

EFFECTS OF WHOLE LANGUAGE PHILOSOPHY,  
IMPLEMENTATION, AND PRACTICE ON ACADEMIC  
ACHIEVEMENT IN READING

by

Kristin Pettey

A Project Submitted to the Faculty of

The Evergreen State College

In Partial Fulfillment of the Requirements for the Degree

Master in Teaching

2013

This Project for the Master in Teaching Degree

by

Kristin Pettey

has been approved for

The Evergreen State College

by

---

Lester Krupp, Member of the Faculty

June 2013

## ABSTRACT

This critical review of literature inspects whether whole language philosophy and instruction is beneficial to students' achievement in reading proficiency. Whole language is compared to other instructional methods, specifically phonics-based instruction, skills-based instruction, and the use of basal readers. This critical review will discuss the controversy behind whole language and the many instructional methods listed above. This review is separated into four sections: the effects of whole language and phonics based instruction; the effects of whole language and skills-based instruction; whole language and early education achievement; and whole language and English Language Learners. The results reviewed here generally support a balanced approach to literacy which includes whole language philosophy and practice but also implements implicit phonics instruction.

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my family. Thank you for helping me realize my dream and always being there to support me. Thank you dad for always pushing me to be the best I can be. Mom, thank you for listening and helping me through this day-to-day process of graduate school. Thank you to all my friends who maintained a fun balance in my life. David, thank you for staying by my side and always having faith and patience that I would succeed in anything I set out to do. I would like to thank Rachel and Jim for lending me an extra pair of eyes through this process. Also, thank you Lester Krupp for working with me and spending many hours working to help me through this. Without you all, this would have never happened.

## TABLE OF CONTENTS

TITLE PAGE	i
APPROVAL PAGE	ii
ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
CHAPTER 1: INTRODUCTION	1
Introduction	1
Rationale	2
Historical Context	4
Definitions	7
Limitations	9
Statement of Purpose	9
Summary	10
CHAPTER 2: CRITICAL REVIEW OF THE LITERATURE	12
Introduction	12
Whole Language and Phonics-Based Instruction	13
Whole Language and Skills-Based Instruction	30
Whole Language in Primary Education and Academic Success	70
Whole Language and English Language Learners and English as a Foreign Language Learners	91
Summary	98
CHAPTER 3: CONCLUSION	99
Introduction	99
Summary of Findings	101

Classroom Implications	106
Suggestions for Further Research	108
Conclusion	111
REFERENCES	113

## CHAPTER 1: INTRODUCTION

### **Introduction**

Whole language is recognized by many reading specialists and elementary teachers alike because of its humane and holistic philosophy of teaching. Whole language is not only praised for its comprehensive approaches in helping students learn to read, but it is also closely observed by today's educational community because of the controversy that surrounds it. The subject of whole language has been a contentious topic for many decades, allowing the education world to ask, "Are the instructional methods in whole language effective in teaching children to read?" This paper will explore the effects of whole language philosophy, implementation, and practice on student achievement in kindergarten through fifth-grade students.

Whole language directly relates to classroom practice because it is a holistic approach to teaching young students how to read and write. This holistic approach could allow students to approach reading from a perspective that is positive. Whole language also allows students to gain the skills of reading and apply them in a way that connects all reading cueing systems instead of isolating them while reading. Whole language is an essential tool students need to function in a society that relies highly on its citizens to be literate. Dewey (1986) suggested that students carry their experience, or prior knowledge, with them as long as they live. Whole language pedagogy can benefit any teacher who wants to have a more holistic and organic approach to reading and writing.

## **Rationale**

The effects of whole language instruction on student achievement in students in the elementary grades is an important question for teachers and people in the educational community because whole language is a philosophy that is rooted in ideas from acclaimed educational and linguistic theorists such as John Dewey, Jean Piaget, Lev Vygotsky, Louise Rosenblatt, and Kenneth and Yetta Goodman. With a practice that focuses so heavily on the experience of the learner, it is important to pay attention to how that practice is executed and to the overall products of these practices (Goodman, Y. M., 1989). If this holistic practice and pedagogy produces high student achievement in reading and writing, then this philosophy is worth incorporating into classroom practice.

The educational community reflects a variety of responses to this pedagogy. There has been an ongoing argument about whether whole language or phonics-based instruction is more effective for student achievement in reading. For example, Pressley, Wharton-McDonald, Allington, Block, and Morrow (1998) stated that “much of the current first grade debate has been between those who favor explicit instruction of beginning reading skills, especially the teaching of phonics, and those who favor an approach playing down systematic instruction in favor of immersion in literacy tasks, the whole language philosophy” (p. 2). Pressley et al. demonstrated that there are mixed reviews about how reading should be taught. Should reading be taught in a way that is holistic and allows the students to construct their own reading skills and ideas? Or should reading be taught in an explicit, instructional way that isolates skills such as phonics? The



controversy is not just limited to skills such as phonics, but stretches out into instructional reading packages such as basal readers and skills-based instruction.

Whole language is a philosophy that is widely accepted by educators who believe in the constructivist theory of learning and who encourage the teaching of skills within a meaningful context and within a whole text (Goodman, Y. M., 1989). In opposition to this approach, the National Reading Panel (NRP) conducted research on reading and writing instruction and presented to the American public of the United States its conclusions on what panelists think is important in reading (Weaver, 2009). According to whole language advocate Constance Weaver, “by deciding to investigate separately the effects of teaching five separate skills, the panel operationally defined reading as mastering a set of skills: phonemic awareness, phonics, fluency, vocabulary, and comprehension” (Weaver, 2009, p. xiv). Weaver explained that the panel did not undergo an investigation of their prescribed program in a real classroom, nor did they compare the effects to different programs. Nonetheless, the NRP’s program was adopted by the federal government and implemented in the Reading First initiative, which in turn was used in the No Child Left Behind Act (Weaver, 2009). NRP’s program isolates individual skills in the reading process, which is the exact opposite of what the whole language philosophy encourages (Weaver, 2009). A concern of whole language advocates is that the NRP’s methods and ideas of teaching reading back supporters of skills-based instruction. Both groups want to see improvement in reading in their students, but each philosophy supports different methods in teaching. This paper will review studies on whole language and phonics, basal readers, and skills-based instruction, and studies that

focus on the effects whole language has on early learners and English Language Learners.

### **Historical Background**

Yetta Goodman (1989) maintained that whole language is rooted as far back as the sixteenth century. She stated that educator John Amos Comenius, referring to a particularly popular children's picture book in Europe, wrote, "It is a little Book, as you see, of no great bulk, yet a brief of the whole world, and a whole language: full of Pictures, Nomenclatures, and Descriptions of things" (Goodman, Y. M., 1989, p. 113). According to Goodman, though Comenius didn't have the sense and meaning of what whole language is today, he was alluding to the idea that children must learn from their own experiences and from familiar material. This is the beginning of a long list of contributors to the whole language movement.

John Dewey was a major educational theorist from the early 20<sup>th</sup> century who put forth ideas that have been integrated into the whole language philosophy (Goodman, Y. M., 1989). Yetta Goodman suggested that Dewey's innovative view on teaching and learning contributed greatly to the whole language movement. His ideas on prior knowledge, experiential continuum, democratic practice, integrative curriculums, and humane educational practice especially influenced the whole language movement (Goodman, Y. M., 1989). Another theorist who contributed to the whole language movement was psychologist, Jean Piaget (Goodman, Y. M., 1989). Piaget developed a theory that explained the child's journey of learning through the world. Goodman (1989) summarized Piaget's constructivist view that "children do not wait for someone to transmit knowledge to them but, rather, learn through their own activity with external

objects and construct their own categories of thought while organizing their world” (p. 116). In other words, Piaget’s theory focused on the child’s construction of the world through active experience and experimentation. Another theorist who contributed to the whole language movement was Lev Vygotsky (1978). Vygotsky, now considered a social constructivist theorist, developed theories of the construction of knowledge through social interactions with other people. Vygotsky’s groundbreaking theory, the zone of proximal development (ZPD), contributed greatly to whole language philosophy. The ZPD is when students are able to reach their level of development with the help of someone who has surpassed that level of development. A more capable other can be a teacher or a student at a higher development level (Vygotsky, 1978). M. A. K. Halliday is a systemic linguist whose ideas contributed greatly to the whole language movement. He suggested that while children are using language they are also learning language (Goodman, Y. M., 1989). Halliday “has developed a system of functional grammar that relates the study of language to the actions within the situational context and to the relationship of the actors involved” (Goodman, Y. M., 1989, p. 117). His work has influenced whole language because it supports student language learning that connects to reading and making meaning out of that reading (Goodman, Y. M., 1989)..

Ideas from theorists like Piaget, Vygotsky, and Dewey contributed to the development of whole language. Educational theorists observed these ideas about learning and contributed to the philosophy of whole language. Kenneth Goodman and Frank Smith contributed to the development of this modern movement in the early 1960s (Goodman, Y. M., 1989). This movement was propelled forward primarily by educators, parents, and administrators who were invested in the method (Goodman, Y. M., 1989).

These progressive thinkers were said to have “developed the theory and research that established the notion of a unified single reading process as an interaction between the reader, the text, and language” (Goodman, Y. M., 1989, p. 117). This philosophy built on the ideas of Dewey, Vygotsky, and Piaget in proposing that students have an interaction with what they are reading; the meaning that students bring to that print involves a transactional experience (Goodman, Y. M., 1989).

Louise Rosenblatt (1988) more fully articulated this transaction between the text and the reader. Rosenblatt’s transactional model aligns with Dewey’s ideas of experiential continuum in that everything is linked to an experience which is carried on within a person for the rest of the person’s learning life.

Verbal communication and reading are essential skills to obtain to become functioning members of society. Whole language focuses on prior knowledge and experiences of the learner as these create a vocabulary for communication through oral and written language (Goodman, Y., 1989). The way that students learn to be literate (read and write) is determined by their experiences in classrooms when they are initially learning how to read (Miller, 2009). It is a teacher’s duty to teach students how to do this by connecting with students’ lives and making reading and writing a pleasant experience for them.

Piaget’s work suggested that students learn by constructing their own knowledge, or schema (Miller, 2009). Vygotsky (1978) suggested that learning can occur when a teacher or more capable peer scaffolds a learning opportunity for a learner who would not be able to understand alone. The process of knowledge construction is still happening,

but through scaffolding from a teacher or a more capable peer (Vygotsky, 1978). Dewey stated that learning happens when there is a connection between what is learned and the student's past experiences (Miller, 2009). These aspects of learning lead to the construction of a philosophy and practice (whole language) that specifically address the student's learning needs.

### **Definitions**

A better understanding of whole language requires certain definitions. Whole language is not an instructional guide one can acquire as a package. Instead, it is a philosophy. As a philosophy, whole language is rooted in progressive ideas that support creativity and construction of knowledge that connects a learner to prior experiences. The phrase *whole language* is used in many professional writings as a broad spectrum of ideas and practices. Whole language has proponents that would define it as a holistic approach to reading and writing in the early years of human life. In fact, whole language is a philosophy that extends upon the idea of natural language. Gilles et al. (1988) stated that "whole language may mean many things to many people, but its essential meaning is that all of language in an integrated form must be presented to students if they are to learn to read and write. Within the complexly organized system of language there are subsystems that work in concert to help humans organize their experiences and mediate meaning" (Gilles et al., 1988, p. 5). Gilles et al. stated that these systems must be kept whole. Furthermore, whole language is a philosophy that teachers use and implement in their own instruction and which looks different from classroom to classroom. Raines (1995) exclaimed that, "teachers who have whole language perspective operate their classrooms with an abundance of children's literature, use a writing process approach,

usually organize the curriculum in integrated, thematic units, teach strategies approaches [*sic*] to inquiry, and find authentic, meaningful ways for children to communicate about their lives and what they are learning” (p. 2).

As described by Raines (1995) above, reading is viewed as a means of constructing meaning from what one reads and what one knows. What the student already knows or has experienced is a strong factor in what the student might comprehend during the reading process. Reading is considered a process that involves “strategies of sampling, predicting, confirming, and integrating meaning” (Gilles, 1988, p. 9). Whole language also focuses on reading as a holistic process. For example, instead of isolating the cueing systems (semantics, grammar/syntax, graphophonemics, and pragmatics) from reading, a reader can pull from all the cueing systems to extract meaning (Anderson, 1984; Gilles, 1988). The semantic system has to do with meaning, the grammar/syntax system has to do with the order of the words to create meaning to the reader, graphophonemics has to do with the sound systems and visual systems of reading, and pragmatics have to do with the type of reading the reader is engaging in (Gilles, 1988).

This paper also focuses on instructional methods that focus on phonics-based and skills-based instruction. Instruction administered in this manner focuses on a particular skill in isolation. For example, phonics is frequently used by students to sound out words. This type of action is supported within a classroom in an isolated fashion without the use of authentic books or texts. Many studies presented in this paper contain treatment groups that received phonics-based instruction (e.g. Griffith et al., 1992). Skills-based instruction is an instructional method that consists of instruction that is teacher-based. Students are taught a lesson and then expected to engage in a worksheet or assignment

that models this lesson (e.g. Davis, 2010). Basal-reading programs are often used within skills-based classrooms where students engage in specifically made workbooks for texts that may have been altered from their original format (e.g. Dahl & Freppon, 1994).

### **Limitations**

This paper is limited to the study of the effects whole language teaching has on student achievement in reading and writing. Also, it is limited to studies that focus on teaching in grades kindergarten through five, because these are the grades where students primarily have instruction in how to read. This analysis of whole language will also focus on comparing the effects of whole language pedagogy with the effects of different instructional methods such as phonics-based and skills-based pedagogies. The use of standardized tests to help acquire quantitative results of a method or practice is a limitation because empirical testing does not always show results that are trustworthy.

### **Statement of Purpose**

The purpose of this paper is to investigate what strategies and philosophies of teaching reading are the most effective for student achievement in reading. This paper specifically focuses on the practice and philosophy of whole language. Gilles et al. (1988) stated that, “it’s essential meaning is that all of language in an integrated form must be presented to students if they are to learn to read and write. Within the complexly organized system of language there are subsystems that work in concert to help humans organize their experiences and mediate meaning” (Gilles et al., 1988, p. 5). Generally, whole language is a way that teachers educate students in all of the many pieces that make up language and connect those pieces to meaningful experiences. This paper

investigates the avenues that teachers take to educate their students in how to read and what the best possible instructional strategies are.

This paper will answer questions about what strategy is best in teaching students how to read. The question this paper strives to answer is “Does whole language philosophy and instruction assist primary students to achieve reading readiness compared to other reading programs like basal readers and skills-based instruction?” This question addresses three aspects:

1. Is whole language an effective philosophy and strategy to use to teach students how to read?
2. Is whole language more effective than phonics-based instruction, skills-based instruction, and basal-reader instruction?
3. Is whole language effective with English language learners?

These questions are important because they allow teachers to explore the importance of reading and the strategies that will help students achieve reading readiness. Reading is essentially a human need because language and text are used to communicate many messages. Therefore, this paper’s purpose is to examine what strategy and philosophy will help students achieve reading proficiency.

### **Summary**

Whole language is a philosophy, practice, and pedagogy that has its roots in progressive ideas and theories developed by John Dewey, Jean Piaget, Lev Vygotsky, and many others. The importance of this practice pertains to teachers, students, parents,



educators, and anyone who is interested in education and learning. Since whole language falls under the category of holistic, humane, and even progressive education, teachers should seriously consider whole language instruction in their teaching practices. This critical analysis of whole language will answer the question of whether or not whole language promotes student achievement in grades kindergarten through five. This analysis will review studies that focused on whole language practice and its effectiveness in comparison to other instructional methods such as skills-based instruction, mastery-based instruction, and phonics-based instruction. This topic is a major controversy in the educational world and deserves a critical analysis.

## CHAPTER 2: CRITICAL REVIEW OF THE LITERATURE

### **Introduction**

Chapter one introduced the purpose of this paper, which is to determine whether or not whole language instruction and philosophy are effective in helping students become proficient readers. Chapter one also discussed the importance of the use of whole language in the classroom because of its humane philosophy and holistic style of instruction. Whole language was defined and a brief history explained how the grass roots philosophy was started and implemented into educational practice. This philosophy can be adopted by teachers to use in their own classroom instruction to advance prior knowledge, experiential learning, and construction of the student's own learning. Chapter one also defined terms that are associated with whole language and with reading instructional practices and assessments. Exposure to whole language philosophy is said to have the potential to allow students to excel in reading and enjoy the activities that go with this instructional style, such as choice, collaboration, and human interaction. Instead of isolating skills needed to read, whole language teachers allow students to engage in real reading with authentic books to stimulate a realistic reading experience (Goodman & Goodman, 1981).

Chapter two critically reviews research that closely studied the effects of whole language practices. Chapter two is organized into four sections: (1) whole language and phonics-based instruction; (2) whole language and skills-based instruction; (3) whole language and primary academic achievement; and (4) whole language instruction and English Language Learners. These articles were chosen for review to determine whether

or not whole language instruction is effective in reading achievement in various instructional settings. Some studies will show that phonics-based or skills-based instruction is more effective than whole language instruction. Some studies will show that whole language practice is effective as opposed to phonics-based or skills-based instruction. Some studies will show that neither instructional method is more effective for student achievement in reading; instead, a balanced approach to reading (whole language and implicit phonics/skills-based instruction) is regarded as an effective approach to students' achievement in reading. As evidence by an overwhelming amount of empirical articles, a whole language approach combined with implicit phonics/skills-based instruction is an effective instructional method to help students achieve reading readiness.

### **Whole Language and Phonics-Based Instruction**

The seven studies reviewed in this section focus on the effects of whole language instruction versus phonics-based instruction. The first and second studies, by Dahl and Scharer (2000) and Dakin (1999), examined whether or not the implicit implementation of phonics in a whole language class was effective for achievement in reading. Donat (2006) and Ryder, Tunmer, and Greaney (2008) examined how the implementation of phonics-based program into whole language classrooms could help students learn how to read proficiently. The fifth study (Griffith, Klesius, and Kromrey, 1992) was a comparative study which asked the question whether whole language instruction or phonics based instruction was more effective to help students learn how to read. Finally, McIntyre (1995) and Baumann, Hoffman, Moon, and Duffy-Hester (1998) conducted surveys on individual teachers' instructional methods and philosophies. In addition to attaining information about instructional methods used within classrooms, researchers

also gathered the teacher's opinions about which instructional method was more effective for student achievement in reading. There have been quarrels between which philosophy or method is more effective on student achievement. Whole language experts argue that isolating phonics instruction can interrupt the reading process (Weaver, 2009). With whole language instruction, educators can teach phonics within the reading process without isolating instruction, thus keeping the reading process intact. The following studies demonstrate the value of a whole language approach to instruction that is not isolated and fragmented, but is holistic and teaches students phonics within a whole, authentic text.

Dahl and Scharer (2000) researched the effects of implicit phonics instruction administered within a whole language classroom. They conducted a quantitative study on 216 diverse elementary school students using a semi-experimental method. The study suggested that phonics could be implicitly taught within a whole language classroom, as long as phonics is not isolated and detached from the reading process.

Dahl and Scharer (2000) reviewed the origins and definitions associated with whole language and phonics-based instruction. With this knowledge, Dahl and Scharer (2000) were able to choose eight teachers from eight different schools keeping those teaching practices and philosophies in mind. The schools selected were suburban public, suburban Catholic, suburban whole language magnet, urban literature magnet, suburban language arts magnet (public), inner city public, and urban whole language school. The demographics of the students in the study when the pretest was administered were 179 European American, 34 African American, 1 Latino, and 2 Asian American. There were 8 classes from 8 different elementary schools that had 9 different teachers who all

conducted whole language instruction in their classrooms. Dahl and Scharer (2000) hired two consultants, a phonics expert and a whole language expert, to uphold credibility and dependability. These two experts followed the study and “oversaw site selection, data-collection and data-analysis procedures” (Dahl and Sharer, 2000, p. 585).

The study started off with a pretest that tested for phonics ability in decoding and encoding. The pretest determined whether the students would be in one of three groups, Group 1, 2, or 3. Group 1 consisted of students who scored a 27-37 on the pretest. Group 2 consisted of students who received a score 15-26, and Group 3 scored 2-14 on the pretest. Before the groups were determined, the pretest was sent to an outside resource to determine the composite scores of each student’s total reading score. After this pretest was administered, Dahl and Scharer (2000) observed reading and writing processes in the classrooms that were chosen for the study. Dahl and Scharer described reading instruction happening in whole groups, where the teacher would read to the whole group a picture book, then would engage in discussion about connections between the text and self and then connect a writing activity to the reading. This was where the phonological instruction was administered. Dahl and Scharer asserted that “a substantial portion (45%) of the phonics events coded in this study occurred in the writing program. Lengthy writing periods provided opportunities for children to grapple with phonics concepts” (p. 588).

At the end of the year, after administering some phonological instruction within multiple whole language classrooms, Dahl and Scharer (2000) administered a posttest. This posttest, which was identical to the pretest, was then compared to the pretest; the averages from all students in each group were compiled together to make a mean score

for each group. Group 1, the group that had the highest scores on the pretest, started at Level 14 (first grade) and rose to Level 28 (fifth grade). Group 2 raised their reading levels from Level 3 (pre-primer) to Level 14 (first grade). Group 3 raised their reading level from pre-primer to Level 4.

Dahl and Scharer (2000) showed various strengths and weaknesses in their study. A main strength of this study was its reliability. The researchers displayed a meticulously detailed method, means of measurement, and data-gathering techniques. The transferability of the study is high because of the detailed demographics of the sites and students. The sample population was large (n= 216), but was confined to the same geographical area. Reliability of the study was high due to the random selection of teachers from the same area and the willingness of the teachers to contribute what they could to the study. The number of students did not drop or increase either. This study shows a weakness in internal validity because of the use of nine different teachers. These nine different teachers could have had slightly different whole language approaches to teaching. The external validity was sound because of the exact same treatment that was administered to each classroom (implicit phonics instruction).

This study ultimately uncovers the misconception that phonics is not taught at all in whole language classrooms. In fact, Dahl and Scharer (2000) inform readers to oppose this dichotomy of whole language and phonics. In Dahl and Scharer's study, phonics was never taught in isolation, but was taught within literature instruction and writing instruction, where students could engage in the whole reading process (sample, predict, correction/confirm, and integration) and cueing systems (syntax, semantics, grapho-phonics, and pragmatics).

Dakin's (1999) study is very similar to Dahl and Sharer's (2010) because of the focus on phonics-based instruction and testing the participants on phonics skills. Dakin (1999) conducted a qualitative study on 12 inner-city Catholic elementary school students, using a practice-based method. Dakin's study researched whether or not explicit phonics instruction within a literacy curriculum was beneficial for students' reading proficiency. The author found that students must be taught strategies to decode words, therefore explicit phonics-based instruction within a literacy program is beneficial for beginner readers.

This study focused on elementary school students who attended an inner-city Catholic school. The study was conducted in San Francisco. The student population included students from the East Bay (20%) and from South San Francisco (13%). Forty-two percent of the students were Catholic, and 57% were non-Catholic; 40% were on free or reduced lunches. The demographics of the school were 59% African American, 16% Filipino American, 7% Hispanic, 5% Asian American, 4% European American, and 9% mixed races. Overall, twelve students were chosen to be focus students for this study, with four specific students chosen from each grade level in grades K-2. In each group of four students, one student exhibited a low reading level, two exhibited a medium reading level, and one exhibited a high reading level.

The research method included data gathering, evaluations, and coded results to ensure an interviewee's anonymity. Each teacher, from kindergarten through second grade, was interviewed to uncover what the educator's teaching philosophy was. Also, these interviews were conducted to obtain the teacher's daily reading routines. After this was established in the K-2 classes, an assessment was administered to the focus students

in each classroom. These assessments measured “level of phoneme awareness, effectiveness of phonics instruction as observed in word attack skills and word recognition/comprehension as it related to phonetic decoding skills” (Dakin, 1999, p. 33). These assessments started with the students reciting the alphabet, which tested for phonemic awareness. Then the students were given a word list to recite. After this, students read a passage where the researcher would keep track of “knowledge of the alphabet, word by word reading technique, self-corrections, reading skill procedures i. e. ‘finger pointing, head movement,’ sight vocabulary proficiency, phonetic decoding skills and difficulties, and rate of reading as it effected story accuracy” (Dakin, 1999, p. 35).

Dakin’s (1999) findings suggested that decoding methods are important for beginning readers. As evidenced by the interviews, these teachers agreed that an explicit phonics program within a literacy program is crucial for students to obtain reading proficiency. The assessments demonstrated a strong occurrence of self-correction while reading. The assessment demonstrated that students with incomplete phonemic awareness found that sounding out words was challenging. Dakin (1999) also found that students without this awareness could not comprehend what they read.

This study had weaknesses in credibility, transferability, and dependability. The methodology does not produce findings that are convincing because of the lack of detail in the data. For example, the data provided for the findings of the study was not described with hard numbers. Instead, there is a generalized description of how some of the students performed on the assessment. There is strength in the transferability of this study because of the detailed demographics that were described in the study. The dependability



of this study is weak because of the researcher being part of the teacher team that was interviewed and researched. There were no measures to prevent bias within the study.

Dakin (1999) wanted to prove that explicit phonics instruction helped students become proficient readers. Her findings indicate that explicit phonics instruction does help students become proficient readers, but her study's reliability and dependability are in question.

Donat (2006) conducted a similar study to Dakin's (1999) where she investigated the effects of the implementation of the reading program *Read Their Way* (RTW) in a district that primarily used whole language instruction. Donat conducted a quantitative study on students from twelve elementary schools in a rural public school district in Shenandoah Valley in Virginia. Donat found that the program RTW was effective in producing satisfactory results in student achievement in reading and writing.

Donat (2006) described RTW as a program that "is comprised of four essential components: phonemic awareness, phonics, contextual reading, and writing. It involves direct explicit instruction of phonemic awareness and phonic skills...RTW combines the phonics and whole language approaches to offer a balance that was expected to accelerate the overall achievement of students in language arts" (p. 307). Upon implementing the program into the Augusta County School District, the district officials set goals for student achievement. Kindergarten students were to reach the first primer level reader, first graders were to reach 1.2 level reader, second graders were to reach 2.2 level reader, and third graders were to reach the 3.2 level reader.

The Phonological Awareness Literacy Screening (PALS) test was administered to the participants (K-3) in fall and spring of the school year. Donat (2006) did not produce a detailed demographics list but alleged that the subgroups were demographically consistent with Virginia statewide demographics. The groups that were analyzed were two separate groups. The separate groups were formed by the first six classes to implement the RTW program and the last six classes to implement the RTW program. These two groups' PALS test scores were compared by dates from 1997-2002. The researchers measured whether or not students needed additional support in literacy while RTW was being implemented.

Donat (2006) found that 63% of first grade students at the fully implemented RTW sites had a second- or third-grade level of proficiency. However, the non-RTW implemented sites had a 38% proficiency in first graders. Donat (2006) expressed these gains in charts that reveal the percentage of students who do need supplementary services. As the percent of implementation of RTW went up, the percent of supplemental services went down. Overall, the achievement on reading was significantly more positive in the RTW groups that had fully implemented the program than in the groups that had not implemented them yet.

Donat (2006) took measures to ensure that the study was reliable, objective, and externally valid. The results of this study were seemingly reliable because of the lengths taken to clearly state all aspects of the study. The study lacked any mention of the demographics of the subjects. Donat (2006) did mention that the reliability of the study was stable because the study's demographics were consistent with the overall demographics within Virginia. What does come into question is the study's definition of

what whole language is and what a balanced curriculum is. RTW seems to be backed by the NRP because of its focus on isolated skills, yet RTW actually is acclaimed because of its use of whole language and explicit phonics instruction. Donat (2006) failed to explain how explicit phonemic and phonics instruction was executed, and it is possible that these skills were taught within an authentic text or oral story. Therefore, the reliability of the description of this program is weak because of the lack of information. It is important for students to learn phonemic awareness and phonics, but the whole language philosophy explains the importance of an uninterrupted reading process and use of the cueing systems. From this study it is impossible to conclude whether RTW was able to engage in the whole language philosophy.

Ryder, Tunmer, and Greaney (2008) also studied the effects of explicit phonemic awareness instruction on a group of students who were in a whole language influenced curriculum. The main purpose of the study was to prove that explicit phonics and phonemic awareness instruction could help students achieve reading readiness. Ryder et al. (2008) found that students who received explicit phonemic instruction scored better on post assessments measuring “phonemic awareness, pseudo-word decoding, context free word recognition, and reading comprehension” (p.349).

The participants in this study included 24 six- and seven-year-old students within a school in New Zealand. These students were randomly assigned to a control group or an intervention group. The control group of 12 students continued class instruction as usual, which was primarily whole language instruction, while the intervention group of 12 students remained primarily whole language based but had an addition of explicit phonics

instruction each day. The school population contained 64% European, 34% Maori, 1% Pacific Islander, and 1% Chinese students.

The method conducted in this study included a pre-assessment to narrow down 24 students from a pool of 64 students. The test administered was the Burt Word Reading Test, New Zealand Revision, which measured the students' experience with phonemes and phonics (Ryder et al., 2008, p. 354). At the end of the study, the students were given a post-assessment that was similar to the pre-assessment. The assessments measured achievement in five areas: "phonemic awareness, phonological decoding ability, context free word recognition ability, accuracy of recognizing words in connected text, and reading comprehension" (p. 356).

The findings indicated that the students who were in the intervention group did better on their post-assessments than the students in the control group. The measures were reported in a table, but the table was never placed into the article. However the results from the two-year follow-up were reported in a table and explained. The means in the intervention group from the pre-assessment and post-assessment were reported, but total mean scores from each group, the intervention and control group, were not reported at all. This lack of information is a major flaw in the validity of the study. The failure to report all information detailing the results of the study's underlying tests is a major flaw, which significantly weakens reliability. The results that were reported were the scores of the two-year follow-up which had students engage in a Burt test. The results demonstrated that the intervention group scored a mean of 52.7 compared to the controlled group's score of 43.7.

Griffith, Klesius, and Kromrey (1992) conducted a similar study where they investigated the effectiveness of whole language instruction versus traditional instruction on students' achievement in phonemic awareness. Griffith et al. conducted a quantitative study on first-grade students from two classrooms in a rural Florida school district, using a semi-experimental method. Griffith et al. found that the students who tested high in phonemic awareness did better than students who tested low on phonemic awareness.

Griffith et al. (1992) conducted their study on two separate groups of students from two different classrooms. These classrooms were located within a Florida district. One of the classrooms used whole language philosophy of instruction, and the other classroom used traditional instructional methods. In the whole language classroom phonics was not explicitly taught, but instead was taught indirectly through writing and spelling activities. The traditional classroom engaged in explicit phonics instruction regularly.

The GKR Phonemic Awareness Test was administered as a pretest to all students in each classroom in September. The GKR Phonemic Awareness test was a test that contained "six subtests: phonemic segmentation, blending, deletion of first phoneme, deletion of last phoneme, substitution of first phoneme, and substitution of last phoneme (Griffith et al., 1992, p.87). After the results were obtained, students were separated into two group in each class, one representing six of the lowest phonemically aware students, and the other representing six of the highest phonemically aware students. Then the students were given a posttest which measured progress in spelling, writing, decoding, comprehension, and writing fluency.

Griffith et al. (1992) found that there was no significant difference between instructional methods; however, there was a statistical difference between high phonemic awareness students and low phonemic awareness students. Griffith et al. explained that the high phonemic awareness students performed better on post-assessments than the low phonemic awareness students. This evidence suggested that the methods used to teach these students were effective for the students who were already doing well on phonics and phonemic awareness, but did not work well for students who did not do well on the pretest for phonics and phonemic awareness.

This study showed various weaknesses and strengths. This study's external validity was in question because its transferability was limited because of the lack of demographics or socioeconomic information needed to replicate this study in another setting. Griffith et al. (1992) offered vague descriptions of the subjects, where they came from, and what type of instruction was administered to each treatment group, therefore rendering the study unreliable. The study showed strength in internal validity because of the meticulous structure of the method. However, the internal validity is questionable when the study used the lowest scoring students and the highest scoring students. This demonstrates a regression in statistical measures (a posttest score that is close to the pretest score). If there was a statistical significance between treatment groups, then the results would not show a valid and significant finding; however, the researchers did not discover significant findings.

McIntyre's (1995) study challenged the results of Griffith et al.'s (1992) study. McIntyre (1995) conducted a qualitative study that followed three teachers' journeys through teaching primary students in a whole language, multi-age classroom. Three

teachers, Donna, Joy, and Tina, were chosen to co-teach 40 first, second, and third graders. The main purpose of this study was to uncover the instructional methods, whole language, skills-based, or balanced approach, that are appropriate to use when teaching developmentally appropriate content. McIntyre's (1995) findings suggested that the participants who engaged in this yearlong study all had beliefs grounded in whole language philosophy; however, their whole language instructional methods were not meeting the developmental needs of students. Therefore, these participants incorporated implicit phonics instruction within their whole language instruction and found that a balanced approach to teaching literacy is effective in helping students learn how to read.

The participants in this study consisted of three teachers. Donna was a highly acclaimed whole language teacher who was sought after in her district for workshops on whole language. The second teacher chosen for the study, Tina, was a student teacher from the local university whose philosophy matched Donna's. The third teacher chosen for the study was Joy, who was a special education teacher who pulled students out of class to administer one-on-one or small-group instruction. These three teachers received a classroom of 40 students who were in first, second, and third grades. The number of students who were from each grade was not provided, and neither was a demographics list.

The method administered in the study consisted of observations done in the classroom, informal interviews, and the Theoretical Orientations to Reading Profile (TORP). Observations were conducted to keep track of the instructional methods that were being implemented within the classroom throughout the year. Informal interviews were conducted during meetings amongst the three participants to make observations on

how their instructional methods were planned. The TORP surveys were administered before the school year started, during the school year, and after the school year had ended. The TORP survey consisted of 28 questions which produced a statement in which the teacher would answer the question with either *strongly agree*, *agree*, *disagree*, or *strongly disagree*. All teachers scored in favor of whole language instruction and philosophy on the survey administered before the school year.

At the beginning of the year, the classroom was described as a place where students were able to have a lot of freedom of choice in what they learned. The structure of the classroom and instruction was loose and was up for interpretation of the student. For example, students were told to read during reading time but were allowed to choose their own book. Sometimes this freedom would be too much for students to handle, and students ended up off task. The students who remained off task fell behind in reading and writing. As evidenced by classroom observations, students who were off task needed more structure within the classroom's activities, which can be viewed as a developmentally appropriate practice to engage in while teaching.

The results indicated that there were some changes in all three teachers' instructional methods as the school year went on. For example, McIntyre reported that two of the three teachers changed their answers to the statement, "Formal instruction in reading is necessary to insure the adequate development of all skills used in reading" (p. 151). Tina changed her answer from *disagree* to *agree*, and Donna changed her answer from *strongly disagree* to *agree*. Both teachers seemed to agree at the end of the year that some students need skills taught to them implicitly so that the meaning behind the lesson in literacy is not lost.



This study shows various weaknesses and strengths. The reliability of this study was strong due to the researcher's full disclosure of the method, data-gathering tools, and the primary TORP survey used for data measurement. Although the reliability of the study was challenged because the teachers already had relationships with the researchers, the full disclosure of method keeps the reliability strong. Confirmability is demonstrated through results that are consistent with previous credible peer-reviewed studies. The transferability of the study is not possible because of the lack of demographics of teachers, students, and the site of the study. The dependability of the study is reliable because of the researcher who was separated from the study.

Baumann, Hoffman, Moon, and Duffy-Hester (1998) studied the effects of whole language instruction and phonics-based instruction on student achievement. Baumann et al. (1998) conducted a quantitative study on 3,199 pre-kindergarten through fifth-grade public school teachers using an experimental method. The researchers wanted to uncover and expand on the phonics versus whole language discussion. They created a survey that investigated the beliefs and practices of teachers within the scope of phonics. They found that teachers value the implementation of literature and literacy-immersion practices along with the implementation of phonics with these activities. The researchers also found that 71% of teachers prefer to immerse their students in literature; however they also include implicit instruction in phonics to create a balanced approach to reading.

The participants of the study included elementary school teachers (89% European American and 93% female) and their students from all over the nation. The average number of students a participant had in her classroom was 25 students. The average demographics of a classroom included 68% European American student, 14% African

American students, 12% Latino/a students, and 6% students from other racial or ethnic groups (Baumann et al., 1998, p. 642). An average of 46% of students lived in middle class homes, 45% lived in lower-income homes, and 8% lived in upper class homes. Fifty-five percent of the students were average-level readers, 24% were below reading level, and 20% were above-level readers.

Baumann et al. (1998) conducted a survey with the help of the Survey Research Center (SRC) at the University of Georgia to understand what the popular teaching styles and practices were in the late 1990s. In a trial run, 3,199 public school teachers across the United States were randomly chosen from a pool of 907,774 teachers. From this trial run, Baumann and the researchers at the SRC were able to closely analyze “bias, clarity, and format” (p. 640). This trial run and subsequent changes to the survey demonstrated the objectivity and internal validity of the study. After the survey was reformatted, “the final Teacher Survey consisted of 53 closed items (multiple choice and short fill-in blanks) and two open items (that included 4 questions, each followed by 3 to 4 write-on blanks)” (p. 640).

The results of this study indicated that teachers valued the implementation of literature-based, literacy-immersion practices, and process-oriented reading and language arts instruction. Most teachers noted that phonics instruction is equally important as literacy immersion (whole language instruction), but that phonics must be implemented within a program that offers literacy as a number one priority. Baumann et al. (1998) stated, “As noted, 71% of the sample indicated that students should be immersed in literature and literacy experiences, and 94% concurred that it was their goal to develop independent, motivated readers who appreciate and enjoy literature” (p. 645). Baumann

et al. found that 89% of teachers prefer a balanced approach to literacy because it is crucial to help students achieve reading readiness.

Baumann et al. (1998) took measures to ensure that their study was reliable, objective, and internally valid. The results of this study were strong in internal validity because of the concise measures taken in the method of the study. Baumann et al. took precise measures to record their data-gathering method. The internal validity is strengthened because of the outside sources that helped construct the quantitative measures, do a trial run of the measures, and take in the data to measure. The external validity is strong as well because of the detailed demographics of the participants. The external validity comes into question because of the large sample size; however the detailed demographics overcome this shortcoming.

The analysis of these seven studies suggests that students benefit from phonics-based instruction; however, whole language instruction is not deemed ineffective. Dakin (1999), Donat (2006), Ryder et al., (2008), and Griffith et al. (1992) all claimed that explicit instruction in phonics and phonemic awareness is effective in helping students know how to read; however, the empirical evidence from all three studies demonstrated limitations to testing a students' success in phonics, not an overall success in the reading process. Dahl and Scharer (2000) found that whole language instruction with implicit instruction in phonics that did not disrupt the reading process was successful in helping students achieve readiness in reading. Findings by Baumann et al. (1998) and McIntyre (1995) suggested that teachers across the nation and within one classroom believe that students should receive a balanced approach to reading which implements whole

language instruction and phonics-based instruction to meet the underlying needs of the students.

### **Whole Language and Skills-Based Instruction**

The studies analyzed so far have investigated the effects of whole language instruction compared to classrooms that focus on phonics-based instruction. Prior studies also analyzed which instructional methods (whole language, skills-based, or a balanced approach to reading) were deemed by teachers to be more effective for student reading readiness. This section of chapter two explores the effectiveness of whole language instruction versus skills-based instruction on student achievement in reading. There are 13 studies in this section, eight of which focus on skills-based or direct instruction treatment, three that focus on basal-reader instruction, and two that focus on student achievement in reading and writing achievement.

The first study by Stahl, Pagnucco, and Suttles (1996), sought to determine which method, process-oriented or traditional oriented, would help students in their reading. Three additional studies (Dahl and Freppon 1994; Davis, 2010; Freppon & McIntyre, 1999) all sought to determine which instructional method, whole language or skills-based, was more effective in increasing students' reading proficiency. A study by Hurry, Sylva, and Riley (1999) wanted to determine which literacy program was more effective for student achievement in reading, a literacy-focused or non-literacy-focused program. The sixth study, by Pinnel, Lyons, DeFord, Bryk, and Seltzer (1994), sought to determine whether the Reading Recovery program was more effective than a Direct Instruction program. Costello (2011) sought to determine whether one program could

help all students in reading efficiency. The eighth study, by Stockard and Engelmann (2010), sought to determine whether Reading Recovery was an effective program to use to achieve reading readiness in students.

The next three studies focus on basal-reading programs. Hollingsworth, Reutzel, and Weeks (1990) focused on which instructional practices in reading are prominent within a first-grade sample. A study by Puorro (1997) focused on whether whole language instruction or basal-based instruction is better to achieve readiness in reading. The eleventh study, by Reutzel and Cooter (1990), focused on the difference between whole language instruction and basal-reader instruction on student reading. The last two studies, by Varble (1990) and Kamberelis (1998), focused on reading and writing achievement.

Stahl, Pagnucco, and Suttles (1996) conducted a qualitative study that followed two separate schools within the same school district. The main purpose of the study was to determine which approach, traditional or process, would help students achieve reading proficiency. Stahl et al. found that students within the traditional classroom were reading more challenging books compared to the process-oriented classroom. The process-oriented classroom demonstrated a better attitude towards reading and writing compared to the traditional classroom.

Stahl et al. (1996) studied 95 first graders from January to May of that particular school year. Two school districts within a southeastern city were chosen. Two separate schools were chosen and then six separate first-grade classrooms were chosen for the study. Three of the classrooms were identified as process-oriented and three were

identified as traditional classrooms. The demographics within each school were approximately 60% European American and 40% African American. Other demographics of the study were not reported.

The method of the study included structured interviews, a test called the Qualitative Reading Inventory (QRI), writing samples, and observations. Structured interviews included questions focusing on the purpose of reading and what exactly makes a good reader. Observations were conducted January-May of the school year. The researchers observed each classroom seven times. In each classroom, two different observers separately recorded data on student engagement. Another researcher observed individual classrooms seven times and took detailed field notes on literacy instructional times. Then the students took the QRI, which measured their ideas of what reading was, what its purpose was, and what their attitude towards reading was.

The qualitative measures found that each classroom had different practices students engaged in daily. A typical day in a traditional classroom involved heavy basal-reader work. Although instruction in traditional classrooms varied from teacher to teacher, the classrooms still worked with basal readers every day. In process-oriented classes, students engaged in whole-group reading, small-group reading, buddy reading or individual reading. The students read trade books and were able to choose their own books to read independently. These process-oriented classrooms were required by the school districts to use the basal readers in their instruction. Some of the process-oriented teachers used the basal readers once a week to teach certain skills. One teacher used worksheets in her instruction for groups of students to fill out. The qualitative measures proved that process-oriented classrooms had little or no sense of a hierarchy of better

readers and the students had a good attitude towards reading as compared to the traditional classrooms.

The results from the quantitative QRI revealed that the traditional classrooms did better on all areas assessed. The QRI assessed students in word recognition (WR), oral recognition (OR), and silent reading (SR). The traditional classrooms scored a mean score of 2.68 in WR compared to the process-oriented classrooms' 1.81 mean score. The traditional classrooms scored a mean score of 2.18 in OR compared to the process-oriented classrooms' mean score of 1.60. The traditional classrooms scored a mean score of 2.3 in SR compared to the process-oriented classrooms' mean score of 1.28. These quantitative measures proved that a traditional instructional method could be more effective in helping students learn how to recognize words.

This study exhibited weakness in the way Stahl et al. conducted the study. The researchers were conducting a study to find out which instructional method was more effective in helping students learn how to read. Unfortunately their quantitative measures fell short. Stahl et al. (1996) conducted a QRI on the students at the end of the study. This assessment was a major contribution to the findings within the study. Unfortunately, the QRI only focused on areas such as word recognition, oral recognition, and silent reading. In no way did the researchers measure the comprehension levels of the students. This failure to measure comprehension levels results in a gap in exploring the whole scope of the reading process. The focus on this assessment is primarily on phonemic awareness and phonics awareness. The authors' qualitative measures are strong due to their detailed and well-planned method. Researchers frequently switched observer roles while in each classroom. This role switching shows strong credibility in the study. Overall, the study's

qualitative measures were strong, but its quantitative measures are weak, therefore weakening this study.

Like Stahl et al. (1996), Dahl and Freppon (1994) studied the effects of whole language instruction and skills-based instruction on student achievement on students through their first two years of elementary school. Dahl and Freppon conducted a qualitative and quantitative study on two school sites that exhibited three main characteristics: “low income levels, most families received public assistance, and the schools’ mobility rates were high” (p. 55). The subjects were administered a pre-assessment prior to the study and a post-assessment following the two-year study. Dahl and Freppon found that the qualitative measures showed that the whole language treatment fostered a positive attitude towards reading in the students; however, the quantitative measures showed no significant difference between the two treatments when comparing the pre-assessments and post-assessments. Students also showed more strategy use during reading.

This study included participants from four classrooms, two whole language and two skills-based classrooms. Initially, 48 students were chosen for this particular study. The area that the study was conducted in included a high rate of diversity, Title I schools, and a high transient population. Out of the 48 initial students, 12 focus students were chosen from the whole language classroom, of whom six were male and six were female, six were African American, and six were European American Appalachian students. In the skills-based classrooms, eight focus students were chosen.



The method of this study included qualitative and quantitative measures to produce credible results that demonstrated achievement in students' reading and writing skills. After the 48 students were chosen from the four skills-based and whole language classrooms, the students were assessed on their written language knowledge. After this was done, the focus students were chosen for close observation by the researcher. The observations happened twice a week for two consecutive school years. The observations included close supervision of the focus students, one student per observation. The researchers kept detailed field notes which followed one student each day in reading and writing activities. Researchers also participated in the classroom by asking the students questions about their learning. Students were quantitatively assessed in six areas: 1) intentionality (written language as a system with accessible meaning), 2) concepts about print, 3) alphabetic principle, 4) story structure, 5) written narrative register, and 6) concepts about writing (Dahl and Freppon, 1994, p.57).

Dahl and Freppon (1994) administered a pretest to each student in both skills-based and whole language classrooms. These pretests included the six areas listed above. The students were tested at the beginning of the year and again at the end of the year. The tests were then used to find the mean and standard deviation, and were analyzed using a mixed-model comparison of variance (ANOVA) where the significance level was  $p < .05$  (Dahl and Freppon, 1994, p. 59).

The results of the tests revealed that there were no statistically different results for either group from pre-assessment to post-assessment with the exception of the written narrative register measure. The whole language group scored higher than the skills-based group on the written narrative register assessment, with an effect size of .07. Overall, the

results of the quantitative study determined that each type of treatment had the same results except for in written narrative register, where the whole language group did significantly better.

The qualitative measures established that there were a number of patterns that arose through the qualitative field notes. The patterns addressed are “interest in accuracy, phonics growth, response to literature, and coping strategies of learners experiencing difficulty” (Dahl and Freppon, 1994, pp. 60-65). In the field notes section on interest in accuracy, all students expressed this need to get it right. It was noted that students in the whole language classrooms felt less pressure to get it right, since invented spelling was supported by the teacher. On the other hand, the skills-based teacher put great emphasis on accurate answers. Phonics growth was consistent in both classroom types, though the whole language students seemed to demonstrate a better awareness of letter-sound knowledge. Assessment criteria for “Response to literature” in the whole language classroom demonstrated that not only were whole language students exposed to more literature, but they were also allowed to make connections to the text and make comments during story time allowing them to build skills in literacy. Coping strategies for when learners experienced difficulty were strong in whole language classrooms, but not in skills-based classrooms.

The internal and external validity of Dahl and Freppon’s (1994) study revealed consistency in method. This study was strong because the researchers recorded their information so that the study could be repeated in another setting. Also, this study involved a set of classes that were controlled and another set of classrooms that received a treatment. In addition to having comparative groups of participants, this study included

a pre-assessment and post-assessment that was the same to measure the success of each class. These assessments allowed the researchers to demonstrate the achievement of each classroom in a consistent way. The study was under public observation and involved many professionals and their opinions, which means that the study strived for unbiased opinions and researchers. The results of this study are reliable because of the potential to repeat the study and repeat the results. The study was highly descriptive and made clear the methods, setting, and participants involved.

Like Dahl and Freppon, Davis (2010) conducted a study that investigated which instructional method was better, whole language (student-centered) or skills based. However, Davis conducted a smaller qualitative study on 19 second-grade students from a western United States elementary school through a comparative study. Davis wanted to understand which instructional method, skills based or student-centered, would affect student achievement, attitude, and motivation while reading. Davis found that whole language philosophy caused students who experienced collaboration and democratic choice in the classroom to react positively to reading.

Two instructional models were implemented in one classroom consisting of a teacher and 19 students, of whom nine were European American, six were African American, three were Asian American, and one was Latino” (Davis, 2010, p. 58). This second-grade classroom consisted of “three students reading significantly above grade level, four students reading above grade level, seven reading at grade level, and five reading significantly below grade level” (Davis, 2010, p. 57). Data collection happened over a four-week period. School records showed that none of the students had had any previous experience with either the skills-based instruction or student-centered (whole

language) instruction. The researcher was actually the classroom teacher, which made her a researcher-practitioner. She had been teaching at the school site for three years. Her experience within the school and her familiarity with the instructional styles utilized at the school gave this teacher-researcher a deeper view in the study.

The method involved observation of the students while the researcher-practitioner was engaged in the classroom. Observations were made by the researcher-practitioner as well as multiple administrators and professional developers. Also this study was implemented in the middle of the school year. The student-centered and skills-based instruction were each implemented for a time span of 10 school days. Students were interviewed and took surveys three times during each 10-day implementation, once before, once in the middle of, and once after the implementation.

The method also involved semi-structured and informal interviews, student questionnaires, assignment evaluation, and focus groups. Interviews were launched prior to each instructional method implementation and after each method. These interviews were modified from the Motivation to Read Profile (MRP). The interviews encouraged students to give genuine insights from the students' perceptions of a particular instructional method. The MRP was also used to form the student questionnaires. These questionnaires were administered before the study to discern the students' attitudes towards reading. The student questionnaire included questions like "The reading activities I did in class this week made me want to read more" and "Choosing my own reading activities made me want to read more," providing data which the researcher used to gauge whether the skills-based instruction or student-centered (whole language) instruction was motivating students to have a positive attitude towards reading. In

addition to interviews, the students took part in a group interview after each instructional method's implementation. The study analyzed the responses to choice, collaboration, challenge, learner control, goal orientation, and environmental context of learning.

In the findings, Davis (2010) asserted that a students' interest in the activity affected their participation and engagement in the literacy task at hand. Students seemed more engaged in a literacy task when students had the opportunity to "make choices, collaborate, and share in control of learning processes and outcomes. Tasks reflective of these characteristics were associated primarily with the learner-centered curriculum" (Davis, 2010, p. 74). This study found that the student-centered (whole language) practices allowed the students to engage in their own education and make choices about that education and engagement. For example, in one of the questionnaires distributed to the students, 57% of students said that when given the opportunity to make a choice during literacy activities, they felt more positively towards that activity. The other 43% of students answered that comment with the statement *kind of agree*. None of the students disagreed with the statement. Also, students' actions were observed during instruction, and Davis recorded in her field notes when a student seemed off task or on task. Within the field notes it was noted that 14 out of 17 students were on task when allowed to choose their own books to read, as opposed to the seven of 17 students who were consistently engaged when the instructional method was skills-based. When students were able to engage in and take charge of their own education, their attitudes towards literacy, reading, and writing changed. This attitude change led to productive students who made academic achievements in reading.

The credibility of this study is low. The study was administered on the same students by the same teacher, which is a strong quality of the study; however, the study was only four weeks long. That means that each treatment was administered for only two weeks. This is a short amount of time to allow any sort of conclusion about the effectiveness of an instructional method on student achievement. The study shows strong transferability because of the detailed notes about the demographics of the students involved and geographic area it happened in. This study is dependable because it could be replicated due to the detailed method, small sample size, and detailed demographics. This study also demonstrated strong qualities within its confirmability. This study was conducted by the teacher in the classroom, but this does not allow different variables like various teachers and treatment groups. The confirmability is strong because outside people made observations within her classroom along with her. These outside eyes and voices contributed to this study's confirmability.

Like Davis (2010), Freppon and McIntyre (1999) studied the effects of whole language instruction and skills-based instruction on student achievement. However, Freppon and McIntyre focused on six students who attended urban public schools in an unidentified Midwestern city, using a semi-experimental method. Freppon and McIntyre found that the three pairs of focus students read similar texts conventionally at the end of the study and demonstrated similar progress in reading development; however, the level of courage, persistence, and application of strategies was clearly different in the constructivist-based whole language classroom.

This two-year study focused on two separate classrooms, kindergarten through first grade, from two urban elementary schools. The schools had elevated rates of

mobility and students who received free or reduced lunches. The two classrooms' teachers used different instructional methods, skills-based and constructivist, to teach their students. The classrooms were selected through interviews and classroom observation. The constructivist and traditional teachers were selected by principals who identified them as masters in their own instructional method.

Three focus students were selected from each classroom to compare their achievement in reading. These students were chosen because of their similar low socioeconomic backgrounds, academic achievement, and demographics. These six students were separated into pairs. Each pair consisted of one student from the skills-based instruction classroom (SB) and one from the constructivist-based (CB) classroom. These students were also paired because of their similar experience in learning. Audrey (SB) and Charlie (CB) were matched because they both had similar experience with learning, which was more advanced than the other participants in the study. Mary Ann (SB) and Jason (CB) were matched because their experience in learning was less than Audrey and Charlie's. Rodney (SB) and Ann (CB) had the least experience in learning, so they were paired up. These students were paired with evidence from assessments that revealed their alphabetic knowledge and written knowledge. Each pair's scores were either exactly or very nearly the same.

The method of the study included two weekly two-hour visits, non-participatory observations, and a particular focus on reading ability of one focus pair per week for two school years. Focus students wore microphones so that the researchers could analyze the reading material and the students' progress in reading more efficiently. The researchers "sat physically close to each child, recorded what he or she did and what his or her

teacher did, noted the child's interactions, and documented the materials used” (Freppon and McIntyre, 1999, pp. 7-8). Any written materials that were used in classroom instruction were photocopied and analyzed by the researchers. The researchers also administered reading pre-assessments and post-assessments. These assessments were analyzed by an outside researcher who had many years of experience as a Reading Recovery educator. This assessor analyzed the reading samples by using Clay’s Running Record. This assessment identified which reading cues students were using or not using while reading, “including oral reading errors, self-corrections, fluency, intonation cross-checking cues, decoding, and stance responses such as effort or persistence” (Freppon and McIntyre, 1999, p. 8).

Freppon and McIntyre (1999) found that the pairs from both instructional classrooms showed similar reading development; however the constructivist-based whole language classroom evidenced positive attitudes towards reading and more meaning-based strategies available to students while they read. The researchers found that the skills-based students did not have the same willingness to try something they didn’t know. The pairs of students from different instructional methods showed similar development in reading and were either at conventional reading level or could read more than a couple of books at conventional level. Therefore, the two teaching philosophies and instructional methods bore similar results except for the love of reading and the development of multiple strategies for reading.

This study demonstrated strengths in dependability, credibility, and confirmability. Freppon and McIntyre (1999) took measures to be extremely detailed in describing how the school district, teachers, and students were selected. This study is



credible because the researchers not only used observations as part of their method, but they also used pre-assessments and post-assessments to measure the students' progress in reading. This study also took measures to represent an unbiased study. The researchers hired an outside professional to assess their focus students' reading samples.

Freppon and McIntyre studied the effects of two separate instructional methods, while Hurry, Sylva, and Riley (1999) studied the effects of reading achievement in schools that were literacy focused and non-literacy focused. The main focus of this study was to uncover whether or not the literacy-program classrooms made more progress in reading compared to the non-literacy-focused classrooms. Hurry et al. (1999) found that the literacy program classrooms made more progress in reading compared to the non-literacy-focused classrooms as evidenced by post-assessments.

Initially, six literacy-based schools were selected for this particular study. Then, for comparison purposes, six more non-literacy schools were selected. The six non-literacy schools were chosen because of their similar scores on the 1993 KS1 Standard Assessment Tasks (SATs). The SATs test is the national test administered in the United Kingdom that measures seven-year-olds' knowledge in reading and other academic areas. In addition to the SATs scores, Reading Recovery tutors were consulted to attain information about the schools to be sure that the comparison schools were close in demographics and other details pertaining to the study. Information from the kindergarten classrooms was obtained through the Institute of Education research team and the students were assessed in September 1995. Other students who enrolled in January 1996 were assessed as well. At the end of the study (May 1997) a total of 179 students were

assessed. Students who were identified with special needs or spoke English very poorly were excluded from the study.

The two types of classrooms focused on in this study used two different types of instructional methods. In the literacy-based classrooms, a program that combined “whole language practices with a structured classroom organization, including a relatively high level of direct teacher instruction” was implemented (Hurry et al., 1999, p. 3). The non-literacy classes followed school policy on teaching literacy or the teacher’s instructional preference.

The method conducted in this study contained pre-assessments and post-assessments. The pre-assessments included assessment in areas such as British Picture Vocabulary Test (BPVS), The British Ability Scale Word reading test (Word Reading), Letter Identification (LI), and Concepts About Print (CAP). The post-assessments included all assessments listed above and a few additions. These additions included Book Level, Dictation (DICT), and Stuart Non-Word Reading (Non-Sense Word Decoding).

The pre-assessments administered to both literacy-based and non-literacy-based classrooms rendered similar results in reading. On the first post-assessment, Hurry et al. (1999) found that the literacy-based classrooms, overall, scored higher than the non-literacy-based comparison classrooms. The literacy-based classrooms scored higher on five of the six reading measures. Three of the assessments, Word Reading, Concepts About Print, and Letter Identification, showed statistically significant differences. The assessment that the literacy-based classrooms did not score higher in was the Stuart Non-Word Reading assessment. Hurry et al. noted that, “the literacy programme children were about 2 months ahead of the comparison children, on average” (Hurry et al., 1999, p. 8).

Comparing the progress of the literacy group and the non-literacy group, the researchers found that the literacy program group was performing at a significantly higher level than the non-literacy comparison group on the Word Reading test ( $t = 2.0$ ,  $df = 177$ ,  $p < 0.05$ ), which means that the students in the literacy-program classrooms did significantly better than the non-literacy-program classrooms. The literacy-based classrooms scored a 90% on this assessment compared to the 72% the non-literacy-based classrooms scored. Similarly, the literacy group performed significantly better than the non-literacy group on the Concepts About Print assessment. The literacy-based classrooms scored a mean score of 18.5 ( $SD = 3.2$ ) on that assessment compared to the non-literacy-based classrooms' mean scores of 16.6 ( $SD = 3.7$ ). These results indicate a statistically significant difference between the comparison group and the control group (Mann Whitney  $U = 2804$ ,  $p < 0.0001$ ). Assessment areas Word Reading, Letter Identification, Book Level, and Dictation rendered similar progress and results. The only area that the non-literacy-based classrooms scored better in was the assessment Stuart Non-Word Reading. The non-literacy classrooms scored a 6.3 compared to the literacy classrooms' 5.2.

This study is reliable because the researchers were detailed in their record-keeping and statistical analysis. The study has internal validity because the independent variable produced a change in the dependent variable. Furthermore, the dependent variable, literacy and non-literacy groups, either changed a little bit or drastically depending on the independent variable, instructional strategies introduced in the different groups. The transferability is strong because the variables could be introduced in a similar

setting, but has the possibility of not translating from England to the United States because of cultural differences.

Hurry et al. (1999) studied the effects of two separate school styles, while Pinnel, Lyons, DeFord, Bryk, and Seltzer (1994) conducted a quantitative study on 403 students to determine whether four different instructional models (three reading Recovery Program variations and one skills-based model) were more effective in students' progress towards reading readiness. The main purpose of this study was to determine if the program Reading Recovery was effective in helping students learn how to read proficiently. The findings suggested that one Reading Recovery treatment showed the most significant gains compared to the other treatment groups.

Pinnel et al. (1994) conducted a study on 403 students. The demographics of the study included 244 European American, 86 African American, one Asian American, and 72 students whose racial or ethnic backgrounds were not released by their schools. The sample included 238 male and 165 female students (Pinnel et al., 1994, p. 15). There were 10 school districts involved in this study: two rural, two suburban, and six urban school districts. Certain school districts could not release any information related to free or reduced lunches, so 131 students were unidentified as needing free or reduced lunches. However, of the other 272 subjects, 166 (60.8%) were receiving free lunch and 11 (4 %) were receiving reduced-price lunch. The teachers who were involved in the study were all teachers who had previously taught at the schools, so no special teachers were brought in to teach the special treatments. The teachers who did teach had experience in the Reading Recovery program, the Reading Success program, direct instruction, or the Reading and Writing program.

Reading Recovery is a program that is similar to whole language in that it helps students learn how to read in a holistic way without teaching explicit phonics skills. Reading Recovery also focuses on reading strategies to help students figure out a word they don't know while they are reading. Reading Recovery included various one-on-one interventions:

Children receive daily 30-minute lessons that follow a basic instructional framework of (a) familiar rereading of easy books, (b) independent reading of an instructional level text, (c) an optional component to develop letter knowledge, (d) composing and writing a sentence, (e) reconstructing a cut-up version of the sentence, and (f) talking about and then reading a new, more challenging text (Pinnel et al., 1994, p. 17).

As evidenced above, the Reading Recovery program is rigorous in its method to help students have the tools to read. Reading Success is a program that is similar to Reading Recovery, including the activities listed above, except the teachers in the study had different theoretical standpoints as evidenced by the Theoretical Orientation to Reading Profile (TORP). The Direct Instruction method was where teachers taught in a way that was skills-based and completely different than the instruction given in Reading Recovery. Although the Direct Instruction model was skills-based, the intervention time was always connected to what was happening within the classroom. Reading and Writing was a program that was framed after the Reading Recovery program, except it was used with groups of children instead of in one-on-one intervention. The study also included comparison groups. These groups were essentially control groups where teachers

continued to use whatever methods they had been using. The teachers also received no additional training like the treatment groups did.

The method of this study included a pretest and posttest design and is self-designated as a split plots design. Pinnel et al. (1994) stated that “each school was in essence a small, randomized trial for one treatment, and the study was designed so that each district provided a set of effect estimates for all four treatments” (p. 15). This study also included some qualitative measures such as observations during instructional time and video-taped lessons. The treatments included a Reading Recovery program treatment, a Reading Success treatment (modeled after Reading Recovery), a Direct Instruction treatment, and a Reading and Writing Group treatment. The treatments were also compared to four controlled comparison groups where the classrooms ran as usual with no treatments. Students were chosen for the study by the individual schools. The schools were asked by the researchers to choose students in a way that was randomized yet particular. The researchers were interested in treating students who were amongst the struggling students in their grade level. Those students were determined by teachers and school administrator and from there four students were chosen randomly from the larger set of low-achieving students. The measures used in this study included pretests and posttests. The pretests that were administered in October 1989 included assessments such as the Mason Early Reading Test, Dictation Task 1, and text reading level assessment. The posttests that were administered February 1990 included assessments such as text reading level assessment, Dictation Task 2, Gates-MacGinitie, and Woodcock Reading Mastery.

The results showed that the Reading Recovery program supported significant growth in reading readiness as compared to other treatments at the end of the actual study (not including follow-up measures which were done significantly after the treatments had stopped). When the pretests were administered, the study still had 324 students. When the posttest was administered the sample size was down to 283, which demonstrates a weakness in internal validity. However, some of the participants were dropped from the sample because of disregard for the researchers' directions on choosing students from the study. Three schools chose the neediest students in the first grade instead of choosing students randomly, so those schools were dropped from the study; therefore, internal validity is upheld. The Reading Recovery (RR) treatment scored a 43.32 mean with a standard deviation of 19.66 on their Mason Early Reading Test. The Direct Instruction group (DI) scored a mean score of 37.74 with a standard deviation of 25.81 on the same assessment. The Comparison Group (CG) scored a mean of 37.72 with a SD of 19.67. On Dictation Task 1 the RR group scored a mean of 6.29 with a SD of 5.36, the DI group scored a mean of 8.09 with a SD of 7.88, and the CG scored a mean of 8.17 with a SD of 6.38. The posttest in text reading level showed that the RR group scored a mean of 10.58 with a SD of 4.74, the DI group scored a mean of 4.31 with an SD of 3.21, and the CG scored a mean score of 4.72 with a SD of 4.04. These scores indicate that the only group presented here that made a significant gain in reading level was the RR group. This indicates that the RR group had a higher overall average reading level than the DI and CG groups. The posttest in Dictation Task 2 showed that the RR group scored a mean of 31.74 with a SD of 6.18, the DI group scored a mean of 24.17 with an SD of 10.45, and the CG scored a mean score of 26.33 with a SD of 7.91. These results, if compared to

Dictation Task 1, indicate that the RR group showed the most gains with a 25.45 average point gain compared to the DI group (16.08 gain) and the CG (18.16 gain). The posttest in Gates-MacGinitie 1 showed that the RR group scored a mean of 36.19 with a SD of 13.12, the DI group scored a mean of 32.23 with an SD of 9.74, and the CG scored a mean score of 31.93 with a SD of 10.46. This indicated that overall the RR group did better than the DI and CG groups. The posttest in Woodcock-R showed that the RR group scored a mean of 39.81 with a SD of 21.35, the DI group scored a mean of 34.16 with an SD of 25.60, and the CG scored a mean score of 30.64 with a SD of 21.56. This also indicated that the RR group did better on this particular assessment compared to the DI and the CG groups.

Overall, the quantitative assessments show that the Reading Recovery group did better on all posttests compared to the Direct Instruction group and the comparison groups. The qualitative measures indicate that the teachers are using methods in the Reading Recovery program that support student reading ability and use of independent reading strategies.

In this example, Dana is reading *Nick's Glasses*, an 8-page illustrated book about a boy who cannot find his glasses because he is wearing them. The text on page 6 says, "Have you looked behind the TV?" said Peter."

Dana read, "Have you looked under the...." She hesitated, glanced at the picture (which did not provide the needed information), and searched the line of print. Then she started over, "Have you looked behind the TV?" said Peter."



At the end of the page, her teacher quickly said, "I like the way you were checking carefully on that page. Show me the tricky part." Dana pointed to the word behind, saying, "It had a b." "Yes," said the teacher, "Under would have made sense. He could have looked under the TV, but that word couldn't be under. I also like the way you checked the picture, but that didn't help enough, did it?" (Pinnel et al., 1994, p. 22).

This example shows that the student was allowed to take chances and use strategies he had been learning without someone telling him what the word was.

The next example is from a Direct Instruction treatment. The teacher is helping the student sound out some words.

In the first example, the teacher had Jeffrey read a series of word cards that formed a sentence, placed right to left on the table in front of him. Indicating that Jeffrey was to start on the left, she pointed to each word and the student read it. She told him the first word, *Bryan*, and Jeffrey repeated it, saying, "Bryan, do you...." At his hesitation, the teacher made an /n/ sound and then told him the word. He went on, "know how to...." The teacher said, "Keep going," and made the /fl/ sound. The student said, "Fly..and...sing," finishing the sentence. Then the teacher said, "Ok, now, let's do them again....No, let's mix them up." The student read the words by taking them one by one from a pile and placing a star sticker on the words accurately identified. (Pinnel et al., (1994), p. 24).

The example given from a Direct Instruction treatment demonstrates a use of out-of-context reading behaviors and the use of what appears to be exclusively phonics. It

appears that the student is only relying on the teacher to help him and his knowledge of the alphabet. These two scenarios look very different. One student is gaining confidence in his ability to read and the other is simply sounding out words without making any connections. Overall, this study indicates that the Reading Recovery program is better in helping students learn how to read in a meaningful way and in making them proficient readers, when compared to the Direct Instruction group and Comparison Group, as evidenced by the quantitative measures.

This study has various strengths and weaknesses. One of the major weaknesses in this study is the internal validity. This study starts out with a certain number of students. As time elapsed, students left the area, therefore changing the number of students and affecting the quantitative measures. However, the internal validity is upheld because of the outside factors that affected the number of participants. The study lost participants due to moving, but the study also lost participants due to three schools' failure to follow instructions. These schools didn't choose their low students randomly; instead they chose the neediest students and inserted them into the treatments. In turn these schools were dropped from the study, making the objectivity strong. The internal validity is strong within its method because of the detailed measures taken to perform the study. The external validity is strong because the researchers recorded demographics with precision. This study is also reliable due to the pretest and posttest measures that were taken and the covariates and variables that were taken into account due to a participation drop discussed earlier.

Costello (2011) conducted a qualitative study that asked the question whether one packaged program (either Direct Instruction or Whole Language inspired) could benefit

all of his students' reading achievement. Costello observed his first-grade students' progress through a literacy program created by Fountas and Pinnell in their book *Guided Reading: Good First Teaching for All Children* (1996), which was used by his school's primary teachers develop curriculum and to track reading ability. Costello kept a detailed teacher research journal where he recorded events that happened in the classroom, his thoughts about the program and student progress, and any questions he had while engaging his students in this program. Costello found that a prescribed one-size-fits-all program is not effective. He found that these programs have positive strategies to help students learn how to read, but ultimately a "teacher's knowledge and experience" in the classroom accompanied with these strategies could be a "powerful learning tool" (Costello, 2011, p. 80).

Costello (2011) conducted a qualitative study on his first-grade classroom in the elementary school in which he was employed. The purpose of the study was to investigate whether or not the literacy programs that the teachers were instructed to use to teach students how to read were effective in reaching all students and in helping them persevere and learn how to read. As a teacher-researcher, Costello kept a research journal where he made daily logs of what he encountered during literacy instruction using the program that was prescribed to the educators to teach students how to read.

Costello (2011) engaged in this study due to the adoption of the Fountas and Pinnell literacy program at his school. The teachers were given a book which "centered on the use of guided reading in accordance with leveled text and suitable assessment tools" (Costello, 2011, p. 71). After this program was adopted, then the running record assessment tool was adopted in order to track students' reading levels and their progress.

Overall, Costello engaged in this study to find out whether or not a packaged program could really help students and whether a whole language, direct instruction, or balanced approach was more beneficial to students' developing reading strategies.

By analyzing his journal, Costello found that there were three prominent themes throughout it: reading levels, running records, and guided reading (Costello, 2011, p. 78). Costello was constantly questioned by administrators about the reading levels of his students in order to gauge the progress of his class. This frustrated him because of this focus on reading levels instead of on the students. Another entry demonstrated that reading levels fluctuated by interest. For example, a student scored level 13 when reading a book of no interest, but when asked to read a book of interest, the student scored a level 16 (Costello, 2011, p. 79). This finding frustrated Costello because the Direct Instruction mentality within education, which focused on levels and the consistent movement through each level, does not account for the affect of student learning.

Costello's study found that the use of a literacy program can be beneficial with the knowledge and skill of a teacher's intuition. He found that the use of a literacy program that was leveled conflicted with his whole language philosophy of enabling students to read books that are of interest to them instead of the prescribed, leveled reading books. He concluded that a balance of literacy instruction is needed to reach the needs of all students.

Costello's (2011) qualitative study is reliable because of his detailed notes throughout one whole school year. He kept track of everything he encountered while teaching and later went through his notes to find common themes. This study is strong in

its detailed notes; however, it shows weakness in various areas. For example, when describing his classroom, Costello failed to record the demographics of his own classroom, therefore making it impossible for others to replicate his results easily. Costello's study showed strength in describing his district and school programs and instructional tools. This study's results may be difficult to replicate amongst different types of demographics throughout the nation, but if the circumstances within the school are the same, the results may replicate.

Like Costello (2011), Stockard and Engelmann (2010) studied the effects of whole language implementation and skills-based instruction on student achievement. However, unlike Costello, Stockard and Engelmann conducted a qualitative study of 48 low-SES kindergarten and first-grade students in four classrooms in two Midwestern cities, using a semi-experimental study, which found that students from the whole language classroom showed significant use of specific strategies in reading and a positive agency. The research conducted in the skills-based classroom determined that the students learned the same skills, but did not view reading or writing as something that went beyond school.

Stockard and Engelmann (2010) focused on two groups of participants from two different areas of the United States. The first group of participants that were selected for this study was located in the Pacific Northwest. Two of the five elementary schools within one district participated in the study. One of the schools used the Reading Mastery (RM) program in grades kindergarten through second grade. The other school was considered a control school, where teachers were supported in their implementation of the *Open Court* program. A total of 68 students in the Northwest region participated in the study. Among these participants were 89% European Americans, 29% of students

who were eligible for free or reduced lunch, and 20% who were classified as eligible for special education (Stockard and Engelmann, 2010, p. 6). The other focus area was located in a rural Midwestern district. This particular district had a total of four elementary schools that had implemented the RM program. In this area, Stockard and Engelmann analyzed data from three cohorts. One kindergarten class was a whole language class, after which these students were placed in an RM class for grades one to three (Cohort 1 in the study). Another kindergarten class focused on a traditional curriculum, but in first grade transitioned to the RM program. A third kindergarten class started with the RM program in kindergarten and stayed in it through third grade (three different schools, including Cohorts 2 and 3). The demographics of this district were 32.8% racial and ethnic minority students (mainly Latino), and 50.6% qualified for free or reduced lunch. Only 7.8% were eligible for special education (Stockard and Engelmann, 2010, p. 7).

This study focused on the effects of the Reading Mastery (RM) program implemented into kindergarten students' reading instruction. The Reading Mastery program does not focus primarily on one isolated skill or subject, but in effect introduces 10% new information in each learning segment. The other 90% of each lesson focuses on skills and concepts that were focal in previous lessons. This study does not make it clear whether or not the reading mastery lessons imposed on the students include whole, authentic texts with inclusive skill building. A weakness of the study is a vague description of what the RM program entails and how it is implemented into the classrooms that were a part of the study.

The students were assessed through the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) system which concentrated on letter naming fluency (LNF), initial sounds fluency (ISF), nonsense word fluency (NWF), and oral reading fluency (ORF). The performance levels on the variables listed above were measured before the students entered kindergarten and were monitored through the end of these students' second-grade year. The progress was computed in a t-value probability chart. The Direct Instruction school at the Pacific Northwest site demonstrated a mean score of 16.1 and an SD of 13.5 in LNF and a mean of 17.4 and a SD of 12.5 in ISF compared to the control group's mean score of 17.8 and an SD of 13.0 in LNF and a mean of 14.9 and an SD of 10.0 in ISF. These scores indicate that there was no significant difference between direct instruction with the program Reading Mastery and the control group. The mean scores of the LNF test at the mid-western site in Cohort 1 showed a mean of 11.9 and a SD of 12.9, cohort 2 and 3 showed means of 11.0 and a SD of 12.2. The mean scores of the ISF test at the Midwestern site in Cohort 1 showed a mean of 9.0 and an SD of 8.2, while Cohorts 2 and 3 showed means of 8.4 and an SD of 7.5. These results also indicate no significant difference between each site and instructional method. The Pacific Northwest site showed progress in NWF. The control group started at 30 points in kindergarten and ended at 85 points in spring of first grade. The RM group at the PNW site began at 27 points in kindergarten and ended at 95 points in spring of first grade. This indicated a 10 point difference between the groups. The Midwestern site showed that the group that received RM in kindergarten (Cohort 1) started at 30 points in kindergarten and progressed to 85 points in spring of first grade. Cohorts 2 and 3 (both Midwestern sites) started at 23 in kindergarten and progressed to 65 points. This indicated that the sample with the

treatment of RM since kindergarten did better in NWF. On the ORF scores at the Pacific Northwest site, both the RM and the control group showed the same progress. The Midwestern site demonstrated that Cohort 1 started much higher at 30 as opposed to Cohorts 2 and 3 which were at -100. Both groups ended up at approximately 700, which indicated more growth within Cohorts 2 and 3. The overall results found in this study indicated that students who had reading instruction through RM from kindergarten on had significant success in reading.

There were various strengths and weaknesses in this study. This study has a questionable internal validity. Though the method was well described and well executed, the treatment used was not described well. The program Reading Mastery was never defined, except that it was a direct instruction model that uses explicit skills. The instructional methods that were used within the classrooms were also not defined in detail. The external validity is strong because of the detailed demographics and the way each site was chosen randomly. The dependability of this study was strong because it rendered results similar to other researchers' findings within the educational community. Although the results only indicated whether students can decode words and sound good while reading a passage, the study did not examine whether the participants comprehend what they read, only whether they can read it. Objectivity of this study is confirmed through its use of randomization of site selection.

The next three studies focus on the specific skills-based instruction called basal-reader instruction. Hollingsworth, Reutzel, and Weeks (1990) conducted a qualitative study that focused on 16 first-grade teachers. The main purpose of the study was to determine what whole language instructional practices in reading are most prominent in a



sample of first-grade classrooms. Hollingsworth et al. (1990) found that whole language instruction only makes up a small percentage of instructional time. Basal-reader and skills-based instruction account for approximately 70% of instructional reading time.

Hollingsworth et al. (1990) conducted a qualitative study that focused on 16 first-grade educators and their classrooms. The 16 teachers were chosen randomly from six school districts in Utah, one of which was an inner city district. The specific student or teacher demographics were not given.

The method Hollingsworth et al. (1990) used for this study was qualitative, including detailed interviews and observation by the researchers. The method included three separate processes of gathering data: 1) observers were to keep track of teacher behaviors during reading instruction by “identifying, timing, and describing” those actions, 2) observers identified and described “classroom environmental components,” and 3) observers engaged teachers in interviews (Hollingsworth et al., 1990, p. 15). The teachers chosen for the study were observed a total of five times during their reading instruction blocks by one observer. Observations occurred approximately three months from November 1987 to January 1988. The total amount of time observers spent in the classroom was 5897 minutes, which is approximately 98 hours.

Hollingsworth et al. (1990) found that only 10% of reading instructional time was “spent on practices consistent with whole language theory” (Hollingsworth et al., 1990, p. 16). The rest of the time was dedicated to instructional methods that were consistent with Behaviorist theory. For example, 17.29% of teachers’ instructional time was spent on basal readers, 13.07% on isolated phonics instruction, 11.08% on transitions, 11% on

looking over assignments, and 9.81% on non-instructional activities (Hollingsworth et al., 1990, p. 19). The researchers concluded that skills-based instruction is still a major contributor to instructional methods that are valued in the public school system.

Although compelling results were founded in this study, it demonstrated weaknesses far too large to overlook. For example, the study could not be redone in another setting because of the researchers' failure to list demographics of students and the teachers. Also the method of the study was vague and would be impossible to replicate.

Hollingsworth et al. (1990) sought to determine which instructional methods are used more often within first-grade classrooms. Puorro (1997) conducted a comparative study that asked that same question. Puorro conducted a quantitative study on two whole language and two basal-reader classrooms in California throughout one school year. She investigated these four classrooms and compared the achievement in reading between the two classroom types. Puorro found that the whole language class demonstrated better scores in reading comprehension. However, the basal-based classroom performed better in "word analysis, vocabulary, and spelling" (Puorro, 1997, p. 14).

Puorro (1997) conducted the study during the 1995-1996 school year. The total number of participants in this study was 95 first graders. Forty-eight of the participants were part of the two whole language classrooms and 47 of the participants were part of the basal-reader classrooms. The teachers selected for this study all had three or more years of experience teaching. Puorro interviewed the teachers to figure out their teaching preferences for whole language or basal readers. It is also important to mention that each

classroom engaged in the “Alpha One phonics programs” (Puorro, 1997, p. 6). Alpha One phonics is a program that conducts phonics instruction in isolation.

Puorro (1997) administered pre- and post-assessments in the academic areas of “reading comprehension, vocabulary, spelling, and word analysis” (Puorro, 1997, p. 11). These areas were the focus of the state mandated California Basic Skills test (CTBS). The pre-assessments were administered in class and used for the study. The results from the CTBS that was administered at the end of the school year were used as a post-assessment in order to gauge success in reading.

Puorro (1997) found that the whole language classes scored particularly higher in reading comprehension. The whole language classes had a mean of 19.1 and the basal-reader classes had a mean of 23.1 in reading comprehension. This four-point difference means that the t-value was 2.6 and showed significant statistical difference at the 0.01 level. Puorro also found that the basal reader classes generally did better in the areas of word analysis, vocabulary, and spelling. The basal classes overall scored better on their standardized test than the whole language students did.

This study was a weak study. The measures used to find achievement in reading between the two comparison groups were weak. Puorro (1997) used assessments that were administered by the teachers to compare to the post-assessment which was the state mandated standardized test (CTBS). This inconsistency in assessments shows a weakness in the scores that were used to show a change over time in the four assessment areas. Also, the researcher did not list the demographics of the students within the classes, which makes replicating this study impossible.

Unlike Puorro's study, Reutzel and Cooter (1990) studied the effects of whole language instruction and code emphasis during reading instruction on the achievement of first-grade students. Reutzel and Cooter, using an experimental method, conducted a quantitative study on 91 first-grade subjects: 53 were first-grade children in two whole language classrooms, and 38 were first-grade children in two basal-reader classrooms (Reutzel and Cooter, 1990, p. 253). At the conclusion of the students' first-grade year, researchers found that the whole language classroom showed statistically significant differences between their pre-assessment and post-assessment when compared to the basal-treated classroom.

In this study, Reutzel and Cooter (1990) assessed the initial reading ability of subjects by using the Gates-MacGinitie-R Reading Survey Test. At the end of the year a reading test survey, Gates-MacGinitie Form 1, Level A Reading Survey Test, was administered. "Reliability coefficients for these tests were computed from the standardized sample using the Kuder-Richardson Formula 20. The coefficient for each subtest was .93" (Reutzel & Cooter, 1990, p. 25). Reutzel and Cooter conducted their study within two school districts in Ohio and Utah. The study focused four first grade classrooms, two in Utah and two in Ohio. Two of these classrooms solely used whole language instruction and two of these classrooms solely used basal-based instruction. Basal readers primarily focus on skills-based learning within a district curriculum, exclusively focusing on one particular reading skill at a time. The whole language classrooms were "rich in print and print-oriented activities" (Reutzel and Cooter, 1990, p. 254).

The results from the comparison study indicate that the whole language classroom was more effective in the areas of comprehension, vocabulary, and overall scores. The whole language scored an overall mean of 67.83, 36.85 in vocabulary, and a 31.70 in comprehension. The basal students scored an overall mean of 61.87, 33.58 in vocabulary, and 28.10 in comprehension. The whole language group demonstrated a standard deviation of 17.76 as opposed to the basal group which had a standard deviation of 17.55. Reutzel and Cooter (1990) claimed that these results indicated statistical significance because the standard deviation represented enough of a difference to believe that the treatment was actually affecting the outcome of the study.

Reutzer and Cooter (1990) concluded that whole language must be effective in helping these students' vocabulary and comprehension levels rise. Reutzer and Cooter (1990) do advise that the whole language approaches taught in the study classrooms differed from teacher to teacher. Ruetzel and Cooter claimed that whole language had a moderate effect on students' achievement in first grade, but claimed a limitation of length and focus of the study. Since this study ran for one whole school year and only focused on first graders, Reutzel and Cooter concluded that the same result may not apply to other grades. If this study were conducted elsewhere, the results may be different due to demographic or sociocultural characteristics. Reutzel and Cooter suggested that the study could have yielded different results due to the type of teacher who taught in the whole language classrooms and the basal-reader classrooms.

The final two studies in this section primarily focus on reading and writing achievement. Varble (1990) conducted a quantitative study that followed 248 second and sixth graders on their achievement in writing structure. Students were in classrooms

identified as either whole language or traditional. The study's method was semi-experimental where a pre-assessment and post-assessment were conducted on the participants. Generally, the results of this study concluded that there was no statistical significance showing that either method of instruction was more effective in student achievement. However, the second-grade whole language class had the ability to create pieces of writing that established deep meaning and content in comparison to the traditional classroom's samples.

Varble (1990) conducted this study on a total of 248 students from elementary schools in various counties in Indiana. The second-grade participants made up 120 subjects in the sample population. Sixty-two of those students were in a whole language classroom and 58 of the 120 were in a traditional classroom. The sixth-grade participants made up 128 subjects in the sample population of subjects. Sixty-two of those participants were in a whole language classroom and 66 of those participants were in a traditional classroom. The teachers' instructional philosophy and instruction were determined by a survey of 10 questions, each one with two possible answers. If a teacher answered at least eight questions by indicating a whole language approach, then their classroom was considered a whole language classroom. If a teacher answered at least eight of the 10 with the more traditional answer, then their classroom was considered a traditional classroom.

The researcher conducted lessons that consisted of pre-writing activities and writing activities that produced writing samples for the researcher to use for her study. The lesson lasted for approximately 45 minutes. The second graders wrote about fantasy and the sixth graders wrote a mystery composition. Each student provided a sample for

the researcher to take and use for rating and quantifying. Varble (1990) was meticulous in creating the rating criteria. Three copies were produced, one for the student and two for the rater. Then the names were blocked and the raters were instructed to use a specific criterion of evaluation. The raters were teachers who were randomly selected in Vigo County, Indiana (Varble, 1990, p. 247).

The results showed that whole language instruction and traditional instruction rendered the same results in the area of mastery of mechanics, but in the area of quality of content, the whole language class had better scores. The researchers found that in the whole language second-grade class there was a statistically significant result of  $F = .0260$  in the area of quality of content. This was found to be statistically significant because the change over time yielded a large change in quantitative measures and in achievement in writing.

Varble (1990) conducted a weak study. The results found no significant difference between the two types of instruction except for outlying factors. This study could have been strengthened by a pre-assessment on the students' writing before engaging in the study. Varble (1990) compared the scores of individual writing samples by looking at how two different people rated the writing on criteria for evaluation. Instead of comparing the progress from before the study to after the study, the results were based on one writing sample that was evaluated by two different raters. These scores were deemed reliable through a rater comparison. Overall, if Varble (1990) compared writing samples from before the study to after the study, then this study would be stronger.

Kamberelis (1998) studied the effects of whole language instruction mixed with a skills-based instruction on student achievement within three different writing genres. Kamberelis conducted a quantitative study on 54 children from one kindergarten classroom, one first-grade classroom, and one second-grade classroom within the United States. Kamberelis found that primary students had an overall understanding of writing genres such poems, informational reports, and story.

Kamberelis' (1998) study involved students from the same school from three different classes. The demographics of this study included 16 kindergarten children (9 boys, 7 girls), 20 first-grade children (9 boys, 11 girls), and 18 second-grade children (8 boys, 10 girls). The group of students was 59% European American, 28% African American, and 13% Asian American. The three teachers who took part in this study all believed in the whole language philosophy; however, they all exhibited different instructional practices. All three teachers used reading and writing together in instruction. The students were exposed to trade books and often had the opportunity to choose what they wanted to write. Each classroom had a time specifically focused on phonics or skills-based activities, but more time was dedicated per day to these areas in kindergarten, less in first grade, even less in second grade.

The method of the study included qualitative measures and quantitative measures. The qualitative measures included observations and interviews with both the teachers and students. The observer was a participant observer who spent time in the classroom for a whole school year. This observer collected data during the spring of the school year. In this data collection, students were to write three compositions at three separate times. These writing compositions were in the genres of story, informational report, and poem.



All grade levels engaged in these tasks. Each writing segment took 20-30 minutes and was audio-taped. Prior to collecting writing samples, the researchers gathered information on how familiar each student was with each genre in question. Kamberelis coded and analyzed the writing samples by looking at “features of text structure, text cohesion, and text register” (Kamberelis, 1998, p. 23). These three features refer to the organization, flow, and phrases that are used only in certain genres. These three sets of data were analyzed using a MANOVA program. If a variable proved to be significant, then an Analysis of Variance (ANOVA) was used to see if the dependent variable was really significant.

The findings indicated that the participants from all grade levels excelled in story writing in all three subsets of text structure, text cohesion, and text register. All three grade levels did better in text structure in the story genre than they did in the genres of informational report and poetry. The mean percentages of every student’s score on text structure within each genre were presented. Kindergarten received approximately 85%, first grade received approximately 90%, and second grade received approximately 88%. Mean percentages for the genre report were put together according to organization and flow specific to informational report texts. Kindergarteners scored a mean percentage of 30%, first graders received a 70%, and second graders received a 70%. This indicates that first and second graders are much more familiar with this type of genre than kindergarteners. The mean percentage of poetic structure was then presented. Kindergarteners scored a mean percentage of 30%, first graders scored a 58%, and second graders scored a 62%. This indicated that first and second graders are more familiar with this structure than kindergarteners. These results indicated that

kindergarteners, first graders, and second graders are much more familiar with the structure and genre of stories as opposed to the informational report and poem genres.

This study had various weaknesses and strengths. This study showed strength in internal and external validity. This study could be replicated in another location at another time because of the detailed demographics that were listed. One shortfall of the external validity was the failure to mention the qualifications and demographics of the teachers chosen for the study. The internal validity is strong due to the detailed method. Kamberelis (1998) explained every step he took through this study. The objectivity of the study was strong due to the use of observer-participants to take field notes and interviews. Also the study is reliable due to the use ANOVA to analyze the variables involved in the quantitative analysis. Overall, this study is strong.

The analysis of these 13 studies overwhelmingly suggests that whole language and skills-based instruction demonstrated similar academic achievement. However, the whole language instructional method did render certain positive results besides academic success. In the first study by Stahl et al. (1996), the results indicated that the students in the skills-based treatment did excel academically beyond the whole language students, but the attitude of the whole language students proved to be more positive than that of the skills-based students. Dahl and Freppon's (1994) study rendered similar findings, that the academic success of the two treatment groups was approximately the same, yet the whole language class had a more positive attitude towards reading compared to the students in the skills-based classroom. Davis (2010) also found that when students were exposed to collaboration and democratic choice, the students reacted positively towards reading. Freppon and McIntyre (1999) also found that students in two separate treatment

groups (skills-based and whole language) experienced the same progress in reading development; however, the students in the whole language class possessed characteristics such as courage and persistence. Hurry et al.'s (1999) findings differed because the students in the literacy program within this study showed more progress in reading compared to students in the non-literacy programs. Pinnel et al. (1994) found that the reading program Reading Recovery significantly helped students achieve reading readiness. Costello (2011) found that one program, whether it be whole language or skills-based instruction, is not the fix-all answer to all students reading struggles. Costello recommends a more balanced approach to teaching reading which includes the teacher's own experience and knowledge to be able to help students succeed in reading. Stockard and Engelmann (2010) found that the program Reading Mastery, which is deeply rooted in whole language philosophy, was effective in helping students in academic success in reading. Hollingsworth et al. (1990) found that even though teachers claim to teach whole language classrooms, whole language instruction only made up a very small part of instructional time. Puorro (1997) found that students in the whole language classroom performed better in comprehension skills; however, the basal-reader classrooms performed better in "word analysis, vocabulary, and spelling" (Puorro, 1997, p. 14). In Reutzel and Cooter's (1990) study, they found that the whole language classroom performed significantly better on post-assessment than the basal-reader treatment classroom. Varble (1990) found that there was no significant difference between student academic achievement in reading and writing, though whole language students in second grade were able to write pieces with deeper meaning. Kamberelis (1998) found that

students within a whole language class with some skills-based variation were able to succeed at writing and understanding genres in writing, particularly in story writing.

### **Whole Language in Primary Education and Academic Success**

The studies analyzed so far have investigated the effectiveness of whole language instruction compared to phonics-based instruction and skills-based instruction (basal readers and direct instruction). Prior studies analyzed also studied teachers' evaluations of which instructional practices helped students become proficient in reading and which ones did not (whole language, skills-based, or a balanced approach to reading). This section of chapter two explores seven studies that followed the academic success or shortfalls of students who were exposed to whole language instruction. The first study, by Ketner, Smith, and Parnell (1997), studied the connection between developmentally appropriate practices and a teacher's philosophical stance (whole language). Freppon (1994) sought to prove that whole language experience and instruction could alter the achievement of one student in a case study. The third study, by Grisham (1993), sought to determine whether students in a whole language or traditional classroom would perform better in reading. A study by Traw (1996) was similar to Grisham's study because he set out to see whether a whole language or phonics-based classroom would show better achievement on a standardized test. The fifth study, by Sears (1999), examined whether first-grade students in a first-grade whole language classroom would show progress in reading. A study by Song and Miskel (2002) sought to discover the opinions of policy makers who made decisions on important policies made in the educational world. The final study in this section is by Morrow, Tracey, Woo, and Pressley (1999). This study strived to determine which instructional method nurtured achievement in reading.

Ketner, Smith, and Parnell (1997) studied the correlation between teachers' ideas of developmentally appropriate practices and their theoretical preference. Ketner et al. conducted a quantitative study on 156 teachers within a single Midwestern suburban school district. The researchers found that K-2 teacher's ideas about whole language philosophy and instruction line up with developmentally appropriate teaching practices.

Ketner et al. (1997) gathered instructional information and philosophies from teachers in a Midwestern suburban school district. Researchers randomly selected 156 teachers from this district to take the survey that included the Primary Teacher Questionnaire (PTQ) and the Theoretical Orientation to Reading Profile (TORP), as well as some further demographic questions. Teachers had a choice to participate in the survey, and their answers were anonymous. Out of the 71 completed surveys the researchers received, 66 were used for the study. The final sample consisted of 18 kindergarten teachers, 12 first-grade teachers, 20 second-grade teachers, and 16 third-grade teachers. White and female describe what the majority of the participants' demographics were. Fifty-seven percent of the teachers had been teaching for 10 years or more, 24% had been teaching 6-10 years, and 17% had been teaching for less than six years.

Teachers were administered The Theoretical Orientation to Reading (TORP) Profile which contained 28 statements which teachers were to answer on a scale from 1 to 5 assessing the degree of agreement or disagreement the teacher had with the statement. The Primary Teacher Questionnaire (PTQ) contained 42 statements to which teachers were to state their degree of agreement on a 4-point scale. The PTQ was distributed to assess teacher endorsement of "developmentally appropriate and traditional practices in

early childhood classrooms” (Ketner et al., 1997, p. 6). The Statistical Package for the Social Sciences (SPSS) was utilized to examine responses by the participants. TORP and PTQ scores were analyzed using a correlation analysis. These scores were then analyzed using a two-way analysis of variances using dependent variables which were the Developmentally Appropriate Practices (DAP) and the Traditional Practices (TRAD) (Ketner et al., 1997, p. 7).

Ketner et al. (1997) predicted that there would be a positive association to developmentally appropriate practices and the whole language approach in teacher’s responses to the surveys. As evidenced by correlation scores with the Primary Teacher Questionnaire (PTQ), whole language instruction is a popular instructional method utilized by primary teachers ( $r = .46$ ). Phonics-based instruction ( $r = -.61$ ) and Skills based instruction ( $r = -.36$ ) correlated negatively with the PTQ. Ketner et al.’s prediction was correct, as evidenced by a strong correlation in TORP scores and the PTQ component DAP scores, with  $r = .47$  ( $p < .01$ ) and TRAD scores,  $r = .52$  ( $p < .01$ ). This meant that Ketner et al. found that there was a strong connection between whole language instruction, developmentally appropriate practices, and achievement in reading.

This study showed strong evidence in internal and external validity, because it is generalizable to other sample populations. This study also demonstrated strong evidence by using a consistent independent variable and a changing dependent variable which determined a correlation between reading achievement and perceived beliefs about whole language instruction. A limitation of this study was the narrow demographic of all white, female teachers. The perspective of teachers of different cultural backgrounds and genders would increase the transferability of the study.

Ketner et al. (1997) studied the instructional methods that were deemed effective according to teacher opinion; however, Freppon (1994) conducted a primarily qualitative study that followed one participant through a tutoring experience that focused on helping the participant achieve readiness in reading. The purpose of this study was to prove that a whole language instructional method could help the participant become a proficient reader and help him gain a better attitude towards reading. Freppon (1994) found that the switch in instructional method, from direct instruction to whole language instruction, enabled the participant to connect meaning to reading, to understand why it was important, and to achieve academic success in reading.

Freppon (1994) studied whether whole language instruction during tutoring sessions would help a student on academic achievement in reading and writing. The participant was a nine-year-old student male from a low-income family, who was a student in an urban elementary school. This student was recommended to a reading tutoring program at the local university to support him in reading. His fourth-grade teacher recommended him to this literacy program because she claimed that reading and writing were his weakest areas in school. In an interview, the participant's fourth-grade teacher stated that the student could not stay on task for more than a few moments at a time and that he had Attention Deficit Disorder (ADD); therefore the teacher was in the process of getting him help through tutors and doctors. Standardized test scores revealed that the participant went from the "stanine of 6 in first grade to 3 by the end of third grade" (Freppon, 1994, p. 230).

Freppon (1994) utilized many methods to record important information while engaging in this study. Freppon conducted interviews with the participant, his parents, his

teacher, and his tutor. The researcher also conducted observations during tutoring sessions. Tutoring sessions included activities that included read-alouds, think-alouds, and observations. The researcher was careful to note whether the responses the participant supplied to the tutor were consistent with what he actually enacted while reading and writing. Observations were conducted at the tutoring sessions and within his fourth-grade classroom. This study was conducted for approximately three and a half months. Within the three and a half months, the researcher attended all tutoring sessions and attended the participant's classroom twice throughout the study.

As Freppon (1994) began her study on this young participant, initial interviews revealed that this student had minimal exposure to print and literature in his home experience. Interviews and read-alouds also indicated that this participant had a negative perception of reading. For example, this interview response indicates the participant's view of the functionality of reading and his attitude towards it.

Tutor: Are there some things you like about reading?

David: They help you get a good education.

Tutor: How about writing?

David: It helps you learn more.

Tutor: Would you say reading is a hard thing to do?

David: Yea, if you don't pay attention to the teacher. (p. 231)

As evidenced by the interview responses David gave, he showed no conception of the reasons behind reading and writing. He did not understand that reading and writing are ways that people communicate. His understanding of reading and writing was that the teacher assigned it and you had to do it. He also stated that reading is hard to do when



you don't pay attention to the teacher, which indicated that he relied heavily on the teacher's instructions to read and write. Otherwise he would not engage in it.

As tutoring sessions went on, it became apparent to the tutor and the researcher that David had minimal strategies to help him while reading. His primary strategy that he used was sounding out the words. When reading any passage he focused on getting the words right, but could not make accurate or meaningful predictions or summarize the passage. It became apparent that David did not attend to meaning while he was reading. It was noted by the researcher that the participant exhibited avoidance activities while engaged in reading and writing. David would concentrate on a reading assignment for a couple seconds, and then he would engage in social activity. This indicated even further that the participant had a negative association to reading, to the point where he would avoid the task at all costs.

Freppon (1994) found that the participant suffered from multiple factors that contributed to his challenges in reading and writing. First, David had had minimal exposure to print in his younger years. David also exhibited fear of failure and would strive to get words right. David also exhibited behaviors that indicated that he could not connect to reading in a meaningful way, which also hindered his achievement. Overall, a combination of ineffective reading strategies, minimal exposure to print, attitude towards reading, and fear of failure contributed to his performance in class and on standardized tests in reading.

Freppon (1994) found that the whole language instructional method helped David discover effective strategies to help him read and make meaning associated to a text.

Over the three and a half month study, David demonstrated great improvement in his reading ability. David became involved in many reading and writing events which were meaningful to him. For example, he read books that he was interested in (so he could construct meaning), he wrote letters and books that were personal to him, and he wrote and illustrated comic strips. David's performance in class excelled from consistent avoidance of reading and writing to minimal avoidance. Freppon (1994) proved that whole language instruction with struggling readers could help them become proficient readers who understand the functionality and meaning of reading and writing. When David was able to connect meaningful circumstances to his reading and writing, he flourished within those two academic subjects.

This study demonstrated various strengths and weaknesses. The confirmability of this study is believable because of similar results rendered by similar studies within the educational community. The dependability of this study is strong because the method was detailed. Research observations were clear and were also reliable because of the use of the constant comparative method that was used to analyze the field notes. The transferability of this study is weak because the demographics of the study were not disclosed. Even though the study focused on one student, this student's ethnicity was not disclosed. The parents' and the teacher's ethnicities were also not disclosed. The credibility of this study is strong because of detailed method the researcher used. The detailed notes and transcripts from the study demonstrated a change in this participant's views and achievement in reading.

Freppon (1994) studied the effects whole language instruction would have on one student; however, Grisham (1993) studied the effects of whole language instruction on

student academic achievement in a comparative quantitative study. Grisham conducted a quantitative study on two whole language and two traditional teachers in four intact suburban fourth-grade classrooms in southern California, using an experimental method. The author found that there was no significant statistical difference in academic achievement between the whole language and traditional classroom groups.

The participants in this study include two whole language and two traditional teachers. These four classrooms were located in Southern California in three suburban K-5 elementary schools. The demographics included lower middle-class and middle-class students. The demographics of the classrooms were similar, with approximately 80% European American students, 15% Hispanic/Latino students, and 5% students from other ethnic groups. All three schools exhibited above-average scores on the Stanford Achievement Test.

The method included procedures such as observations (mid-July, mid-October, and mid-February), pretests and posttests in reading and writing, and reading achievement tests. Also, teachers were interviewed and administered the Deford Theoretical Orientation of Reading Profile (TORP) to determine whether or not their classroom instruction and philosophy was whole language or traditional. After teachers were selected, Grisham collected curricular items such as lesson plans. When the study was conducted, qualitative and quantitative measures were taken. Student attitude and achievement were quantitatively collected. Then qualitative measures were taken through observations and interviews with teachers and school administration. Observations were conducted three times throughout the school year: once in mid-July 1992, another time in mid-October 1992, and a final time in February 1993.

Grisham used quantitative measures to determine student attitude and achievement in reading. Student attitude was measured by the Elementary Reading Attitude Survey. This attitude assessment was administered two times, once in July 1992, and another time in January of 1993. The assessment administered to measure student achievement in reading was the Nelson Reading Skills Test. This assessment was administered twice to the participant students, once in July 1992, and another time in January 1993. In this test students were assessed on their ability to identify synonyms and ability to summarize the passage (reading comprehension).

Grisham (1993) found that the Nelson Reading Skills Test demonstrated that there was no statistical significance between the four classrooms to indicate one instructional method was better than the other. For example, in comprehension one of the two whole language classes showed a positive mean score. One traditional class also demonstrated a positive mean score. The whole language and traditional class that showed negative mean scores showed only small changes. Their mean scores went from 56.22 to 53.92 (WL) and 53.39 to 49.81 (T) showing only -2.3 (WL) and -3.58 (T) variance. The assessments administered to measure students' attitudes towards reading showed scores of half the classes gaining a positive attitude towards reading while the other demonstrated a gain of a negative attitude towards reading. The same classes that showed a decrease in reading comprehension also showed a decrease in a positive attitude towards reading. One of the classes was a whole language class and the other was a traditional classroom.

The findings of the study suggested no statistically significant differences between the whole language classroom instruction and traditional classroom instruction. The attitude surveys demonstrated a better attitude from the students in one whole

language and one traditional classroom. The reading assessment had similar results to the attitude classroom. The same whole language and traditional classroom showed positive gains in all areas of study, especially comprehension, whereas the other whole language and traditional classrooms showed negative means scores in all areas. “The student of skills-based teachers who augmented their programs with discrete skills instruction scored no better than the students of whole language teachers. Means on the total reading test ranged from 48.03 to 52.58 and showed that students from all four classes were roughly equivalent in reading skill” (Grisham, 1993, p. 121).

Overall, Grisham conducted a strong study which found that neither instructional method could render statistically significant scores. All avenues taken to conduct this study demonstrated professionalism, reliability, objectivity, and validity due to its quantitative and qualitative measures. The internal validity of the study was strong because of the deep detail that was taken to the method and data analysis. Grisham (1993) implemented the pre-assessments and post-assessment research design, and also implemented qualitative measures such as observations and interviews. The external validity of the study was strong because of the detailed demographics of participants which could be duplicated in a similar setting. The objectivity of the study was strong because its methods were public and could have been agreed upon by other investigators.

In a comparative study, Grisham (1993) studied the effects of whole language instruction and traditional instruction on student achievement in reading. Traw’s (1996) study is similar in design, yet investigated the effects of whole language instruction and phonics-based instruction on student achievement in reading. Traw (1996) conducted quantitative measures on students’ performance on test scores pre-dating the

implementation, while the implementation took place, and after it happened. Qualitative measures were taken on teachers, principals, and other administrators. Traw found that there was no momentous transformation in achievement in reading as evidenced by the standardized test scores.

Traw (1996) chose the site for this specific study because of his familiarity with the two districts' affinity for whole language philosophy. The two districts chosen for the study were located in Sioux Falls and West Des Moines. These schools were also chosen for their whole language curriculum. This curriculum closely fit the definition of whole language that the author constructed. The author defined it as an instructional method which uses real authentic texts, but which also has traditional skills embedded within the instructional method. The Sioux Falls School District contained 18,000 students. The demographics of this area included mostly middle class families with some lower- and upper-class families. The demographics appeared to be mostly European American, with only 5% being part of an ethnic minority group. The West Des Moines School District consisted of 7,700 students. The demographics of this area were similar to Sioux Falls, with only 6% of people part of an ethnic minority. Most of the people living there were part of the upper-class, with some middle- and lower-class families.

The method of this study included quantitative and qualitative measures. The quantitative measure included a comparison of students' performance on the Stanford Achievement test (SAT). These tests were administered to students in the fourth, eighth, and twelfth grades. Scores of the fourth-grade students were accounted for in this study because the scope of the instructional method pertains to learning how to read. Learning how to read usually happens in the primary grades. Qualitative measures included

interviews with administrators, with three teachers whose philosophies aligned with whole language, and with three teachers whose philosophies did not align. Questions focused on subject matter, instructional methods for teaching reading, teachers' use of standardized test scores to help assess students, and their thoughts on the new curriculum.

Results from the SAT in the Sioux Falls School District indicated no significant change from traditional instruction to whole language instruction. The graphs indicated that before the implementation of new norms and whole language, the schools scored an average of 82-84 on the SAT in fourth grade. When new norms were established, the average score dropped 10 points to a 72. From 1988-1993, the SAT scores indicated a consistent score of 72 with some gains and drops. The comprehension scores showed a gain, rising from an average of 70 in the 1980's to an average of 77 in 1993. This indicates that the whole language implementation led to real gains in comprehension.

The results West Des Moines showed more of a variance in performance as compared to the Sioux Falls district. Test scores on the year the new curriculum was implemented (1987) showed a fourth-grade score of 46.6. These scores fluctuated the following years, demonstrating a 46.5 score in 1988, 47.5 in 1989, 46.9 in 1990, 45.4 in 1991, and 53.1 in 1992. These scores indicate inconsistencies throughout each year with gains and drops in the average score; however, in 1992 there was a 6.6 point gain from when the curriculum started to the most recent test within the study timeline.

In this study, Traw (1996) provided a description of the sample population that focused on a specific geographical area, but did not specifically list the demographics of the studied population. The demographics described offered what the dominant culture

was, but did not list the other demographics of that population, therefore making it a weakness in external validity and reliability. Traw concluded that, “although the limitations of the study reveal a need for more research, standardized test scores in districts that make a philosophical commitment to whole language do not appear to change significantly” (p. 339). Traw also found that, when twenty-one teachers and administrators were asked how often they used qualitative measures (observations) to assess their student’s literacy learning, twenty respondents answered that they used qualitative measures to a high degree. The other person used qualitative measures to a moderate degree when assessing literacy learning. The researcher asked these 21 administrators and teachers how often they used standardized tests as a way to assess their students; only four said they use them to a high degree, four said they used them to a moderate degree, six said they used it to a small degree, and seven said they did not use standardized tests at all. This indicated that standardized tests were not a tool that teachers used often to assess literacy skills. However, teachers and administrators used qualitative measures to assess their students in literacy.

This study had many strengths and weaknesses. A weakness of the study is in the external validity. This study could not be transferred to another location because of the lack of detail in the demographics of the study. The internal validity of the study is questionable because of the many variables involved. There was an unstated number of teachers involved in this study, even though focus teacher and administrators were chosen for qualitative measurements. The fourth-grade standardized tests were ever changing. The number of students fluctuated as the years went on. Traw followed and graphed the standardized test scores from 1979-1993. Also, the students being tested



changed each year. A strength of this measure was the detailed steps in the method that the researcher followed.

While Traw (1996) and Grisham (1993) conducted comparative studies to determine whether one instructional method was more effective than whole language instruction, Sears (1999) conducted a study that only focused on a whole language classroom and its students' achievement. Sears (1999) conducted a semi-experimental quantitative study that followed 15 first-grade students and their progress in reading from November through May. Sears conducted this study within a primary school within a district in Southern California. Sears studied the results from a reading running record she administered to the students which gauged which reading cues they were using. Sears' results suggested that the 15 first-grade focus students in a whole language classroom demonstrated the use of many graphic, semantic-syntactic, and phonological reading cues. Also, Sears (1999) found that even though students were not taught a structured or isolated phonics instruction, these students were scoring high in graphic cues which constitute phonics awareness.

Sears' (1999) study followed 15 first-grade students, who were in a classroom identified as a whole language classroom, for the duration of seven months. Sears visited a primary school within a Southern California school district a total of eight times to gauge the reading progress of eight first-grade boys and seven first-grade girls. The subjects in this study had previously attended kindergarten, and all spoke English as their first language. The classroom was described as a whole language classroom. Not only was reading encouraged, but "writing activities accompanied each lesson and the use of invented spelling was encouraged" (Sears, 1999, p. 95). Sears indicated that the

classroom activities did not contain “structured phonics instruction,” but the teacher did encourage students to give some attention to “sound-symbol relationships in their writing” (p. 95).

Sears (1999) studied the students’ miscues (the reading cues that students are using while making a mistake while reading) during eight separate observation times. This activity consisted of students reading in front of the whole group with the teacher during classroom instruction. The range of the reading selection was 20-226 words. The range of reading errors included a total of 478-734 errors and 26-106 errors per person. Sears (1999) found it important to note that when each observation took place, there were occasions where fourteen out of fifteen students were there, which made the miscue count vary between students.

Sears concluded that “overall, 514 of the 651 response errors, or 80%, were similar graphically, and 276, or 42%, [were] contextually acceptable. Moreover, 29% of errors or miscues were consistent with both sources of information and only 7% represented use of neither phonological nor syntactic-semantic cues” (p. 97). Overall, Sears (1999) found that students were using an array of reading cues. She also suggested that the cue that was used the most was the phonological cue because students would use initial consonant information to help them figure out what a word was. The miscue analysis indicated that 6% of high-progress readers and 12% of low-progress readers’ results were not “graphically or contextually acceptable” (Sears, 1999, p. 100). Her results also indicated that students were relying heavily on graphic cues where high-progress readers scored 88% and low-progress readers scored 68%. Overall, this study demonstrates that, even though the teacher does not concentrate or intentionally teach

isolated phonics skills, students are still learning how to gain (with internalized alphabet facts) phonics awareness.

This study demonstrated that whole language instruction helped students achieve a foundation for reading, but the study exhibited limitations. One limitation is that the study size was small, limited to 15 students. Another limitation is that the demographics of the students were not listed. The study demonstrated strong detail in method and produced reliable results. This study, like many other studies involving whole language, may be difficult to replicate in a similar setting. Whole language is defined several different ways. The definition of whole language is different to many different teachers and is enacted many different ways in many various classrooms.

Sears studied the academic achievement of 15 first-grade students in a whole language classroom. Conversely, Song and Miskel (2002) studied the results from a survey administered to policymakers involved with policies that strongly affect the educational world. Song and Miskel conducted a quantitative study on 103 policymakers who were members of education organizations, using a semi-experimental method. The researchers found that, “overall, interest groups believe that both phonics and whole language should be included in reading instruction, but do not think that either approach should be exclusive basis” (Song & Miskel, 2002, p. 12).

One hundred and three policymakers were chosen for this study. These policymakers were part of particular interest groups that were not described in the study. These people were selected randomly and their names and organizations were changed in the study. “The final sample thus consisted of 103 out of 118 (87.3%) policy actors, with

74 of the participants representing 67 interest groups. Another 31 representing 20 government agencies, and 2 representing both interest groups and government agencies” (Song and Miskel, 2002, p.7). The anonymity of these policymakers was important to the study, but presented a possible weakness. The policymakers were part of interest groups that influenced national reading policy. The primary method of this study included interviews about the interest groups’ perceived influence on national reading policies and their beliefs on reading instructional approaches. “Of the 107 participants, 55 were interviews in person, 50 via telephone, and 2 responded to the interview questions via email” (Song & Miskel, 2002, p. 8). The interviews were recorded and then transcribed into written text. All participants were kept anonymous.

Song and Miskel (2002) made two hypotheses while conducting this study. The first hypothesis asserted that the interest groups would show deep-seated sponsorship for balanced literacy instead of support for other instructional methods (phonics or whole language). The second hypothesis expressed the assumption that interest groups have varying ideas about different instructional methods (balanced methods, phonics based, and whole language based) and have more political influence in these particular methods. Hypothesis 1 was tested by performing a pair-sample-t-test. Of the organizations that were polled on the balanced approach to reading versus the phonics-based approach to reading, 73% of organizations sided with the balanced approach. Of the organizations that were polled on the balanced approach to reading versus the whole language based approach to reading, 79% of organizations sided with the balanced approach. Hypothesis 2 was tested by performing a variance analysis (ANOVA) to compare three instructional methods. The results of the interviews indicated no significant difference between

instructional methods preference (phonics based, whole language based, or a balanced approach). The second hypothesis was disproven because the interest groups did not demonstrate significant perceived influence over policymaking. Therefore, the policymakers did not believe in one approach specifically, but did advocate for using certain aspects from each method to create a balanced approach.

The study conducted by Song and Miskel (2002) revealed strong evidence that a balanced approach to reading is supported by people who make policies about education. The study had a strong internal validity and credibility because of the detailed method design. The external validity of the study was weak due to the ever-changing politics of education and the stances of the policy holders. The objectivity of the study is weak due to the anonymity of the policy holders and their organizations.

Song and Miskel (2002) found that a balanced literacy approach was favored by national reading policy makers and organizations. Morrow, Tracey, Woo, and Pressley's (1999) study is similar in that they asked which instructional method was more effective in student reading achievement. Morrow et al. studied what characteristics of first-grade teachers made students learn how to read efficiently. The purpose of this study was to uncover what instructional method, teaching philosophy, and classroom environment nurtured achievement in reading. Morrow et al. found that all six teachers used a balanced literacy approach to teach students how to read proficiently.

Morrow et al. (1999) conducted a qualitative study on six first-grade teachers. The teachers chosen for the study were from school districts within New Jersey. Teachers were chosen through interviews with administrators. Administrators were advised to

choose a teacher who showed excellent academic and professional success within the previous five years. Morrow et al. (1999) specifically defined this success as excellent student reading scores in all assessments and impressive teacher evaluations. Morrow et al. requested feedback not only from the administrators, but from other teachers and parents as well. After this was done, the administrators were asked to take a small survey which clarified the administrator's evaluation of the educator. All teachers received a master's degree and had nine or more years of experience in the field. The demographics of the schools that were a part of the study were 50% European American, 20% African American, 10% Hispanic, 10% Asian, and 10% from various other backgrounds.

The data was gathered by conducting observations and keeping detailed and organized notes on those observations. Morrow et al. (1999) visited each teacher's classroom for two whole days eight times during literacy instruction. While the researchers were in the classrooms, they took meticulous notes on "literacy instruction such as the schedule of the language arts block, word analysis instruction, comprehension development, language development, assessment strategies, social interaction during literacy instruction, affective teaching characteristics, student engagement, classroom management, and physical environment" (p. 3). Also, informal interviews were conducted with the teachers to find out where their teaching philosophies lay. The interviews revealed that each teacher used a philosophy to teaching reading that resembled a balanced approach which included methodology from whole language and skills-based instruction.

Morrow et al. (1999) found that the observations proved that a balanced approach to reading was effective in students' engagement and achievement in reading. The

detailed observations noted that classroom environment was positive. Each classroom was literacy-rich with many supplemental materials such as an alphabet chart, weather chart, and calendars. The classroom was organized into learning centers where students could engage in whatever content area the centers focused on. These learning centers were where students could go to get all the materials they needed to engage in literacy and content areas. Also students were able to engage in many types of reading situations throughout the week. Students engaged in read-alouds, guided readings, and independent reading. Students were also taught explicit skills during literacy blocks. An example of this is when one of the teachers engaged the students in a book they had already read and were very familiar with. These students had already built comprehension and had read the book thoroughly. After this was done, the students engaged in a skills-based lesson where the students got sentence strips out of order to put in the right order. This particular skill was enacted within an authentic text to which the students had already connected meaning. Therefore, these six teachers were proven to use a more balanced approach to teaching literacy.

This study revealed many strengths and weaknesses. This study demonstrated strength in transferability. The demographics were stated clearly and the study could be replicated elsewhere. The dependability of the study was demonstrated through its hypothesis of whether teachers who were identified as exemplary could demonstrate similar teaching practices to achieve student achievement in reading as agreed on by state assessments and teacher administrators. All teachers demonstrated instructional practices that were a balanced literacy approach. The credibility of this study is strong due to other empirical studies that indicate similar results. Unfortunately the confirmability of the

study was weak because of the failure to record the measures the researchers took to indicate that the observations were checked by many different people, besides just the observers. If these observations and findings had been inspected by an outside party, then the confirmability of this study would be strong.

The analysis of these seven studies suggests that students benefit from a balanced approach to literacy instruction which takes aspects of whole language instruction and phonics/skills-based instruction to make a more balanced approach to literacy. Ketner et al. (1998) and Freppon (1994) found that whole language instruction within a classroom and administered one on one during tutoring sessions fostered academic achievement in reading and a positive attitude towards reading. Grisham (1993) and Traw (1996) found no statistically significant difference between the academic performance of whole language classes and the traditional skills/phonics-based classrooms. Sears (1999) found that, even though students were taught within a whole language classroom, they were still learning important reading cues including grapho-phonics. This indicated that explicit phonics instruction was not needed to help students learn within the grapho-phonics cuing system. Song and Miskel (2002) found that national reading policymakers and organizations favored a balanced approach over whole language or phonics-based approaches. Morrow et al. (1999) found that teachers within the study all preferred and used a balanced approach to teaching reading and saw student achievement.



## **Whole Language and English Language Learners and English as a Foreign Language Learners**

The studies analyzed so far have investigated the effectiveness of whole language instruction compared to classrooms that focus on phonics-based instruction, skills-based instruction (basal readers and direct instruction), and achievement in primary education. Studies previously surveyed also have looked at the opinions teachers and policy makers have about which instructional method has shown to be more effective for their students. The three studies reviewed in this section focus on the effects whole language instruction has on the reading or comprehension achievement of English Language Learners (ELL) or English as a Foreign Language (EFL) students. The first study, by Lim and Watson 1993, sought to determine whether a whole language environment could help ELL students show improvement in reading and language acquisition. The second study, by Kucer and Silva (1999), examined whether or not whole language instruction that was administered to ELL primary students could help them achieve proficiency in reading. Last, a study by El-Koumy (2000) focused on whether or not whole language instruction is effective for EFL college students.

Lim and Watson (1993) studied the effects of whole language instruction on ELL students' achievement. Lim and Watson conducted a qualitative study on a classroom in Columbia, Missouri, which had 10 second- and third-grade students who had been in the United States for less than two years, using a semi-experimental method. The authors found that combining “authentic and natural language experiences with content rich classroom practices leads to optimal subject matter learning” and achievement in student reading (Lim and Watson, 1993, p. 393).

An ESL classroom in Columbia, Missouri was chosen for this study, which was conducted during the summer quarter of school. This study included participants who were second- and third-grade students who had been in the United States for less than two years. Five students' native language was Cambodian, one was Korean, two were Chinese, one was Taiwanese, and one was Arabic. The teacher who taught this class, Betty Belcher, primarily focused on and used whole language philosophy and instruction.

The methodology utilized in this study was qualitative and included close observation by the researchers. Lim and Watson (1993) collected physical data such as observations of classroom activities, worksheets, and answers to problems. With this data, researchers analyzed the effects whole language practice was having on reading and comprehension. The first hour of class consisted of using language to make sense of oral and written texts. The teacher engaged the students in conversation which upheld an interpersonal relationship with the students. The teacher allowed the learners to take the role of active participants in their own learning by initiating conversation instead of simply having students respond to questions initiated by the teacher. The students also had the opportunity to make meaningful experiences in reading through using authentic texts. An example of this is when the teacher was reading the students a book about bats. This book was chosen because a majority of the students showed an interest in bats. This interest in bats helped the students stay engaged in the text. The students were also encouraged to engage in conversation about the book while they read it together.

The results revealed that talk as a means of learning helped students experience learning English in a meaningful way. Since the students were able to experience language learning through a holistic whole language means, they could explore the

meaning of expression and develop their language themselves and engage in it through interesting themes.

This study demonstrated various strengths. This qualitative study is transferable because of the detailed description of the students, teacher, and geographical area. The results are dependable because they are consistent with other results within the academic community. The study is credible because the method and description is clear and well carried out by the researchers. It also appears that the researchers really caught the experiences of the participants.

The next study is similar to Lim and Watson's because it focused on the effects of whole language on ELL students. However, Kucer and Silva (1999) conducted a quantitative study that followed a group of Spanish-speaking students through their journey of learning a new language (English). This study strived to answer the question of whether or not whole language instruction could help students who are transitioning from Spanish literacy to English literacy could help those students demonstrate achievement in comprehension. Kucer and Silva found that out of the four categories assessed (reading, retelling, writing, and spelling); reading and retelling demonstrated the most growth in comprehension during reading as evidenced by miscue analysis.

Kucer and Silva (1999) conducted a yearlong study on a third-grade classroom that was transitioning from Spanish literacy to English literacy. The subjects in this study included 26 third-grade Mexican American students. These subjects were from working class homes and had been chosen for a transitional literacy program by means of their prior work with their second-grade teacher. The teacher chosen for this study was a

Colombia native who had 11 years' experience teaching bilingual elementary classes. This teacher also was attaining her doctorate "in a whole language-oriented language, literacy, and culture program" and her teaching philosophy and instructional method reflected that of whole language instruction (Kucer & Silva, 1999, p. 12). For this span of time, a researcher joined the teacher and the 26 students to be a participant-observer in the classroom. The researcher spent three days a week in the classroom.

Kucer and Silva (1999) studied the four main components that the teacher implemented into her literacy block. These four main components consisted of, "theme-based literacy activities, teacher reading, free reading, and free writing" (Kucer & Silva, 1999, p. 14). The classroom literacy block with four main components allowed the students to engage in regular activities such as, "pairing reading, reader response groups, compare/contrast activities, expert groups, learning logs, writing conferences, modified cloze procedures, and strategy wall charts" (Kucer & Silva, 1999, p. 15). These classroom activities allowed students to engage in listening to reading, reading themselves, analyzing the books, and writing stories themselves.

In order to gauge progress throughout the study, Kucer and Silva (1999) maintained a detailed field notes log and also engaged the students in pre-assessments and post-assessments in the four literacy areas. The four areas that Kucer and Silva assessed the students were, "readings, retellings, writings, and spelling" (Kucer & Silva, 1999, p. 24). The researchers used miscue analysis in order to score the sections on reading and retellings. Writing and spelling were scored according to a detailed holistic rubric that included a four-point scale with detailed description of what each score entailed. To analyze the mean scores produced by the pre-assessment and post-

assessment in all areas, Kucer and Silva used the Friedman Two-Way ANOVA Analysis by Ranks, which is a tool for calculating quantitative measures (p. 24).

Kucer and Silva (1999) found that there was a positive significant change in reading and retelling miscue analysis and spelling but no significant difference in writing. The reading and retelling miscue analysis showed a statistically significant change from fall to spring. The mean score in fall was 38.90 and in the spring the mean score jumped to 52.72. In the analysis it showed an increase in “syntactically and semantically acceptable [sentences] within context” from their pre-assessment to their post-assessment (Kucer & Silva, 1999, p. 26). In the writing assessment, the mean score in fall was 1.30 and jumped to 1.75, which does not warrant a statistically significant change in the students’ writing ability. Kucer and Silva attributed the low achievement in the writing assessment to a lack of explicit instruction in writing. Kucer and Silva claimed that explicit instruction may have helped the students who needed additional support in areas such as “idea development, vocabulary, or syntax” (p. 28).

This study was particularly strong because of its validity, but still had limitations. This study demonstrated strength in describing exactly what happened in the study. The researchers thoroughly described the subjects and the manner that the assessment were created and scored. A limitation of this study was not including a control group of students. This study included a student sample that received the treatment, but there was not a control classroom that the researchers could compare their findings to. If a control group had been included, this would have been an exemplary study, but without this group, there is room for further study and speculation.

El-Koumy's (2000) study is similar to the previous studies due to its research within the whole language realm; however, this researcher sought to compare the effects of whole language instruction and skills-based instruction on English as a Foreign Language (EFL) students' listening comprehension skills. El-Koumy conducted a quantitative study on 96 pretested EFL students, divided into two treatment groups (skills-based and whole language based) for 15 weeks, using a semi-experimental method. El-Koumy found that there was no significant difference between the two treatment groups, only that high ability groups in both treatments did significantly better on post-assessment than the low ability groups.

The participants in the study were EFL students who had been learning English as a foreign language for 9-12 years. These students were enrolled in an EFL program at Suez Canal University in Suez, Egypt. The students' ages ranged from 19-22 and the students "who scored 30 and above were designated as high ability listeners; those who scored 15 or below were designated as low ability listeners; and those who scored between 15 and 30 were excluded from taking part in the experiment" (El-Koumy, 2000, p. 5). In November 1999, the participants were given a placement test to separate students into various listening ability groups. On the test there were 50 multiple-choice questions that tested "discrete listening sub-skills such as recognition of individual sounds, reduced forms, stress and intonation patterns, based on short spoken texts" (El-Koumy, 2000, p. 6). A listening comprehension posttest, which had been confirmed by five university teachers, was administered following the study, which ended in April 2000. Whole language instructional methods (WL) were compared to skills-based instruction (SB) in

implementation and teaching of strategies. A t-test was constructed to determine significance at the .05 level of confidence.

The purpose of this study was to see how whole language approaches affected listening comprehension in EFL students. The participants were separated into four different groups: one whole language low group, one skills-based low group, one whole language high group, and one skills-based high group. The results showed achievement growth in both of the high groups with two separate treatments, with no significant differences between the achievement of the two groups (means WL= 39.13, SB= 36.83). The low group with two different treatments also showed no difference compared to one another's achievement (means WL= 16.50, SB= 17.42). However, the difference between the low groups and the high groups within each treatment showed significant difference in achievement. These results indicate that the high-achieving groups will keep achieving while the low-achieving groups (no matter the treatment) will still do poorly.

This study has various weaknesses and strengths. This study shows a weakness in internal validity. El-Koumy (2000) conducted a study that sought to determine whether whole language and skills-based instruction would have a significant difference in performance and whether the instructional method helped students in the low or high group more. There is a flaw in the method where the researcher changes the participants after administering a pretest. He rejected the students who performed in the middle. This type of design of research would render results that were predictable. Therefore, the students who performed in the lower group would still perform low and the high-performing students would still perform high. An apparent weakness of the study was the researcher's failure to state, in detail, the ways that the instructional methods were

executed within the classroom. The external validity is weak due to the lack of demographics expressed in the study. The objectivity of the study is strong because it could be corroborated by others within the educational community.

Lim and Watson (1993) and Kucer and Silva (1999) both strongly suggest that whole language is effective in helping ELL students achieve proficiency or readiness in reading. The study by El-Koumy (2000) is weak, but it suggests that whole language has no significant effect on helping EFL students learn a new language. Overall, the findings suggest that whole language is successful in helping students attain achievement in learning and becoming literate in English.

### **Summary**

In Chapter 2, the articles reviewed focused on whether or not whole language philosophy, implementation, and practice affect student achievement in reading proficiency. Chapter 2 was divided into four sections; (1) whole language and phonics-based instruction, (2) whole language and skills-based instruction, (3) whole language and primary academic achievement, and (4) whole language instruction and ELL and EFL students. Chapter 3 provides a summary of findings in the sections listed above. Chapter 3 will also consider what the classroom implications are and give suggestions for future studies.



## CHAPTER 3: CONCLUSION

### **Introduction**

Chapter one of this review described the rationale, historical context, definitions, and limitations of this paper and presented a statement of purpose. The main purpose of this paper is to determine whether or not whole language instruction is an effective philosophy and instructional method to help students obtain achievement in reading. Chapter two was a review of 30 studies which were related to whole language instruction and its effects on students' achievement in reading. Chapter three offers a summary of research findings, classroom implications, suggestions for further research, and conclusions to this body of research.

Chapter one offered a rationale that indicated that whole language is a philosophy and instructional method that is controversial, often questioned in its effectiveness towards proficiency in reading. The historical context offered background information on how whole language philosophy and instruction began and how it has become a popular and controversial instructional method. The Limitations section described the scope of this paper, which is limited to whole language philosophy and practice within elementary reading classes. Educational scholars and practicing educators have argued whether students realize a greater benefit from whole language, phonics-based, or skills-based instruction. The purpose of this paper was to determine whether whole language philosophy and instructional methods help students achieve proficiency in reading.

Chapter two reviewed 30 studies which were related to whole language instruction and its effects on students' achievement within reading. Chapter two was

divided into four sections: (1) whole language and phonics-based instruction, (2) whole language and skills-based instruction, (3) whole language and primary academic achievement, and (4) whole language instruction and English Language Learners (ELL) and English as a Foreign Language (EFL) learners.

The first section compared research to determine whether whole language instruction or phonics-based instruction was more beneficial for student achievement in reading. This comparison revealed that implicit phonics instruction within a whole language or holistic program is more beneficial for student achievement in reading.

The second section examined whether whole language instruction or skills-based instruction was more effective in improving students' achievement in reading. Overall, in quantitative measures whole language and skills-based instruction proved equally effective in helping students in reading achievement. However, the qualitative measures of many studies indicated that the whole language classes offered additional benefits.

The third section analyzed the effect whole language instruction had on primary grade students' reading achievement. The analysis found that whole language instruction had a positive effect on student achievement in reading.

The fourth section analyzed the effects whole language had on ELL or EFL students. This section found that whole language instruction was effective in helping develop students' reading skills.

Chapter three includes a summary of findings including an analysis of the strengths, weaknesses, and overall quality of the studies presented in chapter two. Chapter three also includes classroom implications, suggestions for further research, and concluding thoughts.

### **Summary of Findings**

Chapter two analyzed 30 studies that were separated into four sections: (1) whole language and phonics-based instruction, (2) whole language and skills-based instruction, (3) whole language and primary academic achievement, and (4) whole language instruction and English Language Learners and English as a Foreign Language learners. Significant findings from these studies are summarized below.

#### **Whole Language and Phonics-Based Instruction**

The seven studies analyzed for this section investigated whether whole language instruction or phonics-based instruction was a more effective philosophy and instructional method for helping students achieve reading readiness. The findings suggest that whole language instructional methods are beneficial to student achievement in reading; however, two strong studies indicated that a balanced approach to literacy is helpful in teaching students how to read (Baumann et al., 1998; Dahl and Scharer, 2000). Dahl and Scharer found that students more rapidly progressed in reading levels when provided whole language instruction along with the implementation of implicit phonics instruction. Baumann et al. conducted a strong study which determined elementary school teachers' attitudes towards the two instructional methods. The majority of teachers surveyed stated that a balanced approach to literacy is imperative to students' success in reading. A moderately strong study by McIntyre (1995) demonstrated the instructional methods and ideas of teachers on a smaller scale. This researcher found that three teachers, who had been whole language advocates for years, changed their mind about literacy instruction for primary age students and came to believe that a balanced approach to literacy was more effective in helping students achieve reading readiness. Three weak

studies indicated that phonics instruction was more effective in helping students achieve reading readiness (Dakin, 1999; Donat, 2006; Ryder et al., 2008). All studies indicated that explicit phonics instruction was more effective in improving student achievement in reading, yet the measures ended with phonics testing, not tests on comprehension; in so doing researchers did not examine the main reason one learns to read. The last study, by Griffith et al. (1992), was weak and indicated no significant difference between instructional models; however, there was a difference between the low and high groups. The high group showed higher levels of achievement than the low group.

This section included three studies that were quantitative, two that were qualitative with some quantitative measures, and two that were qualitative. The quantitative measures included pre-assessment and post-assessment work. The qualitative studies that included quantitative measures also included pre- and post-surveys. The qualitative measures administered were primarily observations, informal interviews, and field notes. The researchers' data included detailed descriptions of treatment method, implementation, and delivery, and how students reacted.

### **Whole Language and Skills-Based Instruction**

This section analyzed 13 studies that compared whole language and skills-based instruction, to determine which was more effective in helping students achieve reading readiness. The overall findings of this section indicated that both instructional methods are effective in helping students achieve in reading readiness or competence. However, the whole language treatments also fostered an increased positive attitude in students towards reading. One moderately strong study indicated, through quantitative measures, that skills-based instruction helped students excel beyond the students in the whole

language classroom (Stahl et al., 1996). However, qualitative measures indicated that whole language students' exhibited attitudes towards reading were more positive than those of the students who surpassed them. The post-assessment from three strong studies indicated no significant difference in reading improvement between the skills-based and the whole language classrooms (Dahl and Freppon, 1994; Davis, 2010; Freppon and McIntyre, 1999).

All studies came to the same conclusion through qualitative measures: students within the whole language classrooms had a more positive attitude towards reading than the skills-based students. Three strong studies (Hurry et al., 1999; Pinnel et al., 1994; Reutzel and Cooter, 1990) and two weaker studies (Stockard & Engelmann, 2010; Varble, 1990) found that student reading post-assessment performances were higher in whole language classrooms than in skills-based classrooms. These whole-language classrooms also indicated the same positive attitude towards reading through qualitative measures. Costello (2012) conducted a study as a researcher practitioner. He found that no single prescribed program could meet the needs of every student; he advocated for a balanced approach.

The studies in this section are a mix of quantitative and qualitative. Most of the studies were comparison studies that included both quantitative and qualitative measures to ensure precise results. The study designs included pre-assessment, post-assessment, and some included mid-assessments on many measures. The studies were administered to students within kindergarten through third grade. The qualitative studies focused on smaller groups of participants and confirmed the results indicated by the quantitative studies. This body of research strongly indicates that whole language is a philosophy and

instructional method that helps students achieve reading readiness and the motivation and attitude to persevere over challenges in reading.

### **Whole Language and Primary Academic Achievement**

Of the seven studies reviewed in this section, five were exceptionally strong studies. The following are the findings of the strong studies as they relate to whole language and academic achievement. Freppon (1994) indicated that one-on-one lessons with a student using whole language instruction could help him become a reader with a positive relationship with literacy. Grisham (1993) indicated no statistical difference between instructional methods. Sears (1999) indicated that students in a whole language classroom progressed in comprehension and phonics skills in reading without any explicit instruction in phonics. Morrow et al. (1999) indicated that teachers preferred a balanced approach to literacy. Ketner et al. (1997) indicated that a balanced approach to literacy is beneficial to student academic progress in reading. Overall, the strong findings in these studies indicate that students within primary education must be taught with an instructional method that is similar to whole language.

Three of the studies in this section are qualitative (Freppon, 1994; Morrow et al., 1999; Sears, 1999). The measures taken in each study involve surveys, interviews, observations, and field notes. Four of the studies within this section were quantitative (Grisham, 1993; Ketner et al., 1998; Song & Miskel, 2002; Traw, 1996). Two of the quantitative studies took measures that included pre-assessment and post-assessment in two treatment groups. One of the quantitative studies included national surveys on instructional methods and teachers experiences with those methods (Song & Miskel, 2002). Two of the quantitative studies (Song & Miskel, 2002; Traw, 1996) were

relatively weak in design but their findings were still important to the body of research. Both studies were quantitative in design; however, their internal and external validity were weak.

### **Whole Language Instruction and English Language Learners and English as a Foreign Language Learners**

This section included three studies that analyzed whether whole language instruction was effective in helping ELL or EFL students in academic achievement in reading. One strong qualitative study suggested that whole language instruction benefitted students in achieving reading readiness in their second language (Lim & Watson, 1993). Another moderately strong study found the same results through quantitative measures (Kucer & Silva, 1999). A moderately weak quantitative study (El-Koumy, 2000) found that whole language instruction did not show significant test results compared to other instructional methods.

The qualitative study by Lim and Watson utilized methods that included observation and field notes. Additionally, the researchers used classroom experiences and worksheets to provide evidence of achievement in reading. The quantitative study by Kucer and Silva included measures such as pretests, posttests, and miscue analysis. This study also included some qualitative measures, including observations, but based findings on the quantitative measures. While the findings in Kucer and Silva suggested academic success for ELL students in reading, qualitative measures made by Lim and Watson confirmed any quantitative findings made by Kucer and Silva. El-Koumy's design contained weak internal validity that led to a weak result .

The summary of findings evaluated and exposed the overlying patterns that emerged from the body of research. The following Classroom Implications section will evaluate the significance of these findings and will explore which instructional methods and approaches are most likely to support students' academic achievement in reading.

### **Classroom Implications**

The benefit of whole language instruction, as it relates to students' academic achievements in reading was somewhat supported by the body of research. In stating this, the body of research suggested that whole language is just as supportive as other instructional methods, and in some quantitative studies, whole language was proven to be more effective than other instructional methods (e.g. Freppon, 1994; Pinnel et. al., 1994; Reutzel & Cooter, 1990; Sears, 1990). While the classrooms studied varied in methods used, the support was rooted in whole language instruction. These studies found students' reading improved through receiving whole language instruction and that this improvement was reflected in both quantitative and qualitative measures.

Another implication for classroom practice that was strongly supported by many studies was a balanced approach to literacy instruction. This strong trend was supported by comparison studies. Even though the studies revealed no statistical differences between the two treatment methods (whole language and skills/phonics-based instruction), qualitative measures (observation, interviews, or surveys) indicated that students within the whole language classrooms had a more positive attitude towards reading compared to students within the other treatment groups (e.g. Dahl & Freppon, 1994; Davis, 2010; Freppon & McIntyre, 1999; Grisham, 1993). Of the studies that found (through quantitative measures) that phonics/skills-based instruction was significantly



better than whole language, one of the studies indicated that the whole language group had a better attitude towards reading than the other treatment group (Stahl et al., 1996).

An implication for classroom practice addressed in some studies was that of teachers' opinions and experiences with whole language instruction, phonics/skills-based instruction, and balanced instruction. Surveys focused on teachers' opinions (Baumann et al., 1998; Ketner et al., 1998; Song & Miskel, 2002) found that the majority of respondents preferred, and observed higher achievement with, a balanced approach to literacy instruction.

Other case studies found that a balanced approach was useful in supporting students' academic achievement in reading (Dahl & Scharer, 2000; McIntyre, 1995; Morrow et al., 1999). Newer studies showed a trend toward a balanced approach to literacy. It would be advantageous for any literacy teacher in primary school to adopt the whole language philosophy, even as a part of a balanced approach to literacy that includes explicit or implicit phonics instruction. This idea was also supported by Costello's qualitative study, in which he observed his own classroom (Costello, 2012).

Research has found the implementation of the whole language philosophy and instructional method to be just as effective as, or more effective than, other instructional methods in supporting academic achievement in reading. Further, studies have shown that the whole language philosophy fosters a positive attitude towards reading and helps students develop reading strategies that improve reading comprehension (Dahl and Freppon, 1994; Davis, 2010; Freppon & McIntyre, 1999; Varble, 2004). Other studies indicated that explicit or implicit phonics or skills-based instruction was important to administer while engaging in literacy instruction (Dahl & Scharer, 2000; McIntyre, 1995;

Morrow et al., 1999). The varied results found in the body of research indicate that no single method works best for all students. Therefore, teachers should employ literacy methods that will provide the greatest benefit the individual students.

### **Suggestions for Further Research**

The body of research presented contains weaknesses which indicate important areas for further research that could provide a more complete body of evidence regarding whole language instruction's effect on reading readiness.

Some studies included here support explicit phonics and skills instruction in the classroom (Dakin, 1999; Donat, 2006; Griffith et al., 1992; Ryder et al., 2008). These studies employed quantitative measures and determined that phonics/skills-based instruction was more effective in supporting students in reading. The researchers arrived at these conclusions because of the limited scope of the studies used. All of these studies relied exclusively on tests measuring skills in phonics. The failure of these studies to broaden their scope to include other important reading cues produced an incomplete representation of how people learn to read. The inclusion of other measures relating to comprehension may have produced different findings. The measure of comprehension should be included in future studies to ensure that a more comprehensive view of student achievement in reading is accounted for in the corpus.

A few studies examined the beliefs of teachers and administrators regarding which instructional method was more effective in supporting student reading (Baumann et al., 1998; Ketner et al., 1998; Song & Miskel, 2002). These studies found that a balanced approach to literacy was preferred by teachers and administrators, but these studies neglected to gather classroom data such as demographics, teacher procedures, and student

progress. These studies are flawed because they rely on opinions. These opinions were not supported by empirical data on student achievement in reading. To strengthen the results of future studies, researchers should obtain classroom data to support the opinions and experiences provided by the survey respondents.

These studies were also flawed because there were many variables not considered in the study design. The researchers did not account for differences between teachers' claims of classroom practices and actual classroom practice (practices that could be observed). It is possible that a teacher who answered on the whole language side really did not practice whole language. Therefore, the validity of future studies would be improved by the inclusion of the researchers' direct observation of classroom activity.

Studies whose design included pre/post-assessments which split students into low and high groups of achievement in a particular area contain a design flaw. If students are separated into groups that represent high-achieving students and low-achieving students, the results are likely to represent the high group achieving higher scores than the low-achieving group.

When conducting any study, the reliability of the study is stronger when a control group is utilized. In a few studies researchers did not utilize a control group to which they could compare their findings. For example, Costello (2012) conducted a study where he was the researcher/practitioner. He only administered his treatment on one group and did not compare his findings to a control group. Inclusion of a control group is a standard research practice, and including such measures in future studies is likely to produce more reliable results.

Communication is an essential human need. Acquisition of this skill strongly correlates to academic success in school and financial success in society. This body of research did not include studies that examined the long-term effects of whole language instruction on student achievement. Such studies could provide insight into whole language instruction's effect on and support for students in their long-term reading achievement. The academic community would benefit from long-term research studies that examine whether students who received whole language instruction achieved greater academic success and a higher completion rate than those who did not receive whole language instruction.

Further research focused on whole language instruction of students from a low socioeconomic group and their achievement in reading could provide teachers with greater insight into teaching methods that benefit this segment of society. This body of research included two studies that focused on whole language instruction within a low socioeconomic demographic. Results indicated that whole language instruction with implicit phonics instruction was beneficial to student achievement in reading proficiency; however, the majority of these studies were conducted on upper class and middle class students. Further research to determine whether whole language philosophy and practice are effective within classrooms that have students from the low socioeconomic demographic is recommended.

Lastly, the majority of the research was conducted on European American students. Further research should focus on the effectiveness of whole language philosophy and instruction when working with students from more varied cultural backgrounds. Further research to determine whether whole language philosophy and

practice are effective within classrooms that have students from varied cultural backgrounds is recommended.

### **Conclusion**

Chapter one introduced the main point of this paper: whether whole language philosophy and instruction is supportive in students' academic achievement in reading. The rationale and historical context was provided as well. The historical context described the formation, implementation, and current role whole language plays within the educational world today. Chapter two critically reviewed 30 articles that examined whole language in relation to academic achievement. Chapter two was divided into four sections: (1) whole language and phonics-based instruction, (2) whole language and skills-based instruction, (3) whole language and primary academic achievement, and (4) whole language instruction and English Language Learners and English as a Foreign Language. Chapter three comprised a summary of findings, classroom implications, and suggestions for further research.

Whole language philosophy, implementation, and practice was proven by this body of research to be more effective than, or just as effective as, other instructional methods in helping students achieve reading proficiency. However, the body of research also suggests that a balanced approach to literacy is beneficial to students' achievement in reading. Whole language philosophy is focused on the uninterrupted reading process which allows students to gain meaning while reading (Gilles, 1988). This body of research suggests that whole language can be carried out within the classroom setting along with phonics and skills-based instruction. These skills do not need to interrupt the reading process, but can be incorporated into an authentic reading and writing experience.

The whole language and balanced literacy approach can be utilized by teachers to guide their important work of teaching students how to read.

## REFERENCES

- Anderson, G. S. (1984). *A whole language approach to reading*. Lanham, MD: University Press of America.
- Baumann, J. F., Hoffman, J. V. Moon, J., & Duffy-Hester, A. M. (1998). Where are teachers' voices in the phonics/whole language debate? Results from a survey of U.S. elementary classroom teachers. *Reading Teacher*, 51(8), 636-650.
- Costello, D. (2012). The impact of a school's literacy program on a primary classroom. *Canadian Journal of Education*, 35(1), 69-81.
- Dahl, K. L., & Freppon, P. A. (1994). *A comparison of inner-city children's interpretations of reading and writing instruction in the early grades in skills-based and whole language classrooms*. Washington, DC: Office of Educational Research and Improvement.
- Dahl, K. L., & Scharer, P. L. (2000). Phonics teaching and learning in whole language classrooms: New evidence from research. *Reading Teacher*, 53(7), 584-594.
- Dakin, A. B. (1999). *The effectiveness of a skill based explicit phonics reading program K-2 as measured by student performance and teacher evaluation*. (Unpublished master's thesis) Dominican College, San Rafael, California.
- Davis, L. (2010). Toward a lifetime of literacy: The effect of student-centered and skills-based reading instruction on the experiences of children. *Literacy Teaching and Learning*, 15(1-2), 53-79.
- Dewey, J. (1986). Experience and education. *Educational Forum*, 50(3), 241-52.
- Donat, D. J., (2006). Reading their way: A balanced approach that increases achievement. *Reading & Writing Quarterly*, 22(4), 305-323.
- El-Koumy, A.S. (2000). *Effects of skills-based versus whole language approach on the comprehension of EFL students with low and high listening ability levels*. Retrieved from <http://www.eric.ed.gov/PDFS/ED449670.pdf>
- Freppon, P. A. (1994) Understanding the nature of reading and writing difficulties; An alternative view. *Reading & Writing Quarterly*. 10(3), 227-238.
- Freppon, P. A., & McIntyre, E. (1999). A comparison of young children learning to read in different instructional settings. *Journal of Educational Research*, 92(4), 206-218.

- Gilles, C., Bixby, M., Crowley, P., Crenshaw, S. R., Henrichs, M., Reynolds F. E., & Pyle D. (Eds.). (1988). *Whole language strategies for secondary students*. New York, NY: Richard C. Owen.
- Goodman, K. S. (1989). Whole-language research: Foundations and development. *The Elementary School Journal*, 90(2), 207-221.
- Goodman, Y. M. (1989). Roots of the whole-language movement. *The Elementary School Journal*, 90(2), 113-127.
- Goodman, K. S., & Goodman, Y. M., (1981). *A whole-language, comprehension-centered reading program. Program in Language and Literacy Occasional Paper Number 1*. Tucson, AZ: Arizona University College of Education.
- Griffith, P. L., Klesius, J. P., & Kromrey, J. D. (1992). The effect of phonemic awareness on the literacy development of first grade children in a traditional or a whole language classroom. *Journal of Research in Childhood Education*, 6(2), 85-92.
- Grisham, D. L. (1993). *The integrated language arts: Curriculum enactments in whole language and traditional fourth grade classrooms*. (Doctoral dissertation). Retrieved from Ebscohost (ED359495)
- Hollingsworth, P. M. , Reutzel, D., & Weeks, E. (1990). Whole language practices in first grade reading instruction. *Reading Research and Instruction*, 29(3), 14-26.
- Hurry, J., Sylva, K., & Riley, J. (1999). Evaluation of a focused literacy teaching programme in reception and year 1 classes: Child outcomes. *British Educational Research Journal*, 25(5), 637-649.
- Kamberelis, G. (1998). Relationships between children's literacy diets and genre development: You write what you read. *Literacy Teaching and Learning*, 3, 7-53.
- Ketner, C. S., Smith, K. E., & Parnell, M. K. (1997). Relationship between teacher theoretical orientation to reading and endorsement of developmentally appropriate practice. *Journal of Educational Research*, 90(4), 212-20.
- Kucer, S. B., & Silva, C. (1999). The English literacy development of bilingual students within a transitional whole language curriculum. *Bilingual Research Journal*, 23(4), 7-53.
- Lim, H., & Watson, D. J. (1993). Whole language content classes for second-language learners. *Reading Teacher*, 46(5), 384-93.



- McIntyre, E. (1995). The struggle for developmentally appropriate literacy instruction. *Journal of Research in Childhood Education, 9*(2), 145-156.
- Miller, P.H. (2009). *Theories of developmental psychology*. New York, NY: Worth Publishers.
- Morrow, L. M., Tracey, D. H., Woo, D. G., & Pressley, M. (1999). Characteristics of exemplary first-grade literacy instruction. *The Reading Teacher, 52*(5), 462-476.
- Pinnel, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly, 29*(1), 8-39.
- Pressley, M., Wharton-McDonald, R., Allington, R., Block, C., & Morrow, L. (1998). *The nature of effective first-grade literacy instruction* (CELA Research Report No. 11007). Retrieved from <http://cela.albany.edu/1stgradelit/index.html>
- Puorro, M., (1997). *The whole language approach versus the basal reading approach and the effects on reading achievement scores*. (Master's thesis). Retrieved from EBSCOhost. (ED405551.)
- Raines, S. (1995). *Whole language across the curriculum*. New York: Teachers College Press.
- Reutzel, D., & Cooter, R. B. (1990). Whole language: Comparative effects on first-grade reading achievement. *Journal of Educational Research, 83*(5), 252-57.
- Rosenblatt, L. M. (1988). *Writing and reading: The transactional theory* (Technical Report No. 416). Retrieved from University of Illinois at Urbana, Center for the Study of Reading website: [www.ideals.illinois.edu/bitstream/.../ctrstreadtechrepv01988i00416\\_opt.pdf](http://www.ideals.illinois.edu/bitstream/.../ctrstreadtechrepv01988i00416_opt.pdf)
- Ryder, J. F., Tunmer, W. E., & Greaney, K. T. (2008). Explicit instruction in phonemic awareness and phonemically based decoding skills as an intervention strategy for struggling readers in whole language classrooms. *Reading and Writing: An Interdisciplinary Journal, 21*(4), 349-369.
- Sears, S. (1999). The development of reading strategies in a whole language classroom. *Reading Psychology, 20*(2), 91-105.
- Song, M., & Miskel, C. (2002). *Interest groups in national reading policy: Perceived influence and beliefs on teaching reading* (CIERA Report No. 3-023). Retrieved

from University of Michigan School of Education, Center for the Improvement of Early Reading website: <http://www.ciera.org/library/reports/inquiry-3/3-022/3-022a.pdf>

- Stahl, S. A., Pagnucco, J. R., & Suttle, C. W. (1996). First graders' reading and writing instruction in traditional and process-oriented classes. *Journal of Educational Research*, 89(3), 131-44.
- Stockard, J., & Engelmann, K. (2010). The development of early academic success: The impact of direct instruction's "reading mastery". *Journal of Behavior Assessment And Intervention In Children*, 1(1), 2-24.
- Traw, R. (1996). Large-scale assessment of skills in a whole language curriculum: Two districts' experiences. *Journal of Educational Research*, 89(6), 323-339.
- Varble, M. (1990). Analysis of writing samples of students taught by teachers using whole language and traditional approaches. *Journal of Educational Research*, 83(5), 245-251.
- Vygotsky, L. S. (1978) *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Weaver, C. (2009). *Reading process: Brief edition of reading process & practice* (3<sup>rd</sup> ed.). Portsmouth, NH: Heinemann.