is also identified. Though significant, these relationships are all at the moderate level. Pallant (2007) identified that relationships at the .8 or .9 level are of concern. According to Meyers et al. (2006), moderately correlated dependent variables are the ideal situation for a MANOVA procedure.

Homogeneity of Variance with outliers removed. A Levene's test for homogeneity of variance was analyzed to check the assumption of equal variances across both

experimental and comparison groups. Refer to Table 18 for the Levene's statistic, which was not significant ($p < .025$) indicating equal variance in the groups. A multivariate analysis was utilized for this study; therefore, Box's Test of Equality of Covariance matrices examined the covariance between the dependent variables. These results were not significant ($p < .025$) indicating that the assumptions of homogeneity were met and the matrices were equal, refer to Table 19.

Table 18

Levene's Test of Equality of Error Variances of Pretest Scores with outliers removed

	Levene Statistic	df	Significance
Core Reading Words	.002	29	.966
Core Writing Words	.246	29	.624
Phonemes	.730	29	.400
Spelling	.275	29	.604

Table 19

Box's Test of Equality of Covariance Matrices of Pretest Scores with outliers removed

Box's M	12.079
F	1.023
df1	10.000
df2	3660.167
Sig.	.421

***Data Analysis*.** For the purpose of this study the Wilks' Lambda was reported to compare the means of the experimental and comparison groups. Results indicated no significant difference between the mean scores for the dependent variables: (a) core reading, (b) core writing words, (c) phonemes, and (d) spelling. Refer to Table 20 for the results of a multivariate analysis of variance test comparing experimental and comparison groups of pretest scores with outliers removed.

Table 20

Results for a Multivariate Analysis of Variance Test Comparing Experimental and Comparison Groups with outliers removed

Multivariate Test	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilks' Lambda	.779	1.847[a]	4.000	26.000	.150	.221

a. Exact statistic

Posttest Data with outliers removed. Data were reviewed for accuracy and missing values. Missing data were identified, impacting the group size for each variable with the total sample (n = 34), core reading words (n = 31), core writing words (n = 31), phonemes (n = 34), and spelling (n =34).

Descriptive Statistics with outliers removed. Descriptive statistics were utilized to examine the posttest raw scores with outliers removed for core reading, core writing words, phonemes, and spelling assuming equal variances between the groups once outliers were removed. Descriptive statistics for the total sample are presented in Table 21.

Table 21

Descriptive Statistics for Core Reading Words, Core Writing Words, Phonemes and Spelling Posttest Scores with outliers removed

	Core Reading Words	Core writing Words	Phonemes	Spelling
N				
Experimental	17	17	17	17
Comparison	14	14	17	17
Mean				
Experimental	22.470	38.790	53.590	10.590
Comparison	22.570	40.330	54.350	11.590
Standard Deviation				
Experimental	2.267	14.902	10.730	2.647
Comparison	2.954	12.396	10.799	1.698
Skewness				
Experimental	0.098	1.144	0.810	0.530
Comparison	-1.139	0.101	0.008	-0.302
Kurtosis				
Experimental	-0.080	0.257	-0.837	-0.406
Comparison	2.341	-1.320	-1.014	0.456

Outliers and Data Normality with outliers removed. An evaluation of univariate and multivariate outliers was conducted for further assessment of data normality. The normality

of the distribution of raw scores for core reading words, core writing words, phonemes, and spelling for the experimental and comparison groups were examined. Stem-and-leaf plots and histograms were inspected for all four variables. There were no extreme values found in the experimental or comparison groups in core reading, phonemes, and spelling. Core writing words had no extreme values in the comparison group, though there were three values identified in the experimental group. Box plots confirmed three outliers in the experimental group for core writing words.

Skewness and kurtosis levels were analyzed with some significant results. Kurtosis levels for the comparison group were above the ± 1 for the core writing words variable. As stated earlier, for small groups the use of skewness and kurtosis levels along with the Shapiro-Wilk procedure is powerful in identifying normality. See Table 22 for follow-up results utilizing the Shapiro-Wilk, with all values above the .001 significance cutoff.

Table 22

Shapiro-Wilk Test of Normality of Posttest Scores with outliers removed

	Significance
Core Reading Words	
Experimental	.361
Comparison	.180
Core Writing Words	
Experimental	.010
Comparison	.094
Phonemes	
Experimental	.006
Comparison	.013
Spelling	
Experimental	.469
Comparison	.523

Mahalanobis distances were calculated to test for multivariate normality. The Mahalanobis distance was computed using a chi-square criterion of 4 degrees of freedom at a $p < .001$ confidence level, resulting in a critical value of 18.47 (Meyers et al., 2006). The maximum Mahalanobis distance value is 12.807, below the critical value of 18.47.

Correlations with outliers removed. Pearson product-moment correlation coefficients were analyzed to ensure no violation of linearity between the dependent

variables. Refer to Table 23 for all correlations, which are below the .80 or .90 area of concern stated by Pallant (2007).

Table 23

Pearson Product-Moment Correlations between Core Reading Words, Core Writing Words, Phonemes, and Spelling of Posttest Data with outliers removed

	Core Reading Words	Core Writing Words	Phonemes	Spelling
Core Reading Words	-	.375*	.308	.319
Core Writing Words		-	.116	.447*
Phonemes			-	.024
Spelling				-

**p < .05 (2-tailed)

Significant relationships ($p < .05$) are identified between core reading words and core writing words. An additional significant relationship ($p < .05$) between core writing words and spelling is also identified. Though significant, these relationships are all at the moderate level, therefore, not of concern.

Homogeneity of Variance with outliers removed. A Levene's test for homogeneity of variance was analyzed to check the assumption of equal variance across both experimental and comparison groups. Refer to Table 24 for the Levene statistic, which was not significant at $p < .025$, indicating equal variance in the groups. A multivariate analysis was utilized for this study; therefore, Box's Test of Equality of Covariance Matrices was examined. Refer to

Table 25 for these results, which were not significant ($p < .025$) indicating that the dependent variable covariance matrices are equal across the levels of the independent variable.

Table 24

Levene's Test of Equality of Error Variances of Posttest Scores with outliers removed

	Levene Statistic	df	Significance
Core Reading Words	0.796	26	.380
Core Writing Words	12.752	26	.001
Phonemes	1.166	26	.290
Spelling	2.671	26	.114

Table 25

Box's Test of Equality of Covariance Matrices of Posttest Scores with outliers removed

Box's M	17.843
F	1.484
df1	10.000
df2	3231.873
Sig.	.139

Data Analysis. Preliminary assumption testing was conducted after addressing all violations noted. This allowed for a one-way between groups multivariate analysis of variance (MANOVA) to be performed to investigate the use of a reading support program with an imbedded progress monitoring program on struggling readers' literacy skills (n = 28). Four dependent variables were included: core reading words (n = 14), core writing words (n = 14), phonemes (n = 14), and spelling (n = 14). The independent variable was the use of a reading support program with an imbedded progress monitoring program or a reading support program with no progress monitoring.

For the purpose of this study, the Wilks' Lambda statistic was reported, since it is the most typically reported according to Meyers et al. (2006). Results indicated no statistically significant difference between struggling readers in a reading support program with an imbedded progress monitoring program and those who were in a reading support program without progress monitoring, $F(4, 23) = 1.47$, ns. Refer to Table 26 for the results for a multivariate analysis of variance test comparing experimental and comparison groups. Table 27 reviews the mean scores of the four dependent variables indicating slightly higher scores for struggling readers in the comparison group than for those in the experimental group for all four dependent variables.

Table 26

Results for a Multivariate Analysis of Variance Test Comparing Experimental to Comparison Group Posttest Scores with outliers removed

Multivariate Test	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilks' Lambda	.796	1.467[a]	4.000	23.000	.245	.203

Table 27

Mean Scores for Core Reading Words, Core Writing Words, Phonemes, and Spelling for Experimental and Comparison Groups Posttest with outliers removed

	Core Reading Words $n = 28$	Core Writing Words $n = 28$	Phonemes $n = 28$	Spelling $n = 28$
Mean				
Experimental	22.47 ($n = 14$)	38.79 ($n = 14$)	53.59 ($n = 14$)	10.59 ($n = 14$)
Comparison	22.57 ($n = 14$)	40.33 ($n = 14$)	54.35 ($n = 14$)	11.59 ($n = 14$)

Research Question 2

Data Preparation. Research Question 2 examined the impact of a reading support program with an imbedded progress monitoring program and reading self-concept on a student's literacy skills (core reading words). The *Statistical Package for the Social Sciences* (*SPSS*; Pallant, 2007) was used to analyze assumptions as well as interpret the analysis of the multiple regressions. Stevens (2002) recommends that a sample size of about 15 per predictor is needed for a reliable equation. The sample size of progress monitoring ($n = 40$) and reading self-concept ($n = 38$) meet this criterion.

Descriptive Statistics. Descriptive statistics for the total sample are presented in Table 28. Refer to Table 29 and 30 for the descriptive statistics for the experimental and comparison.

Table 28

Descriptive Statistics for Reading Support Group, Reading Self-Concept, and Core Reading Words for the Total Sample

	Mean	Standard Deviation	N
Reading Support Group	0.480	0.506	40
Reading Self-Concept	3.632	0.339	38
Core Reading	22.120	3.255	34

Table 29

Descriptive Statistics for and Reading Self-Concept for the Experimental group receiving Progress Monitoring

	Core Reading (n = 19)	Reading Self-Concept (n = 17)
Mean	21.63	3.58
Std. Deviation	3.50	0.35
Skewness	-1.47	0.24
Standard error	0.52	0.55
Kurtosis	3.88	0.95
Standard error	1.01	1.06

Table 30

Descriptive Statistics for Core Reading Words and Reading Self-Concept for the Comparison group not receiving Progress Monitoring

	Core Reading Words (n = 15)	Reading Self-Concept (n = 21)
Mean	22.73	3.68
Std. Deviation	2.92	0.31
Skewness	-1.23	0.06
Standard error	0.58	0.51
Kurtosis	2.48	-0.97
Standard error	1.12	0.97

Outliers and Data Normality. Mahalanobis distances were calculated to test for outliers. The Mahalanobis distance was computed using a chi-square criterion of 2 degrees of freedom at a $p < .001$ confidence level, resulting in a critical value of 13.82 (Meyers et al., 2006; Pallant, 2007). The maximum Mahalanobis distance value was 7.378, below the 13.82 critical values.

Pearson Correlations were analyzed to examine the correlation of the two predictors, progress monitoring and reading self-concept. Refer to Table 31 for Pearson Product-Moment Correlation results. Meyers et al. (2006) stated that correlations above .75 are

highly correlated and of concern. Multicollinearity is not a problem in this analysis since all coefficients were below .20.

Table 31

Pearson Product-Moment Correlations of Core Reading Words, Progress Monitoring, and Reading Self-Concept in the Multiple Regression

	Core Reading Words	Reading Support Program	Reading Self-Concept
Core Reading Words	-	-0.17	0.06
Reading Support Program		-	-0.15
Reading Self-Concept			-

**p < .05 (2-tailed)

To further analyze problems with multicollinearity which may not be evident on the correlation matrix, the Tolerance and Variance inflation factors (VIF) were examined. Refer to Table 32 for coefficients collinearity statistics. Tolerance "is the amount of a predictor's variance not accounted for by the other predictors" (Meyers et al., 2006, p. 182). Pallant (2007) and Meyers et al. (2006) identify values below .10 as suggesting the possibility of multicollinearity. VIF "is the reciprocal of the tolerance and measures the degree of linear association between particular independent variables and the remaining independent variables in the analysis" (Meyers et al., 2006, p. 212). Pallant (2007) and Meyers et al. (2006) identify values above 10 as indicating a concern with multicollinearity. All values for this analysis were within the expected range.

Table 32

Coefficients of Collinearity Statistics for Progress Monitoring and Reading Self-Concept

	Tolerance	VIF
Progress Monitoring	0.98	1.02
Reading Self-Concept	0.98	1.02

Meyers et al. (2006) stated a viable means for detecting statistical assumptions violations is the analysis of residual scatterplot. Refer to Table 33 for a Casewise diagnostics table, which identified one case that had a standardised residual value above the identified value 3.0, as stated by Meyers et al.

Table 33

Casewise Diagnostics depicting residual outliers

Case Number	Standard Residual	Posttest Core Reading Words Score	Predicted Value	Residual
9	-3.179	11	21.53	-10.534

Pallant (2007) recommends the analysis of Cook's Distance to evaluate whether this case (#9) had any undue influence on the results as a whole. According to Tabachnick and Fidell (2007), cases with values larger than 1 are potential problems. Cook's distance value of .256 is below this value, suggesting that case number 9 is not influencing the results.

Data Analysis. Standard multiple regression was used to assess the ability of progress monitoring and reading self-concept to predict literacy skills of struggling readers. Refer to Table 34 the model summary assessing the success in predicting. With a small group size, Pallant (2007) suggests reporting the adjusted R Square value rather than the R Square value. Meyers (2006) supports this by stating that researchers using a sample size of less than 60 should report the adjusted R Square value. The R Square value of .03 indicates that progress monitoring and reading self-concept explain 3% of the variance in core reading words, with no significance. The Adjusted R Square value of -0.37 reports that progress monitoring and reading self-concept explain 3.7% of the variance in core reading words, though with no significance. No further analysis is necessary since this tells us that the prediction of the core reading words by using progress monitoring or reading self-concept is no better than chance.

Table 34

Standard Multiple Regression Model Summary

Model	R	R Square	Adjusted R Square
1	.174[a]	.030	-.037

a. Predictors: (Constant), Progress Monitoring and Reading Self-Concept Score

CHAPTER FIVE: SUMMARY AND CONCLUSIONS

This Chapter contains six sections elaborating on this study. An overview of the first four chapters of this study is provided. Next, findings are organized by research question, analysis, and synthesis of the research findings for each question. This is followed by a limitations section, which expands on assertions made in Chapter Three, along with issues that surfaced throughout the study. The implications section provides suggestions for what should be done as a result of this study regarding the use of progress monitoring with struggling readers. Suggestions for future research offer ideas on what should be done to further investigate in this area. The final section is a summary of this research project.

Overview of the Study

The National Reading Panel (NRP) identified five components essential to a child's learning to read: (a) phonemic awareness, (b) phonics, (c) vocabulary, (d) fluency, and (e) comprehension. The *No Child Left Behind Act* (NCLB) of 2001 redefined education built on four basic principles: (a) stronger accountability for results in educational achievement, (b) increased flexibility and local control of funding, (c) expanded options for informed parental choice, and (d) an emphasis on teaching methods are supported through scientific research. NCLB also requires all children in grades 3-8 to be tested every year regarding reading achievement.

The Reading First initiative identified instruction and assessment as being intricately linked. To ensure that all students become successful readers, four types of assessment were identified that must be conducted at the classroom level. These four types of assessments are screening, diagnostics, progress monitoring, and outcome assessments (Hosp & Hosp, 2003; United States Department of Education, 2006).

Curriculum-Based Measurement (CBM) is a tool used by special education teachers to monitor progress in basic academic skills assisting them in decision-making. Olinghouse et al. (2006) stated that CBM is a scientifically validated progress monitoring instrument. Progress monitoring consists of data collection on regular intervals to keep track of children's academic development. Changes to instruction are based on the interpretation of progress monitoring data (Fuchs & Fuchs; n.d.; Olinghouse et al., 2006).

As part of IDEA 2004, along with progress monitoring, Response to Intervention (RtI) also became required in general education classrooms (McCook, 2006). RtI consists of three tiers of intervention incorporating quality classroom instruction, scientifically-based supplemental small group instruction, and specific intensive and explicit instruction. School districts and states must develop these programs including specific progress monitoring processes.

This researcher designed the current study to evaluate a specific early intervention reading program's use of an imbedded progress monitoring system. This study also examined the impact of progress monitoring and a student's reading self-concept on predicting post literacy skills (core reading words).

Participants for this study were first and second grade struggling readers. All students were given the *DRA2* (Beaver, 2006) with the lowest 20% of students then being administered the *Dominie* (DeFord, 2004). Students were then identified to participate in the reading support programs. The researcher sent a total of 102 consent forms to parents of these identified struggling readers. A total of 40 forms (39%) were received giving permission for children to participate in this study. A possible reason for the low return rate could be that the process for the distribution of the parent consent forms varied between

buildings. It was this researcher's intention to have a unified distribution in all seven participating schools. Administrators each had their own methods for disseminating information to their parents. Therefore, at each school, staff members distributed the study documents and permission forms in their own manner.

Once permission was received, students were administered the *Reading Self-Concept Scale* (Chapman & Tunmer, 1995b). All pretest data were collected for all participating: (a) *Dominie* subscale scores and (b) *Reading Self-Concept Scale* results. Students in the experimental group received *Leveled Literacy Intervention* (*LLI;* Fountas & Pinnell, 2009), a reading support program with an imbedded progress monitoring system. Students in the comparison group received a reading support program already being utilized by their school staff. Upon completion of the 14-week intervention time period, all students in the experimental and comparison groups were then administered the same subtests of the *Dominie.*

The specific research questions addressed were:

1. Is there a significant difference in literacy skills (core reading words, core writing words, phonemes, and spelling) between struggling readers who are part of a reading support program with an imbedded progress monitoring system and those who participate in a reading support program without an imbedded progress monitoring system?
2. To what degree and in what manner do progress monitoring and reading self-concept predict literacy skills (core reading words)?

The researcher used the SPSS Version 15.0 (2006) as the statistical analysis tool. A Bonferroni correction was utilized to adjust the alpha level of .05, which was divided by 2,

resulting in an alpha level of .025 (Huck, 2008; Meyers et al., 2006), which was employed for all statistical tests. For Research Question 1, the researcher utilized a quasi-experimental design employing a two-group multivariate analyses of variance (MANOVA) to determine the differences in mean scores in four subscales of the *Dominie.* The four subscales included, core reading words, core writing words, phonemes, and spelling.

For Research Question 2, the researcher applied a correlational design with multiple regression to determine the proportion of variance using the combination of predictor variables, progress monitoring and reading self-concept, in relation to core reading words posttest means.

Findings and Discussion

This section presents the results from the statistical analyses performed in Chapter Four. The initial section reviews and discusses the findings regarding Research Question 1. It is followed by a section that reviews and discusses the findings regarding Research Question 2.

Research Question 1

Is there a significant difference in literacy skills (core reading words, core writing words, phonemes, and spelling) between struggling readers who are part of a reading support program with an imbedded progress monitoring system and those who participate in a reading support program without an imbedded progress monitoring system?

A Wilks' Lambda statistic indicated no statistically significant difference in struggling readers' literacy skills (core reading words, core writing words, phonemes, and spelling), between struggling readers in a reading support program with an imbedded progress monitoring system and those who were in a reading support program without

progress monitoring. While this did not require further analysis, the dependent variables were analyzed separately in the Tests of Between-Subjects Effects; using a Bonferroni adjusted alpha level of .025, resulting in no statistically significant differences for any of the four dependent variables (core reading words, core writing words, phonemes, and spelling).

The mean scores for each of the dependent variables were reviewed for the experimental and comparison groups. This analysis revealed that for each of the independent variables, the comparison groups mean score was slightly higher than the experimental group means score. Refer to Table 35 for posttest mean scores for each of these groups.

Table 35

Posttest Mean Scores for Experimental and Comparison Groups

	Experimental	Comparison
Core Reading Words	22.47	22.57
Core Writing Words	38.06	40.21
Phonemes	53.59	54.35
Spelling	10.59	11.59

Vygotsky's (1978) zone of proximal development theory supports that each and every student develops at his or her own pace. Vygotsky also believed that learning and child development were directly related though not in equal measure or at the same time. This supports the concept that one program of instruction may not assist each student to master the material presented. The use of progress monitoring may assist educators in guiding students from their zone of proximal development to their potential development, but in this instance, it was no better than the traditional strategies employed in the current reading support program. A conclusion could be acknowledged that these students may not have been in

their zone of proximal development, therefore, not ready to move on to their potential development.

Results from this study support the finding of Foorman, et al. (1998), which indicated that not all instructional approaches have the same impact on student achievement. The growth individual students make is dependent on each student's needs. In addition, McIntyre et al. (2008) concluded that neither scripted nor non-scripted approaches to early reading instruction were more effective for phonics or reading achievement of first-grade struggling readers. In this researchers' study, the experimental reading support program utilized *Leveled Literacy Intervention* (*LLI*; Fountas & Pinnell, 2009), which is a scripted approach to reading incorporates an imbedded progress monitoring system. The comparison reading support program utilized non-scripted programs, which were developed in each building.

Results of this study may be compared with the findings of Olinghouse et al. (2006) who found that it is important to choose an appropriate progress-monitoring assessment based on the students' skill level, the targeted remediation, and the goals of the intervention. This researcher wonders if the imbedded progress-monitoring assessment was focused on the participating students' needs.

Roehrig et al. (2008) surveyed reading coaches who identified barriers when attempting to use assessment data to inform instruction. These barriers included (a) coach availability and quality of support received, (b) a breakdown between receiving assessment results and knowing what to do with children, (c) teacher knowledge, and (d) willingness of teachers to examine the effectiveness of their practices using student assessment. The current study did not utilize reading coaches, though the K-6 Language Arts Coordinator managed the regular progress monitoring meetings. The above barriers were out of the Coordinators'

control, though they may have impacted the results of the present study. Participating educators may not have been completely comfortable utilizing the *LLI* program or the progress monitoring approach. While there was ongoing training throughout the study, it may have been beneficial to have more detailed training prior to the beginning of the study.

Research Question 2

To what degree and in what manner do progress monitoring and reading self-concept predict literacy skills (core reading words)?

A standard multiple regression indicated no significance in progress monitoring and reading self-concept predicting literacy skills (core reading words). The R Square value of .03 indicated that progress monitoring and reading self-concept explain 3% of the variance in literacy skills (core reading words).

The present study resulted in a mean score of 3.64 for the total Reading Self-Concept Score. A score of 1 indicates low reading self-concept, while a 5 indicates a high reading self-concept. A mean of 3.64 for the sample of struggling readers demonstrates a slightly positive reading self-concept. This supports the conclusion of Chapman and Tunmer (1995a) that during the first 3 years of school, children's attitudes toward reading are very positive. Rider and Colmar (n.d.) support the finding that by eight years of age, children have firmly established reading self-concepts. The finding that following the 4th year, children indicated less optimism supports the need for providing reading support programs as early as possible in children's educational programs in order for them to continue to have positive reading self-concepts.

The current research results of a slightly positive reading self-concept during the pretest phase supports the suggestion made by Chapman and Tunmer (1997) that "reading

self-concept is more a consequence than a cause of reading performance" (p. 287). This statement suggests that students' positive reading self-concepts were not a cause of their poor reading performance.

Limitations of the Study

The most significant limitation of the study was the small sample size. Gall, Gall, and Borg (2003) along with Meyers et al. (2006) support the fact that statistical power increases automatically with sample size, creating increasingly more stable and precise estimates of population parameters. A second limitation is that the experimental group was introducing a new program, which they were trained in prior to and simultaneously with the study period. The comparison group was implementing a program with which, they were familiar with and had been utilizing for a number of years.

Threats to Internal Validity

Threats to internal validity are imposed by extraneous variables that have not been controlled by the researcher. Therefore, the observed effect can be attributed to the extraneous variables rather than solely to the experimental variables. This researcher attempted to control extraneous variables which could threaten the internal validity of this study. The 14-week time period of the study addressed issues of maturation and history, by reviewing pretest and posttest records though this was still of concern. To minimize effects of treatment implementation training for *LLI* all participants were trained prior to and throughout the study period. Meetings regularly addressed any questions regarding students as well as implementation and use of *LLI*.

In an attempt to control for demoralization of respondents, all participating educators in the experimental and comparison groups were initially trained in *LLI*. All educators in the

comparison group were aware that they would continue their training in the use of *LLI* during the following school year. Limitations arise from experimental treatment diffusion if the experiment is perceived as desirable compared to the comparison condition. In the current study, the comparison group educators continued utilizing the reading support program with which they were familiar. The experimental groups were trained in a new reading support program. The educators in the comparison groups were familiar with the program they were using while those in the experimental group were learning a new program. The educators in the treatment group may have lacked some skills needed to thoroughly implement the *LLI* program, impacting the growth made by their students as well as their posttest scores.

Threats to External Validity

External validity is the extent to which the findings of a study can be applied or generalized to other settings. Isaac and Michael (1997) state "By involving a wide variety of classes from several settings it is possible to achieve an even higher degree of external validity" (p. 77). They state this is especially true when considering interaction of the selection and treatment. Therefore, this researcher involved seven schools in total, three in the experimental group and four in the comparison group, to achieve a higher degree of external validity. The reactive effects of experimental procedures (Isaac & Michael, 1997) may produce effects that limit the generalizability of the findings. The experimental groups were observed at least monthly. Observations rotated between the three experimental schools to minimize this effect. Students though, may have felt nervous or uneasy with being the center of focus, therefore impacting their final posttest results.

Implications

With the use of a Response to Intervention (RtI) model incorporating progress monitoring, this study provided a variety of implications. *Leveled Literacy Intervention* (*LLI*; Fountas & Pinnell, 2009) is one of a variety of programs which focuses on progress monitoring of students. This program incorporates all five of the components essential to a child's learning to read, identified by the National Reading Panel (NRP). The use of progress monitoring with struggling readers is appropriate due to federal laws requiring yearly testing of students in grades 3 through 8 specifically related to reading achievement.

Implications include the need for continued professional development on the implementation of the LLI program or any new reading support program. Once educators are comfortable and confident using the program, it is expected that growth should be observed. Research reviewed in Chapter Two stated that children learn in different ways (Vygotsky, 1978) and that, not all instructional approaches have the same impact on student achievement (Foorman, et al.1998). An implication may be that more than one reading support program should be available to meet the needs of all children. In this way, struggling readers could be placed in the program that best meets their needs.

Suggestions for Future Research

It is imperative to conduct future research in this area. This research should focus on the use of progress monitoring in a longitudinal study, over time. Studies over time may also include reading self-concept which is measurable over a minimum of a six-month period. Research studies should be completed on a variety of reading support programs with imbedded progress monitoring systems with a large sample size, for the same length of time. These studies may identify which programs meet the various needs of diverse students.

Educators may need more specialized training in the use of progress monitoring in order for results to be significant. This may need to occur after educators have utilized the techniques and had training in the program for a year or more, prior to beginning a study.

The development of progress monitoring systems to meet the specific varied needs of struggling readers is another vital area for future research. Another component to research in the future is program based progress monitoring versus teacher developed progress monitoring to meet each students' needs. An examination of a program's effect may also be considered for future research, it may also include qualitative data. This may include the perceptions of the individuals delivering the treatment and program for the comparison groups.

Summary

The initial question of this research study was related to the impact of progress monitoring with struggling readers on their literacy skills. Findings yielded that there was no significant difference between students who participated in a program with progress monitoring imbedded into the program and those who did not receive a program incorporating progress monitoring. Means for both groups did increase from pretest to posttest, which shows that reading support is essential in improving literacy skills. Educators need to consider the most effective and efficient way to improve student reading achievement. Implications of these results are important for educators and administrators to consider when implementing an RtI model in schools.

The second question considered how effective progress monitoring and reading self-concept were in predicting students' literacy skills as measured by core reading words. Again, results did not demonstrate statistical significance. Educators must consider these

results when designing programs to support struggling readers in order to examine whether or not these programs are being delivered effectively.

References

Allington, R. L. (2001). *What really matters for struggling teaders: Designing research-based programs.* New York: Longman.

Armbruster, B. B., Lehr, R., & Osborn, J. (2003). *Put reading first the research building block of reading instruction: Kindergarten through grade 3.* (C. R. Adler, Ed.) Center for the Improvement of Early Reading Achievement.

Beaver, J. M. (2006). *Developmental reading assessment K-3* . Parsippany, NJ: Pearson Learning.

Blachman, B. A., Fletcher, J. M., Clonan, S. M., Schatschneider, F. D., Shaywitz, B. A., & Shaywitz, S. E. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow-up. *Journal of Educational Psychology , 96* (3), 444-461.

Boersma, F. J., & Chapman, J. W. (1992). *Perception of ability scale for students.* Los Angeles: Western Psychological Services.

Bookbinder, G. E. (2002). *Salford sentence reading test.* London: Hodder and Stoughton.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally College Publishing Company.

Capizzi, A. M., & Fuchs, L. S. (2005). Effects of curriculum-based measurement with and without diagnostic feedback on teacher planning. *Remedial and Special Education, 26* (3), 159-174.

Chapman, J. W., & Tunmer, W. E. (1995a). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology , 87* (1), 154-167.

Chapman, J. W., & Tunmer, W. E. (1995b). *Reading self-concept scale. Psychology in education portfolio*. Windsor, U.K.: The NFER-NELSON Publishing Company Ltd.

Chapman, J. W., & Tunmer, W. E. (1997). A longitudinal study of beginning reading achievement and reading self-concept. *British Journal of Educaiton Psychology , 67*, 279-291.

Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. (2000). Early reading-related skills and performance, reading self-concept, and the development of academic self-concept: A longitudinal study. *Journal of Educational Psychology , 92* (4), 703-708.

Children's Educational Services. (1987). *Test of oral reading fluency.* Minneapolis, MN.

Christensen, C. A., & Bowey, J. A. (2005). The efficacy of orthographic rime, grapheme-phoneme correspondence, and implicit phonics approaches to teaching decoding skills. *Scientific Studies of Reading , 9* (4), 327-349.

Connecticut Economic Resource Center. (2009). *CERC town profile: Fairfield, Connecticut*. Retrieved Novemnber 21, 2009, from http:// cerc.com/TownProfiles/Customer-Images/fairfield2009.pdf

Daly, E. J., & Murdoch, A. (2000). Direct observation in the assessment of academic skills problems. In E. S. Shapiro, & T. R. Kratochwill, *Behavioral Assessment in School* (pp. 46-77). New York: The Guilford Press.

DeFord, D. (2004). *Dominie reading and writing assessment portfolio*. Parsippany, NJ: Pearson Education.

Driscoll, M. P. (2005). *Psychology of learning for instruction.* New York: Pearson Education.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test-revised.* Circle Pine, MN: American Guidance Service.

Foorman, B. R., Fletcher, J. M., Francis, D. J., Mehta, P., & Schatschneider, C. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology , 90* (1), 37-55.

Fountas, I. C., & Pinnell, G. S. (2009). *Leveled literacy intervention program guide: Green system, Levels A-J, Lessons 1-110* . Portsmouth, NH: Heinemann.

Fuchs, L. S., & Fuchs, D. (n.d.). *What is scientifically-based research on progress monitoring?* Retrieved November 1, 2008, from National Center on Student Progress Monitoring: http://www.studentprogress.org

Fuchs, S., & Fuchs, D. (2000). Analogue assessment of academic skills: Curriculum-Based measurement and performance assessment. In E. S. Shapiro, & T. R. Kratochwill, *Behavioral Assessment in School* (pp. 168-201). New York: The Guilford Press.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Montioring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children , 71* (1), 7-21.

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

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational research an introduction* (7th ed.). New York, NY: Allyn and Bacon.

Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills.* Eugene, OR: Institute for Development of Education Achievement.

Hosp, M. K., & Hosp, J. L. (2003). Progress monitoring: An essential factor for student success. *The Utah Special Educator , 24* (4), pp. 26-27.

Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding,word reading, and comprehension: Do the relations change with grade? *School Psychology Review , 34* (1), 9-26.

Huck, S. W. (2008). *Reading statistics and research (5th ed.).* Boston, MA: Pearson Education.

Iaquinta, A. (2006). Guided reading: A research-based response to the challenges of early reading instruction. *Early Childhood Education Journal , 33* (6), 413-418.

Issac, S., & Michael, W. (1997). *Handbook in research and evaluation (3rd ed.).* San Diego, CA: Educational and Industrial Testing Services.

Jastak, & Wilkinson. (1984). *The wide range achievement test.* Willmington, DE: Jastak Associates.

Kaufman, A. S., & Kaufman, N. L. (1985). *Kaufman test of educational achievement*. Circle Pines, MN: American Guidance Service.

Kovaleski, J. F., Gickling, E. E., Morrow, H., & Swank, P. R. (1999). High versus low implementaiton of instructional support teams: A case for maintaining program fidelity. *Remedial and Special Education , 20* (3), 170-183.

Levy, B. A., & Lysynchuk, L. (1997). Beginning word recognition: Benefits of training by segmentation and whole word methods. *Scientific Studies of Reading , 1* (4), 359-387.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Newbury Park: SAGE Publications.

Lovett, M. W., Lacerenza, L., & Borden, S. L. (2000). Putting struggling readers on the PHAST track: A program to integrate phonological and strategy-based remedial reading instruction and maximize outcomes. *Journal of Learning Disabilities , 33*, 458-476.

McCook, J. D. (2006). *The RTI guide: Developing and implementing a model in your schools.* Horsham, PA: LRP Publications.

McIntyre, E., Rightmyer, E. C., & Petrosko, J. P. (2008). Scripted and non-scripted reading instructional models: Effects on the phonics and reading achievement of first-grade struggling readers. *Reading & Writing Quarterly , 24* (4), 377-407.

McMaster, K. L., Fuchs, D., Fuchs, L., & Compton, D. L. (2005). Responding to Nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children , 71* (4), 445-463.

Meyers, L.S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research design and interpretation.* Thousand Oaks, CA: Sage Publications.

National Institute of Child Health and Human Development. (2000a). *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.* (NIH Publication No. 00-4769). Washington, DC: Government Printing Office.

National Institute of Child Health and Human Development. (2000b). *Report of the national reading panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* (NIH Publication No. 00-4754). Washington, D C, U.S.: Government Printing Office.

No Child Left Behind. (2004, September 21). *Education Week* .

Olinghouse, N. G., Lambert, W., & Compton, D. L. (2006). Monitoring children with reading disabilities' response to phonics intervention: Are there differences between intervention aligned and general skill progress monitoring assessments? *Exceptional Children , 73* (1), 90-106.

Pallant, J. (2007). *SPSS survival manual a step by step guide to data analysis using SPSS for windows (3rd ed.).* New York, NY: Open University Press.

Rider, N., & Colmar, S. (n.d.). *Reading achievement and reading self-concept in year 3 children.* Retrieved March 14, 2009, from Australian Association for Research in Education: http:www.aare.edu.au/05pap/col05347.pdf

Roehrig, A. D., Duggar, S. W., Moats, L., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction: Identifying potential supports and challenges. *Remedial and Special Education, 29* (6), 364-382.

Safer, N., & Fleischman, S. (2005). How student progress monitoring improves instruction. *Educational Leadership, 62* (5), pp. 81-83.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441.

Simmons, D. C., Coyne, M. D., Kwok, O.-m., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilites , 41* (2), 158-173.

Statistical Package for the Social Sciences. (2006). SPSS 15.0 guide book. Chicago: SPSS Inc.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Strategic School Profile, 2008-2009 for the Fairfield Public School District. (n.d.). Hartford, CT: State Departmentof Education.

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics-second edition*. New York, NY: Harper and Row Publishers.

Torgesen, J. W., Wagner, R., & Rashotte, C. (1997). *Test of word reading efficiency.* Austin, TX:: Pro-Ed.

Tunmer, W. E., & Chapman, J. W. (2006). Metalinguistic abilities, phonological recoding skill, and the use of context in beginning reading development: A longitudinal study. In M. Joshi, & P. G. Aaron, *Handbook of Orthography and Literacy* (pp. 617-635). Mahwah, NJ: Lawrence Erlbaum Associates.

United States Department of Education. (2006). *Helping readers achieve and succeed.* Retrieved May 1, 2009, from http:www.ed.gov/nclb/methods/reading/strivingreaders.pdf

United States Department of Education. (2007). *Nation's report card shows record gains: Achievement gap continues to narrow as student population becomes more fiverse.* Retrieved May 1, 2009, from http://www.ed.gov/nclb/accountability/achieve/report-card2007.pdf

United States Department of Education. (2008a). *Progress by our dchools and the U.S. department of education.* Retrieved May 1, 2009, from http://www.ed.gov/print/nclb/accountability/results/trends/progress.html

United States Department of Education. (2008b). *Reading first: Student achievement, teacher empowerment, National Success.* Retrieved May 1, 2009, from http://www.ed.gov/nclb/methods/reading/readingfirst.pdf

Valencia, S. W., & Buly, M. R. (2004). Behind test scores: What struggling readers REALLY need? *The Reading Teacher, 57* (6), pp. 520-531.

Vaughn, S., Thompson-Linan, S., Kouzekanani, K., Bryant, D. P., Dickson, S., & Blozis, S. (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education , 24* (5), 301-315.

Vaughn, S., Wanzek, J., Murray, C. S., Scammacca, N., Linan-Thompson, S., & Woodruff, A. L. (2009). Response to early reading intervention: Examining higher and lower responders. *Exceptional Children , 75* (2), 165-183.

Vygotsky, L. S. (1978). *Mind in society.* Cambridge: Harvard University Press.

Wagner, R., Torgesen, J., & Rashotte, C. (1999). *Comprehensive test of phonological processing.* Austin, TX: PRO-ED.

Wechsler, D. (1974). *Wechsler intelligence scale for chilrdren-revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Wechsler intelligence scale for children--third edition (WISC-III).* San Antonio, TX: Psychological Corporation.

Wiederholt, J. L., & Bryan, B. R. (1992). *Gray oral reading tests--third edition (GORT--3).* Austin, TX: Pro-ed.

Woodcock, R. W. (1987). *Woodcock reading mastery tests--revised (WRMT-R).* Circle Pines, MN: American Guidance Service.

Woodcock, R. J., & Johnson, M. B. (1989). *Woodcock-Johnson psychoeducational battery-revised (WJ-R).* Chicago: Riverside.

Appendix A: Model Lessons Schedule

Date	**Lesson Focus**
February 25, 2009	**Components of Leveled Literacy Intervention** • **Re-reading for fluency** • **Running Record** • **Word work connections to Reading and Writing** • **Writing** • **New Book**
March 11, 2009	**New Book Orientation** • **How is meaning conveyed?** • **What are appropriate language structures to introduce to readers?** • **What vocabulary issues are appropriate to introduce?** • **How many words are introduced to students before reading a text?**
March 25, 2009	**Colleague Visit** • **What components did you witness during the lesson?** • **What teaching points did you observe?** • **How was teaching reinforced in writing?** • **What concerns did you see in the lesson?**
April 8, 2009	**Fluency Opportunities-Consider and discuss how fluency is developed in the following sections of the lesson:** • **Familiar rereading** • **Word work** • **Writing** • **New book**
April 22, 2009	**Writing** • **View video of writing lesson** • **Model layout of practice page and writing page** • **Taking words to fluency** • **Sound/letter boxes**
May 13, 2009	**Strategic Processing of Text** • **Elkonin Boxes** • **Sound boxes/Letter Boxes**
May 27, 2009	**Word Work** • **Making and Breaking Words** • **Procedures for learning** • **Generative words (onsets, rimes)**
June 9, 2009	**Data Monitoring** • **Data Collection** • **Next Steps**

Appendix B: Literacy Intervention Student Visit – Record

Literacy Intervention Accelerating Student Visit - Record

Host: _____ School: _____ Date: _____
Colleague/s: _____ ❒ Initial visit ❒ Follow-up
Student: _____Weeks: _____Lessons: _____ Entry text level: _____ Current text level: _____
***Before* the visit:** host *analyses the last 3-5 lesson records and running records* identifying trends in the student's current level of literacy processing. Please ensure all records are available for the visit.

***During* the lesson**: Colleague/s record as much information as possible regarding teaching & student literacy processing.

***After* the lesson**:

Acknowledgement of student learning i.e. what can the student currently control in literacy?	What does the student need to learn how to do next?	What contingent teaching is needed? Be specific / explicit re: prompts. Include references to theory .
Reading continuous text		
Writing continuous text		
Other eg. LID, WM&B, CUS		

Appendix C: Progress Monitoring Sheet

Text Level										
T(44)										
S										
R										
Q										
P										
O										
N										
M										
L										
K										
J										
I										
H										
G										
F										
E										
D										
C										
B										
A										
	Week 1-2	Week 3-4	Week 5-6	Week 7-8	Week 9-10	Week 11-12	Week 13-14	Week 15-16	Week 17-18	Week 19-20
Mark Student Progress: **O-90-94 % Accuracy (Instructional)** ***-95-100% Accuracy (Independent)**										

Appendix D: Letters of Consent

To: Western CT State University Instructional Leadership Doctoral Committee
From: (name of superintendent)
Date: November 21, 2008
Re: Doctoral Study Approval

Please allow this letter to serve as permission for Mrs. Teresa Samuelson to conduct research in (name of school and district) as partial fulfillment of the requirements for the degree of Doctor of Education in Instructional Leadership from Western Connecticut State University. Teresa will be conducting this research under the supervision of (name of supervisor) and in cooperation with (name of LA curriculum leader). The focus of this research is to measure the effect of data analysis and progress monitoring on student acquisition of early literacy skills.

_____
Superintendent of Schools
Name of District

To: Western CT State University Instructional Leadership Doctoral Committee
From: (name of principal)
Date: November, 2008
Re: Doctoral Study Approval

Please allow this letter to serve as permission for Mrs. Teresa Samuelson to conduct research in (name of school and district) as partial fulfillment of the requirements for the degree of Doctor of Education in Instructional Leadership from Western Connecticut State University. Teresa will be conducting this research under the supervision of (name of supervisor) and in cooperation with (name of LA curriculum leader). The focus of this research is to measure the effect of data analysis and progress monitoring on student acquisition of early literacy skills.

_____
Principal

January, 2009

Dear Colleagues,

We all want our children to be successful readers. In order to monitor the progress of struggling readers in our schools, I have developed a research project as part of my work with Western Connecticut State University's Instructional Leadership Doctoral Program. The project focuses on closely monitoring the progress of first and second grade struggling readers by gathering and reviewing data regularly, in order to improve their reading achievement. The progress monitoring of struggling readers research has the approval of the Superintendent of Schools and the Board of Education have approved given permission to conduct this research in the (Name of school district). It also has been approved by the Institutional Review Board at Western Connecticut State University, which is the University's research review committee.

The experimental group including Language Arts Consultants and/or Early Literacy Tutors in three of the elementary building will be trained this winter by (LA Coordinators name). The training will consist of assessing students through the use of three subscales of the *Dominie Reading and Writing Assessment Portfolio* to measure the reading skills of your students. Throughout the winter and spring they will then meet biweekly with (LA coordinator) to monitor progress of each student. (LA coordinator) will ask you to gather student work and data to discuss student progress and strategies for implementation. This will also be utilized in the research study.

The comparison group including Language Arts Consultants and/or Early Literacy Tutors in six to seven of the elementary buildings will also be trained this winter (LA coordinator). The training will consist of assessing students through the use of four subscales of the Dominie Reading and Writing Assessment Portfolio to measure the reading skills of your students. They will not meet in the bi-weekly progress monitoring of students.

Students will also complete the Reading Self-Concept Scale, which will be administered one-on-one. You will ask students 30 yes or no questions regarding how they feel about reading. The assessment takes about fifteen to twenty minutes to administer. This information will be collected throughout the program and will be reported based on group results, not individual student results. Please be assured that any information that you provide will be held in strict confidence by the researcher.

I appreciate all your time and input in this research project. Please sign and return the attached acknowledgement of your participation in this research study.

Sincerely,

Teresa Samuelson

Progress Monitoring of Struggling Readers

If you agree to have this information collected about your students, please complete the following information. Your signature indicates that the research study has been explained to you, that your questions have been answered, and that you agree to take part in this study.

_____ _____

Print Name School

_____ _____

Signature Date

Return to: Teresa Samuelson
(name of school and address)

January, 2009

Dear Parent or Guardian,

We all want our children to be successful readers. In order to monitor the progress of struggling readers in our schools, a research project was developed by Teresa Samuelson as part of her work with Western Connecticut State University's Instructional Leadership Doctoral Program. Her project focuses on closely monitoring the progress of first and second grade struggling readers by gathering and reviewing data regularly, in order to improve their reading achievement and self-concepts as readers. The progress monitoring of struggling readers research has the approval of the (name of school district). It also has been approved by the Institutional Review Board at Western Connecticut State University, which is the University's research review committee.

Language Arts Consultants and/or Early Literacy Tutors working with your child will be utilizing three subscales of the *Dominie Reading and Writing Assessment Portfolio* to measure the reading skills of your child. In the winter and spring some of the Language Arts Consultants and Early Literacy Tutors will be trained on interpreting these subscales using the progress monitoring model. They will meet biweekly with (name of LA coordinator), the district K-6 Language Arts Curriculum Leader to monitor the progress of each student. Each student's work will be reviewed and appropriate strategies and suggestions specific to the weekly monitoring will be discussed and provided.

Students will also complete the Reading Self-Concept Scale. This will be administered one-on-one by an educator. Your child will be asked 30 yes or no questions regarding how they feel about reading. The assessment takes about fifteen to twenty minutes to administer. This information will be collected throughout the program and will be reported based on group results, not individual student results. Please be assured that any information that you provide will be held in strict confidence by the researcher.

In order to improve my child's school program I understand that information will be collected about how my child is doing in school and in the Progress Monitoring Program. I understand that participation in the data collection is voluntary and that I may decide to withdraw my child from the data collection process at any time. If I withdraw my child from the data collection process, this will not change my child's school program in any way or affect my child's grades in school. I know that I can contact the Project Director at any time and that I can receive a final report of the research results in aggregate form upon request to the Project Director:

Teresa Samuelson
(name of school and address)

Sincerely,

Teresa Samuelson

Progress Monitoring of Struggling Readers

If you agree to have this information collected about your child, please complete the following information. Your signature indicates that the research study has been explained to you, that your questions have been answered, and that you agree to take part in this study.

_____	_____	_____
PRINT YOUR CHILDS NAME	SCHOOL	GRADE

_____	_____	_____
PRINT YOUR NAME (PARENT/GUARDIAN)	YOUR SIGNATURE	DATE

I acknowledge that the signer of this consent form has been informed of and understands the nature and purpose of this study and freely consents to participate.

Signature of Person Who Obtained Consent: _____

Please return this form to: Your Child's classroom teacher at his or her school

A copy of this form has been included for your records.

This research project has been reviewed and approved by the WCSU Institutional Review Board. If you have questions concerning the rights of the subjects involved in research studies please call the WCSU Assurances Administrator, at (203) 837-8281.