DEVELOPING A PROFILE TO PREDICT STUDENT RESPONSE TO TREATMENT

FOR FAST FORWORD PROGRAMS

by

LAURIE ANN GLAZENER

A DISSERTATION

Presented to the Department of Educational Methodology,
Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Education

December 2010

DISSERTATION APPROVAL PAGE

Student: Laurie Ann Glazener

Title: Developing a Profile to Predict Student Response to Treatment for Fast ForWord Programs

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Education degree in the Department of Educational Methodology, Policy, and Leadership by:

Dr. Paul Yovanoff   Chair
Dr. Keith Hollenbeck   Member
Dr. Joseph Stevens   Member
Dr. Jeffery Sprague   Outside Member

and

Richard Linton   Vice President for Research and Graduate Studies/Dean of
                           the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2010

DISSERTATION ABSTRACT

Laurie Ann Glazener

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

December 2010

Title: Developing a Profile to Predict Student Response to Treatment for Fast ForWord
Programs

Approved: _____
Dr. Paul Yovanoff

Matching reading deficits to appropriate intervention programs is a challenge

given the number of treatment options available to educators. The Fast ForWord (FFW)

computerized intervention series has been marketed as a way to treat underlying causes of

poor reading skill; i.e., substandard levels of basic language skill, phonemic awareness,

and/or phonics application. If the programs work as claimed, then completion of Fast

ForWord should improve the next reading subskill developed after phonics, oral reading

fluency. Part 1 of this study involves a treatment ($n = 72$) versus comparison ($n = 84$)

group two by two ANOVA to evaluate that hypothesis. No effect for FFW is found

($p = .84$). Application of decision rules from Response to Intervention (RTI) models

classifies positive changes in risk category at a greater rate for the comparison group

($n = 31$) than for the FFW group ($n = 20$) ($X^2 = 3.81$, (1), $p = .05$). Pre-intervention

language scores for the FFW group are compared to assist with intervention placement

decisions. Differences in mean language scores are not significant ($p = .85$) between the

two groups [positive response ($n = 19$) versus low response ($n = 57$)]. In a binary logistic regression of quartile membership for language scores, no score ranges predict membership ($X^2 = 4.75$, (8), $p > .05$). Measuring treatment effect with ORF is not recommended. The use of pre-intervention language and ORF scores below the 25th percentile as indicators of a positive change in oral reading fluency following FFW treatment also is not recommended. However, future research that considers language scores along with other curriculum-based measures of prereading skill either as pre-intervention indicators or outcome measures is recommended.

CURRICULUM VITAE

NAME OF AUTHOR: Laurie Ann Glazener

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene

DEGREES AWARDED:

Doctor of Education, Educational Methodology, Policy, and Leadership, 2010,
    University of Oregon
Master of Science, Curriculum and Instruction, 1990, University of Oregon
Bachelor of Science, Elementary Education, 1987, University of Oregon

AREAS OF SPECIAL INTEREST:

Change Leadership
Three-Tiered Instructional Models

PROFESSIONAL EXPERIENCE:

Principal of Lacomb School, Lebanon Community School District, Lebanon,
    Oregon, 2010-present

First Grade Teacher, Lebanon Community School District, Lebanon, Oregon,
    2009-2010

Administrator of Special Programs, Lebanon School District, Lebanon, Oregon,
    2007-2009

Teacher on Special Assignment, Springfield School District, Springfield, Oregon,
    2003-2007

Elementary School Reading Specialist, Springfield School District, Springfield,
    Oregon, 1992-2003

Primary Teacher (Kindergarten-2nd Grade), Springfield School District, Springfield, Oregon, 1989-1992

Special Education Teacher, Sweet Home School District, Sweet Home, Oregon, 1988-1989

PUBLICATIONS:

Glazener, L. (2004). *Sensorcises: Active enrichment for the out-of-step learner*. Thousand Oaks, CA: Corwin Press.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Identifying the Problem

Becoming a fluent reader is central to achieving academic success in our schools today. While average reading scores from the 2009 National Assessment of Educational Progress (NAEP) were the highest on record, a large proportion of those assessed were still not proficient in reading by NAEP standards *(*National Center for Education Statistics [NCES], 2009). The percentage of students operating at or above the proficient level was 31% for fourth grade (up from 30% in 2005) and 30% for eighth grade (up from 29% in 2005). Alternately, these same statistics indicate that 69% of fourth-grade and 70% of eighth-grade students were not considered proficient readers. These NAEP results create cause for concern for school systems because, without considerable intervention, students failing to become proficient readers by the time they have completed elementary school have only a 20% chance of becoming proficient readers as adults (Shaywitz et al., 1999).

This is not a recent trend. The percentage of low-literate students in the nation's schools has remained relatively constant over the past decade (The Education Trust, 2009). This trend continues despite substantial increases in funding for providing interventions (Bali & Alvarez, 2004; Campbell, Hombo, & Mazzeo, 2000). Clearly, there is a need to change the way we provide reading instruction to students who fail to develop reading proficiency

Federal Response to Poor Reading Skill

The federal government has invested millions of dollars in research and professional supports to states for implementing proven instructional models that include the use of evidence-based reading practices and curricula. A multidisciplinary approach to improving literacy rates has been employed at the national level, as evidenced by the distribution of funds to various agencies and stakeholders. For example, via the National Institutes of Health and Child Development, federal money has supported the exploration of neurological causes of dyslexia for the purpose of informing intervention protocols (Eden, 2002; Habib, 2000; Shaywitz et al., 2001). Several regional support centers are available for the provision of technical assistance to schools and school districts that implement research-validated practices in reading instruction (Denton, Vaughn, & Fletcher, 2003. Federal funding has supplemented states' efforts to work with struggling readers, and federal sponsorship for meta-analysis of best practices in reading has been instrumental for the creation of national policy.

Critical Research Supporting Legislative Decisions

In 1997, the National Research Council, a branch of the National Science Academy, published the results of an extensive study called *Preventing Reading Problems in Young Children* (National Research Council [NRC], 1998). This report detailed practices in reading instruction believed to prevent reading disabilities through early education and intervention. During the Clinton Administration, the United States

Congress used these findings to craft Goals 2000, legislation that promised every child in America would be reading by third grade.

Congress also directed the National Institutes of Health and Child Development (NICHD) to review all of the reading research available in order to determine proven practices that treat and prevent reading disabilities/dyslexia. As a result, this medical research arm of the government created a panel of experts to conduct a thorough review of available evidence. Members of this group, called the National Reading Panel (NRP), came from the fields of education, medicine, psychology and business. Results of the NRP study were published as the *Report of the National Reading Panel: Teaching Children to Read* (National Reading Panel [NRP], 2001) and have been influential in developing core reading programs, assessment and monitoring systems, and interventions that have a high likelihood of preventing the development of a reading problem.

The NRP report was also extremely influential in creating a more directive congressional response to low literacy by setting requirements for use of federal funds in schools. The No Child Left Behind Act of 2001 (NCLBA) and the Individuals with Disabilities Educational Improvement Act of 2004 (IDEIA) now require, rather than recommend, that remediation programs (a) use research-based curriculum and instructional practices, (b) devise staffing plans that place highly qualified personnel with the most disadvantaged students, and (c) implement parent-involvement policies that align with proven practices. Federal funding is contingent on these prerequisites.

The change to IDEIA (2004) that significantly impacted reading instruction and service delivery models was the provision that allowed states to develop criteria for

defining learning disabilities without waiting for the reader to fail. Prior to this authorization, learning disabilities were defined as a significant discrepancy between potential and performance. Quantifying significance often meant that a student would need to be two years behind his or her peers in a specific academic subject area in order to be eligible for special education identification of a learning disability. The wait-to-fail method for identifying specific learning disabilities evolved into a method of identification that considers the individual's response to instruction (RTI), as defined by the NRP (2001) report. These allowances for alternate identification processes impacted schools by indirectly requiring implementation of the report's model of reading instruction practices.

Instruction Models Supporting Response to Intervention

Although no specific model for early intervention was written into IDEIA (2004), the system of instruction that supports early intervention involves the use of evidence-based instruction and grouping practices for delivering the reading curriculum in the elementary school. Key procedures within this system of instruction assure that each individual receives targeted instruction that is (a) based upon individual need, (b) provided in a timely fashion, and (c) monitored and adjusted frequently based upon data. These data typically come from measures of automaticity with subskills in reading.

There are usually three levels of instructional response that support the acquisition a standardized outcome: Tier 1, Tier 2, and Tier 3 supports. This three-tiered instructional model provides a schema for organizing instructional resources by defining standardized

outcomes and measuring individual growth toward the end goal. Instructional responsibility is divided up through these models with the classroom teacher providing strong initial teaching of the evidenced-based content, as well as temporary skill groupings for individualized support when necessary. Specialists are responsible for assuring effective interventions and providing additional instruction when necessary.

The impact of the NRP (2001) report continues today. According to this report, there are five critical areas of reading instruction that must be addressed every day in our elementary schools. These components of reading are phonemic awareness, phonics, fluency, vocabulary, and comprehension. Instructional targets within each of these reading-comprehension dimensions vary by stage of reading development, and in the United States of America, we have mapped those stages to grade-level expectations. These expectations are rooted in automaticity theory, which explains the way complex skills such as reading are mastered.

Automaticity

Automaticity theory assumes the human brain has a finite amount of attentional capacity, constraining focus on multiple tasks concurrently (LaBerge & Samuels, 1974). Therefore, to be skilled at a complex task, one needs to perform the subskills associated with the task, without conscious thought—automatically (Schwaneflugel, Meisinger, & Wisenbaker, 2006; Walczyk, 2000). This theory helps us understand stages of development for mastering sophisticated skills such as walking or driving. When we learn to walk we do so by mastering incremental steps that support that process as observed in

rolling over, crawling, and furniture cruising. Each of these behaviors provides a critical foundation for mastering the next. The same is true for the complex task of learning to read.

Understanding the process of learning to read in terms of automaticity begins with explaining that the language skills necessary for negotiating a nonprint world need to be developed before the skill of phonemic awareness is possible. In other words, reading proficiency begins with a solid grasp of oral language. Next comes an understanding that the words heard are made up of sounds (phonemic awareness) and that these sounds are associated with specific letters or letter combinations (phonetic analysis). Once these skills have reached the level of automatic retrieval, the individual begins increasing the pace at which words are retrieved (fluency). When one or more of these subskills are not automatic, then the desired reading outcomes of vocabulary and reading comprehension are not fully achieved (Logan, 1997; Schwaneflugel et al., 2006) and reading becomes an inefficient process. In order to detect the likelihood of reading disabilities, elementary schools use curriculum-based measures (CBM) to monitor students' automaticity with these prereading skills.

Further automaticity with phonemic awareness and phonics application in the early years of reading development (kindergarten through first grade) impacts word-reading fluency directly (Good, Simmons, & Kame'enui, 2001; O'Connor & Jenkins, 1999)

To monitor the progress of reading skill development, targets of performance are set based upon grade level and time of year and measured with CBM. For example,

kindergarten reading instruction spotlights skills associated with phonemic awareness and includes the expectation that automaticity with that skill set and all it entails will be accomplished in kindergarten. While vocabulary development and the use of basic comprehension skills are practiced in oral discussions or through acting out understanding, mastering print-based skills is also expected (e.g., basic decoding skills such as learning letter sound associations, a few common site words, and the ability to read CVC words like dad, pig, and but, are the emphasis of kindergarten curriculum). By the end of second grade, the use of phonics is expected to be automatic, and oral reading fluency begins to be the next target for automaticity development. Knowing how well an individual student is performing the targets for automaticity is based on measures of rate and accuracy with the skills related to phonemic awareness, phonics and fluency (Speece & Case, 2001).

Curriculum-based measures efficiently provide indicators of reading health. Most require less than three minutes to administer for each student, allowing schools to monitor progress frequently without significant disruption to the learning process. Commonly used CBMs include (a) one-minute timed readings of grade-level passages measuring oral reading fluency (ORF); (b) reading grade-level passages with missing words, supplying the word that bests fits into the sentence from a list of three or more (cloze exercises); and (c) rapid naming (RAN) activities that require students to name as many items on the testing protocol as they can in one minute. These measures are gathered at least three times a year to monitor progress and identify potential skill deficits.

School District Response to Poor Reading Skill

In recognition of the urgent need to intervene in reading development, many districts throughout the United States have implemented a three-tiered instructional plan. The three-tiered instructional model, as applied in the school setting, requires the use of decision rules (a) to guide the reading specialist and teachers in selecting the content and intervention for a specific target, and (b) to determine the amount of additional instructional time needed. Diagnostic assessments with CBM are used to determine skill mastery targets for the individual. Percentile ranks associated with raw scores guide instructional teams of classroom teachers, reading specialists and administrators in determining the amount of instruction needed to mediate the skill deficit.

Instructional Intensity

Tier 1 instruction is for students on track who need only the recommended 90 minutes of instruction to make progress in the general reading curriculum. Students scoring near the 25th percentile need strategic support and are considered at some risk for not meeting state assessment standards. These individuals would receive Tier 2 instruction in addition to the Tier 1 lessons. Typically, these lessons provide an additional 30 minutes of targeted skill development 4 to 5 days a week until the deficit skill is established at the expected rate of automaticity. Individuals with scores at the 10th percentile or below are in need of substantial support (Tier 3 intervention), as they are at

risk for continued reading difficulties. Tier 3 tutoring increases intervention intensity by using small groups for up to 120 minutes of additional reading instruction each day.

Targeting Skills

Once progress-monitoring data determine the level of instructional intensity needed, the process for structuring the initial intervention begins. Using CBM, the specialist will test each subskill supporting the targeted outcome. Typically, these targets are related to ORF probes. Subskills supporting fluency include phonics and phonemic awareness. Probes for automaticity with phonics are administered to determine whether there is sufficient automaticity with phonics to support the development of fluency. If phonics skills are intact, the intervention then focuses on developing passage-reading fluency specifically. If phonics skills are weak, then a test of phonemic awareness is used to guide the decisions for the intervention target. Once the skill profile for the individual is established, instruction begins based on the subskills that need improvement. Determining the proper curriculum and/or instructional strategy involves understanding the types of interventions available.

Selecting Interventions

With a sense of urgency and little time to wait, schools are turning to commercial programs that target research-based reading skills with methods not yet proven in published evaluations. This is particularly true for students who are nonresponsive to traditional interventions. The best guidance available for evaluating these products comes

9

from the research describing attributes of a successful remediation program. There are generally two types of intervention curriculum: (a) those that target the big ideas in reading development (phonemic awareness, phonics, fluency, vocabulary and comprehension), and (b) those that target the skills supporting reading development (sensory processing and basic language). The literature refers to the former as alphabetic interventions and the latter as pre-alphabetic interventions. It is helpful for practitioners to know whether a curriculum is aimed at alphabetic or pre-alphabetic skill.

Summary of the Issue

Learning to read is a problem for some students. Regardless of the cause, elementary schools are being called upon to assist these students in becoming proficient readers. Several commercial products have been marketed to treat the skill deficits associated with poor reading proficiency and are successfully used in Tier 2 interventions.

One to 10% of readers continue to have complex reading deficits after a series of Tier 2 interventions. For these individuals, sometimes referred to as treatment resistors, complex interventions are required. The science of determining key intervention targets has allowed practitioners to help many students with reading problems, but curricula have stopped at the alphabetic level. There are now models of reading that provide a road map for addressing the base skills supporting the alphabetic skills needed for reading: sensory processing and oral language. These interventions are classified as pre-alphabetic in nature.

Pre-alphabetic interventions on the market generally target three skill sets simultaneously: basic language, phonemic awareness, and phonics. A few also include

methods designed to improve sensory-processing confounds associated with auditory

processing. One product series used in schools, clinics, university research hospitals and

private practice are the Fast ForWord (FFW) computer-based intervention programs. This

study examines extant data from a school-based program evaluation of FFW.

Purpose of Investigation

FFW is a multifaceted curriculum-delivery system that claims to inculcate the

subskills related to reading fluency, including interventions for sensory processing and

basic language, the subskills supporting phonemic awareness. These programs are both an

alphabetic and a pre-alphabetic intervention. However, little is known about the

characteristics a teacher would observe in a student who would benefit from the FFW

program.

With FFW, students work at their own pace, receiving corrective feedback from

the computer program throughout the lessons. This is an attractive feature for intervention

specialists who have been limited to providing support for the few students they can work

with in a group and still make progress (about six). This feature allows one supervisor to

oversee as many students as there are computers, while the student still gets

individualized corrective feedback and repeated practice with the skill being mastered.

This makes FFW products attractive to schools, as it significantly reduces the cost of

fulfilling student-intervention obligations. However, the investment of instructional

minutes and the setup costs associated with purchasing and running a computer-based

intervention must be carefully weighed and deemed worthy of what is truly a considerable

investment. The purpose of the study is twofold: (a) to determine which pre-intervention probes could provide data to match the student to the skills targeted in the intervention, and (b) to provide data that would be useful for making policy decisions regarding the purchase and use of these programs.

CHAPTER II

LITERATURE REVIEW

Learning to read is difficult for some students in schools today. Using systems of instruction that provide intensified skill development for those behind in reaching age-specific benchmark targets has been effective in reducing the number of reading disabilities identified in schools. However, there is a group of students who are nonresponsive to intervention with the alphabetic curriculum traditionally used in these models. Medical models for understanding dyslexia and accompanying treatment implications have expanded the number of possible interventions to use in schools. While alphabetic interventions are based on research from the field of education, pre-alphabetic treatments for reading challenges come from the fields of neuroscience and language therapy.

Interventions

The National Reading Panel (2001) report considered the five constructs that language supports (phonemic awareness, phonics/decoding, fluency, vocabulary and comprehension), but did not investigate the effectiveness of language intervention relative to reading skill acquisition. Most alphabetic curricula seek to develop automaticity with phonemic segmentation, letter names, letter sounds, and whole word reading in order to improve oral reading fluency.

In Figure 1, the foundation of language skill clearly supports all aspects of reading development. Evidence-based alphabetic programs tend to spotlight one block in the reading model for repeated practice and mastery, beginning with phonemic awareness and extending through fluency development. When these alphabetic interventions are substituted for regular reading instruction without additional time devoted to increasing vocabulary and comprehension skill level commensurate with a student's peers, the learning gap rarely closes (Denton et al., 2003; Vaughn, Moody, & Shuman, 1998). When done well, exposing students to remedial and general curriculum simultaneously, alphabetic interventions are sufficient for remediation of reading difficulties and bringing most students to grade level (Anthrop, 2002; Aylward et al., 2003; Coyne, Kame'enui, Simmons, & Harn, 2004). Even then, approximately 1% to 10% of these students do not catch up with their peers (Al Otaiba & Fuchs, 2002; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Shaywitz et al., 1999; Wolf & Katzir-Cohen, 2001). Alphabetic interventions
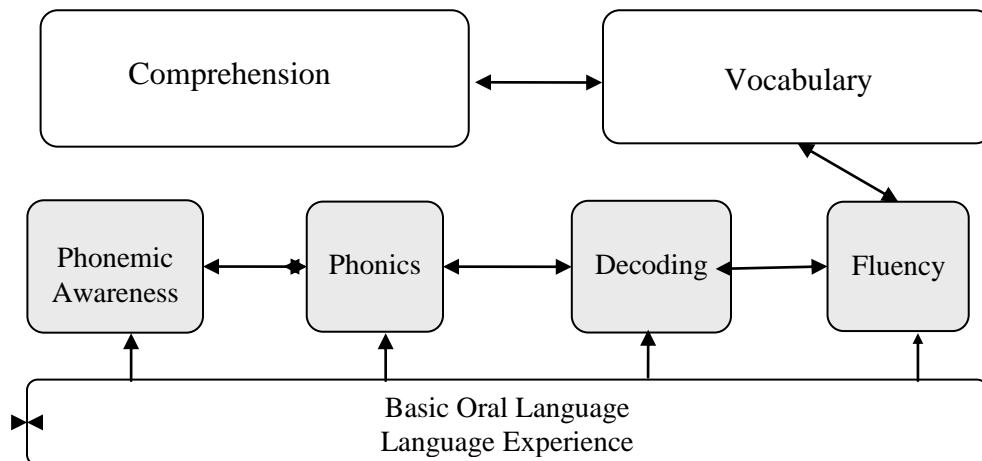

FIGURE 1. Reading intervention blocks based on the NRP (2001) report.

14

reflecting the research synthesized in the NRP (2001) report are commonplace. In the

school setting, many pre-alphabet intervention systems have been used for individuals

with severe language disabilities for a number of years, but only recently have

pre-alphabetic interventions been considered for use as part of reading intervention

programs.

Pre-Alphabetic Interventions

Physicians and speech language pathologists examine reading difficulties with a

different set of procedures and theory. Literature from the medical field provide similar

conclusions about reading disabilities/dyslexia but offer a few key insights that can

inform school-based reading intervention programs. From the medical view,

nonproficient readers are of normal intelligence (Ross-Kidder, 2003; Shaywitz et al.,

1999; Torgesen, Rashotte, Alexander, & MacPhee, 2003) and are without significant

impairments involving sight or hearing acuity (Ramus, 2003; Talcott et al., 2002), which

rules out native intelligence or sensory acuity issues as reasonable explanations of poor

reading behaviors. Educational research has corroborated these conclusions as well.

However, differences exist between the fields with regard to areas of intervention. For

instance, sensory-processing issues have been implicated as treatable symptoms of

dyslexia in studies using functional magnetic resonance images (Helenius, Salmelin,

Richardson, Leinonen, & Lyytinen, 2002; Kujala et al., 2000). Treating these processing

issues as part of a reading curriculum is viewed as somewhat controversial in most

research conducted by educators in the United States. Studies by speech pathologists and

15

neurologists have found that most struggling readers have language-skill deficits as well (Joanisse, Manis, Keating, & Seidenber, 2000; Share & Leikin, 2004). The most prevalent factors/symptoms include (a) poor phonemic awareness (Brown et al., 2001; Engel-Edlar & Rosenhouse, 2000; Ramus, 2003; Ramus et al., 2003); (b) substandard language comprehension (Joanisse et al., 2000; Tallal, Miller, Bedi, Wang, et al., 1996); (c) poor syntax (Vellutino, Fletcher, Snowling, & Scanlon, 2004); and (d) attentional deficits (Faeoetti et al., 2003; Fawcett, Nicholson, & Dean, 1996). Where alphabetic intervention design ignores these language deficit findings, the medical field embraces this as a starting point.

Relating these findings to the alphabetic model for developing reading skill is best accomplished through visual synthesis of the current medical findings regarding the etiology of reading disabilities. The model in Figure 2 shows that poor phonemic awareness is potentially related to the failure of the motoric, sensomatic, auditory, and/or visual processing systems and possibly involves the attentional, working memory and executive function systems at some level. Curriculum addressing pre-alphabetic concerns generally do not involve printed words. Rather, the exercises target language development, noted as foundationally necessary by the NRP and one or more of the processing systems shown in the model developed by Alexander and Slinger-Constant (2004). According to Alexander and Slinger-Constant (2004), Ramus et al. (2003), and Vellutino et al. (2004), the visual representation in Figure 2 is useful for understanding the processing systems that support each of the subskills of reading comprehension (language, phonics and fluency). Given poor performance with language or phonemic

16

awareness, a possible treatment course could target development of any of the components of phonemic awareness shown in Figure 2. In that way, each of these blocks represent another potential area for targeted practice.
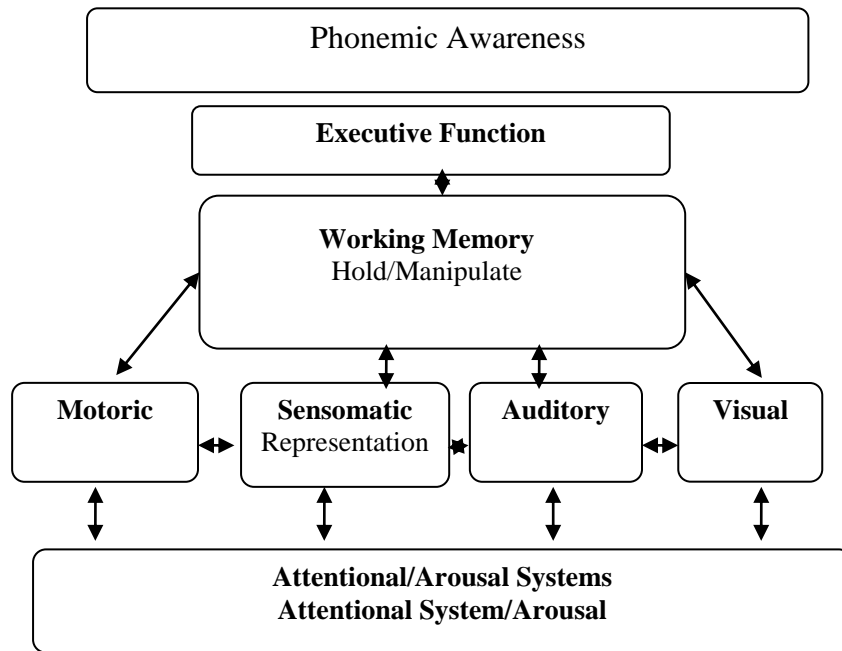
FIGURE 2. Structures supporting phonemic awareness.

Commercially Available Pre-Alphabetic Interventions

Three products have emerged from the literature as potentially effective for treating the processing deficit and the language skill deficit simultaneously: Earobics, LiPS, and Fast ForWord. Each of these programs has been designed to be administered in small-group settings or with computer software, providing frequent opportunities to respond, with immediate and specific feedback for the learner. The differences between

the products can be found in the program delivery method (computer-based versus teacher-based) and the thoroughness with which the construct of language is addressed.

Lindamood Processing System (LiPS)

This is a multisensory, teacher-driven curriculum usually administered in groups of one to four. Exercises focus first on locating and naming sounds in the environment, then one-to-one correspondence with sound sequences and objects, followed by actual letter-sound correspondence. The LiPS treatment systems have had notable results in improving phonemic awareness in elementary-aged students (Pokorni, Worthington, & Jamison, 2004). The instructional design includes multiple opportunities for practice and feedback, but the time on task is reduced, even for individuals in groups of one to four. Evidence of neurological changes following treatment has established credibility for the program's ability to change neuro-firing patterns so they more closely represent those of a nondisabled reader (Simos et al., 2002).

Although the program is aimed at improving phonemic awareness, it does little to develop the language usage and comprehension skills commonly missing from the truly challenged reader's profile. Further, there is nothing in the company literature that suggests the goal of the program is to improve a sensory-processing deficit. In fact, the exercises in the program are designed to present information from multiple sensory channels in hopes of using the strength of one system to attain the phonemic skill needed. The narrow focus of the program, the staff-intensive nature of instructional delivery and

lack of sensory-specific training are considerations for decisions about the purchase or use of LiPS as a pre-alphabetic intervention.

Earobics

A second program that has shown evidence of correcting language-processing deficits is Earobics, published by Cognitive Concepts (1999). Earobics is a computer software program designed to treat the language-processing deficit by overteaching the sounds of language, the phonetic representations associated with the sounds, and many discrete language skills that interfere with comprehending spoken language. This is an adaptive training program that teachers can manually modify in response to student performance levels.

Although the company literature makes claims of treating auditory-processing deficits, the neurological evidence to support these claims is missing from the literature reviewed. This product may be an effective treatment for those nonresponsive to alphabetic treatments alone, but evidence outside of company-based reports is missing from the literature.

Fast ForWord

The third program series is the Fast ForWord family of products produced by Scientific Learning Corporation, which include Fast ForWord Language (FFW-L) and Fast ForWord Language to Reading (FFW-LR). The products are designed to improve basic language skill, phonemic awareness, phonics application, auditory memory, and

sensitivity of the brain to discern sounds of various frequency and duration (Scientific Learning, 2010). Auditory processing is trained by synthetically slowing speech so that learners develop the ability to make distinctions between similar sounds—e.g., /b/ and /d/. Other training exercises bolster language-specific skills related to grammar and syntax, as well as auditory memory. The Fast ForWord (FFW) programs are described in much greater detail in the Methods chapter of this dissertation, while the reported effects of treatment on specific skills are analyzed here. This is the product under investigation in my study.

Effects of Fast ForWord

The Scientific Learning Corporation (SLC) has published over 150 reports documenting results of FFW training on specific populations (Scientific Learning, 2010). Although many of the studies were well designed, and considered to be scientifically sound experiments by evaluators at the What Works Clearinghouse (2002), the bias of the document source cannot be overlooked. Therefore, the only SLC articles given consideration in this review are those that describe foundational research prior to the release of FFW (Merzenich et al., 1996; Tallal, Miller, Bedi, Byma, et al., 1996).

The published research-based literature regarding FFW is mixed with the basis for evaluation. For instance, SLC built the products based on a theory that reading disabilities are caused by an inability to hear rapidly transitioning stimuli. Arguing that this theory is inaccurate, several practitioners have found that changes in auditory process rate are

unrelated to improved reading (Gillam, Crofford, Gale, & Hoffman; 2001; Habib et al., 2002; Marler, Champin, & Gillam, 2001; Thibodeau, Friel-Patti, & Britt, 2001).

A second source of research comes from medical practitioners who used neural-imaging data to measure effects for FFW and confirmed the theory that FFW changes neural pathway development to the maps of typical readers (Gaab, Gabrieli, Deutsch, Tallal, & Temple, 2007; Stevens, Fanning, Coch, Sanders, & Neville, 2008; Temple et al., 2003; Temple et al., 2000), a process that cannot be duplicated in the school setting. If outcome measures from these studies include pre-alphabetic or alphabetic skills, the results are reported later in this review.

The third group of professionals to investigate FFW were speech and language therapists, most of whom were in private practice or worked for university clinics. These practitioners conducted independent single-subject studies or field tests of FFW in clinics, using growth in scores for basic language and early reading skills to measure effects (Deppeler, Taranto, & Bench, 2004; Gillam, Crofford, et al., 2001; Hook, Macaruso, & Jones, 2001; Loeb, Stoke, & Fey, 2001). These researchers, primarily speech therapists in clinical practice, tested the products for the purpose of determining whether the synthetically slowed speech made an impact on student receptive language. A few reported results in terms of alphabetic reading skills, and those results are included in this review.

The fourth source for research comes from school-based program evaluations measuring (a) FFW effects on changes in state test scores from one year to the next (Borman, Benson, & Overman, 2009; Cohen et al., 2005; Divine & Botkin, 2008; Hall,

21

2002; Overbay & Baenan, 2003; Rouse, Krueger, & Markman, 2004); and (b) other

subskills associated with reading development (Pokorni et al., 2004; Troia, 2004; Troia &

Whitney, 2003).

Using the stages of development for reading as an organizing structure, this

review of literature reports studies based on the effects of program completion. First, a

discussion of the effects of FFW on pre-alphabetic skills (basic language and processing)

is conducted. Then a review of the studies testing for program effects on alphabetic skills

(phonemic awareness, phonics, and comprehension) illuminates trends in score changes

and the participation criteria used to enroll individuals in the study. Missing from the

review are studies that test effects on reading vocabulary and fluency, as these variables

have not been used as growth indicators in the studies available for consideration.

<div align="center">Effect of FFW on Pre-Alphabetic Skills</div>

Pre-alphabetic skills are those associated with the use of oral language and

sensory processing. No alphabetic understanding is needed to use these skills. They are

the very base support of reading development in the automaticity-based understanding of

reading skill development.

Processing Skills

Two studies measured effects of FFW on auditory processing with tools available

for use by the school-based practitioner. In Australia, Deppeler et al. (2004) used a single-

subject experimental design with pre-, post-, and follow-up measures of language ability

that included subtests of auditory processing, memory for sentences, and digit span. Both measures are easily administered in the school setting and considered CBM. Subjects were primary-aged children (6 to 9 years old, mean age = 8 years). Treatment occurred each weekday for an average of 16 days.

At the conclusion of the experiment conducted by Deppeler et al. (2004), one subject experienced clinically significant gains (> 1 *SD*) on the digit span measure. Another subject had similar gains on the memory for sentences probe. At the one-year follow-up, two different students made clinically significant gains on those measures. Those with significant growth at posttest maintained their gains at follow-up as well. Language abilities at the beginning of the study ranged from normal (*n* = 3; scores > than the mean), to mild delays (*n* = 2; scores < 1 *SD* below the mean), to severe impairments (*n* = 3; scores > 2 *SD* below the mean) based on CELF (Semel, Wiig, & Secord, 1995) or TOLD: P (Newcomer & Hammill, 1997).

The second study was conducted by Stevens et al. (2008). These researchers used neural imaging technology to look for effects of treatment on brainwave patterns and skills simultaneously. Using these two sources of information as a basis, Stevens et al. found an effect for increased auditory attention when measured by neural-imaging techniques and other assessments. According to their research, FFW appears to train auditory attention for students with low basic language scores prior to intervention. However, the relationship between improved processing rates and improved language skill is not directly attributable to the impact of FFW. There may be other mediating factors that are not explained in the model.

Oral Language Ability

Corporate-sponsored field trials of FFW prototype games took place in 35 clinics throughout the United States. Over 300 clients were tested on measures of auditory processing and basic language pre- to postintervention at each site and scores were reported to SLC (Scientific Learning, 2010). Of the children who participated in this field trial, 90% made clinically significant gains on at least one measure following treatment (Scientific Learning, 2010). Since that time, the products include far more exercises than those used for the laboratory and field trials.

The studies designed to duplicate corporate field trial outcomes (Gillam, 1999; Gillam, Frome-Loeb, & Friel-Patti, 2001; Loeb et al., 2001; Veale, 1999; Turner & Pearson, 1999) did not yield results as robust as those reported by Scientific Learning (Merzenich et al., 1996; Tallal, Miller, Bedi, Byma, et al., 1996). This has caused criticism of the FFW products, including conclusions that SLC could not back up claims that these products changed neural firing patterns as stated in the promotional literature (Gillam, 1999; Veale, 1999). Effects of FFW on changing neuronal activation patterns have since been documented in peer-reviewed sources (Aylward et al., 2003; Stevens et al., 2008; Temple et al., 2003). The issues related to duplication of results for the initial field trials remain a concern.

Independently, single-subject program evaluations from several researchers have determined that FFW exercises significantly improve expressive-language ability, when practiced in the clinic (Gillam, Frome-Loeb, et al., 2001; Turner & Pearson, 1999) or

24

home setting (Loeb et al., 2001). Although several reports such as these are available for thorough analysis, this report is written with the reading intervention specialist in mind. Therefore, studies that reported effects on language scores exclusively are not included in this examination of literature. It is likely that the degree of language ability prior to treatment would be helpful data for matching poor readers to promising reading outcomes. The remainder of the review will discuss the effects of FFW on various reading skills. To assist the reader, this discussion will report the results of others who studied FFW by using the reading skill development pathway explained in the NRP (2001) report: phonemic awareness, phonics, fluency, vocabulary, and comprehension.

Phonemic Awareness

Cohen et al. (2005) used a multicenter random-control design to compare two computer-based language interventions: FFW products and an eclectic mix of computer games matched to FFW but without modified speech. Seventy-seven clients with severe language impairments, who ranged from 6 to 10 years old, were enrolled in the study. Although the study predominately used measures of basic language to report treatment effects, two reading-related subtests were also included in the assessment battery: word reading, and rhyming.

Participants were divided into three groups: (a) those using FFW ($n = 23$), (b) those using the other computer language intervention ($n = 27$), and (c) a noncontact control group ($n = 27$). The purpose of the study was to determine whether the synthetically generated speech used in the FFW computer program caused significantly

25

different outcomes between groups. ANOVA analyses compared group scores and showed a significant effect ($p < .05$) favoring the FFW participants for the skill of rhyming. Both treatment conditions produced similar outcomes on all other indicators.

Like Cohen et al. (2005), Pokorni et al. (2004) compared language development programs using pre/post measures of basic language skill. However, they used far more prereading skills as part of the assessment battery (phonemic awareness, letter word identification, phonics application) and included reading comprehension measures (cloze tasks). Participants were students ranging from 7 to 9 years old who were enrolled in a 20-day summer school program based on teacher or speech therapist recommendation.

In the study conducted by Pokorni et al. (2004), each of the participants were randomly assigned to one of three treatment curricula when they reported for summer school: (a) Lindamood Phoneme Sequence Program (LiPS), (b) Earobics, or (c) FFW. Site 1 enrolled 32 students and assigned each to one of the three intervention programs. Site 2 enrolled 30 students and assigned individuals to two programs: FFW or Earobics. Post-summer-school data available for analysis were collapsed between sites, leaving 20 students in the FFW group, 16 in the computer-based therapy group using Earobics software, and 18 in the LiPS intervention, which is a therapist-client treatment protocol. The results of this experiment found that all three treatment conditions caused clinically significant gains (score changes > 1 *SD* pre to post) on the phonemic awareness measures. No significantly different gains for any group or any other reading skill were evident.

Matching pre-intervention skill to likely outcomes is part of the job of the school-based intervention team. Typically the team tests skills and matches them to curricula that

address the deficit skill pattern. While reviewing articles reporting treatment effects for FFW, I considered pre-intervention characteristics of the population in order to discern patterns that may be useful for intervention teams. With regard to phonemic awareness, individuals with documented severe language deficits (scores > 1.5 *SD* from the mean) are likely to make significant progress on phonemic skills (Cohen et al., 2005; Pokorni et al., 2004), but FFW was not better than other language intervention curricula in producing the effect.

<div align="center">Effect of Fast ForWord on Alphabetic Skills</div>

Alphabetic skills are those that require the individual to interact with print, including phonics application, fluency, vocabulary and comprehension. The outcome measure used for each of the constructs associated with alphabetic skills varies greatly.

Phonics

A variety of phonics skills have been used to test the effectiveness of FFW. These include word attack, letter naming, and word reading. The development of phonics skill begins with recognizing that letters make sounds and sounds make words. As the reader develops fluid use of phonics skill at the letter level, instruction begins to focus on words at the syllabic level. Assessments for each of these types of phonics application have been used to test FFW. The subtle differences between these skills will not be summarized in this portion of the review, but categorized broadly as phonics skill.

A program evaluation conducted by Hook et al. (2001) compared effects of Fast ForWord ($n = 11$) versus Orton-Gillingham ($n = 11$) for 6- to 9-year-old participants placed in summer school because of low language skill. This was a two-phase evaluation. Phase 1 was designed using a pre-, post-, and follow-up data-collection cycle measuring expressive language, and the reading subskill of phonics over a two-month period. Phase 2 compared the FFW participants to a control group matched on pretreatment language ability who did not receive the summer school interventions. Those two groups were compared annually for the two years following the summer school program.

During Phase 1, the FFW group experienced significant gains on composite scores for speaking (t (10) = 4.24, $p = .002$; effect size: $r^2 = .64$) and syntax (t (10) = 5.45, $p < .001$; effect size $r^2 = .75$) when compared with scores from those in the Orton-Gillingham treatment condition. Both groups made statistically significant gains on phonics and phonemic awareness, but neither program was more effective than the other in producing those gains. There were no other significant gains recorded for either group during the first phase of this study.

The second phase of the experiment involved the 11 students who participated in the FFW group and nine noncontact control cases. Both groups were screened using the Woodcock Johnson Reading Mastery Test and a standardized basic language assessment. At the end of Year 1, statistically significant effects for reading skills were found for the FFW group in the area of word attack ($Eta^2 = .22$, $p < .001$). No other effects for FFW maintained at the two-year mark.

Although the effects of FFW for basic language were most often immediate (Gillam, Frome Loeb, et al., 2001; Loeb et al., 2001; Turner & Pearson, 1999), effects for reading skills were mostly delayed. Evidence of delayed response to FFW was found for auditory-processing skills (Deppeler et al., 2004) and for phonics as well (Hook et al., 2001). This makes sense in light of SLC claims that FFW creates the neural foundation for reading skill development (Scientific Learning, 2010).

However, the samples sizes used in these studies have been less than 30, suggesting caution be applied when making generalized statements in relation to the potential of delayed effects—especially since FFW had immediate effects on phonics skills in the study reviewed next.

Troia and Whitney (2003) analyzed data from a Migrant Education summer school program in the state of Washington. Students participated in FFW training or academic enrichment activities at several sites across the state without using a random assignment process. The summer school program location determined the type of program the participant was involved with. To analyze the effect of FFW (treatment, $n = 99$) on reading skills, students were matched on age and English proficiency levels ($n = 92$) with other students who were in the non-FFW summer school program. Unlike Hook et al. (2001), who found a delayed effect for FFW on the reading subskill of phonics, Troia and Whitney (2003) obtained an immediate pre/post effect for phonics application ($ES = .31$, $p < .05$) favoring FFW completers.

A secondary data analysis was performed by Troia and Whitney (2003) to determine whether those with severe language deficits responded to FFW treatment

differently than the entire sample. For that analysis, cases were sorted to include only those with significantly deficient pre-intervention scores (25th percentile or below) for expressive language, receptive language, or English language proficiency. Twenty-five FFW participants were within this subgroup; 23 members from the comparison group had scores in that range. Students with low basic language (expressive or receptive language skills) did not score differently from one another on any reading related subskill pre- to postmeasure. Among students with low English language proficiency, FFW participants significantly outscored their peers on tests of letter-word identification ($ES = 1.03$, $p < .05$) and word attack ($ES = .98$, $p < .05$). No long-term follow-up analyses were reported. As a pre-intervention characteristic to screen for, substantially low scores on measures of limited English proficiency may be useful for guiding intervention team decisions.

The following year, Troia (2004) completed another study, this one involving native English-speaking first- through sixth-grade students enrolled in one of two skill-support programs during the regular school year. Participants were students who failed the state reading assessment the year prior to intervention. Parental consent for participation in the evaluation portion of this project was required, and 13 families of students assigned to the control condition did not return consent forms. Therefore, the treatment group ($n = 25$) was larger than the comparison group ($n = 12$), but equivalent with respect to age and IQ.

Pre/post assessments included tests of expressive language, phonics application, and sight word reading. When posttests were compared, there were no statistically significant differences between group test scores to report for any of the measures.

30

However, students with expressive language scores below the 25th percentile at pretest significantly improved word-blending skills posttest (t (20) = 2.73, $p < .05$, $d = 1.17$) when compared with skill-matched peers. This adds to a pattern of evidence suggesting that extremely low oral language scores based on a student's primary language may predict improved reading skill (Cohen et al., 2005; Troia, 2004; Troia & Whitney, 2003).

Teacher referral to intervention programs may not be a reliable method for making treatment decisions, however. Given, Wasserman, Chari, Beatie, and Egan (2008) sought to determine if the two computer programs, FFW and SuccessMaker (SM; Computer Curriculum Corporation, 1995), could be used separately or in combination to improve reading for seventh-grade students. The method of referral to this study was teacher perception of at least a one-year skill gap between the student and his or her peers. Treatment outcomes were measured by changes in phonics skills and comprehension.

Given et al. (2008) used a complex experimental design to explore this question. The evaluation included five possible treatment conditions for the 12-week term. The groups were divided as follows: (a) those spending 12 weeks in FFW treatments exclusively ($n = 12$); (b) those who spent 6 weeks in FFW and 6 weeks in SM ($n = 15$); (c) reversing the order of intervention, those who spent 6 weeks in SM programs followed by 6 weeks in FFW ($n = 11$); (d) those with 12 weeks' experience with SM ($n = 14$), and (e) students who enrolled in an enrichment class (control condition, $n = 13$). Although the study reported the use of a randomized method to assign treatment condition, the method for doing so was unexplained. All groups were established to be statistically similar to one another on the basis of IQ and pre-intervention skill probes.

A repeated-measures ANOVA showed a significant effect for time for all groups, but no between-groups effects emerged for the reading skills assessed—meaning all versions of treatment, including enrollment in an enrichment class, had similar effects on reading skills for this seventh-grade group. Without differential effects for the FFW treatment group to report, the authors concluded that the expense of the FFW program was not justified.

More recently, the effects of two computer-assisted learning programs on language and reading scores were reported based upon a clinical trial of FFW products (Frome-Loeb, Gillam, Hoffman, & Marquis, 2009; Gilliam et al., 2008). The clinical trial was funded by the National Institutes of Health and Human Development. The study compared four groups of 6- to 10-year-old summer school attendees from two geographically different regions of the United States. Each site used two different computer-based interventions, one client-therapist treatment protocol and an enrichment program, as comparison groups. Random assignment to treatment condition was conducted as the student arrived on site, with the first child going to Group 1, the next child to Group 2, the third child to Group 3, and so forth. Each of the four groups had 54 participants. Participants were recruited based on known language development concerns.

The purpose of the study was to determine which type of intervention system worked best, controlling for computer effects and program type. To eliminate biases associated with human judgment, interventionists were blind to the purpose of the study, individual student characteristics, and the other treatment conditions used in the evaluation. The same was true for parents and for the statisticians gathering evaluating

the data. Pre-, post-, and follow-up measures of language and pre-alphabetic skills were taken for each child.

The researchers concluded that FFW was no better at treating language disabilities or improving reading skill than any of the remaining groups at posttest. However, at follow-up, a delayed effect for word blending was obtained for both computer-based interventions. Tests of significant differences between follow-up scores for each group (FFW versus Other Computer) yielded no statistically significant differences. Yet, the non-FFW computer condition produced only a moderate effect ($d = .62$), while FFW participants experienced a large effect ($d = .79$) for word blending six months posttreatment.

Little can be concluded from this study with regard to pre-intervention characteristics that may be useful for making decisions about the best candidate for treatment with FFW, but the evidence of delayed effects for FFW with regard to skills supporting comprehension continues to grow.

Reading Comprehension

Studies that tested effects of FFW on reading comprehension come mainly from school-based program evaluations. Most often the program evaluation designs were quasi-experimental as very few have random controls in place. The comparison group was often a noncontact group that had not been evaluated with any assessments outside of the dependent variable. Source documents reporting results from program evaluations are addressed to school district leaders for the purpose of supporting decisions regarding the

use or purchase of the intervention. The current study involves data from a program evaluation. Therefore, considerations of white papers were entertained (Divine & Botkin, 2008; Hall, 2002; Overbay & Baenan, 2003;). Results reported in peer-reviewed sources regarding FFW's effects on comprehension follow the discussion of white paper reports (Borman et al., 2009; Rouse et al., 2004).

White Papers

In 2002, Hall submitted a program evaluation of FFW based on results of the Texas Assessment of Academic Skills (TAAS) and the Texas Learning Index (TLI) to the general superintendent of the Dallas Independent School District. During the 2001-2002 school year, nearly 1,350 students underwent FFW treatment. These students came from 10 different schools where individuals were assigned to FFW based upon teacher recommendation and a failing TAAS or TLI score. The comparison group was not formally assigned; rather, it included all other students in the district not participating in FFW who (a) had failed the TAAS or TLI indicators the same year as the training group, and (b) had similar demographic descriptors as the treatment group. Outside of results on the TAAS and the TLI as pre- and postmeasures, no other assessment data were collected.

Several details regarding treatment effects by location were reported in the document, but in the end, treatment effect was determined by data from 94 treatment participants and 94 noncontact peers in Grades 3 through 8. The research controlled for grade level; those completing FFW treatment experienced a statistically significant change in reading scores following completion of FFW ($Eta^2 = .01$, $p < .05$), but a low

effect size for the change. A regression analysis found limited English proficient (LEP) status negatively related to score changes.

The findings of Hall (2002) contrast with those of Troia and Whitney (2003), who found severe deficits with English language skill predictive of favorable treatment outcomes on phonics and phonemic awareness. This may be due to the fact that reading comprehension is a much more complex skill, one not sensitive to growth. If so, then effects for FFW are better represented by measures related to the subskills of comprehension. More likely, the LEP status indicator used in the school setting is too broadly qualified to capture meaningful information. The same is likely to be true with the indicator called Special Education.

The second white paper reviewed came from a program evaluation published by Overbay and Baenan (2003) for the Wake County Public School system in Florida. The dependent variable was the Florida End of Grade (EOG) assessment, and the evaluation involved students from Grades 3 through 8 who failed to meet expectations on the EOG assessment for the 2001-2002 school year. During the 2002-2003 school year, 616 students participated in at least one of the FFW training programs, including 80 students in third grade, 69 in fourth grade, 82 in fifth grade, 60 in sixth grade, 142 in seventh grade, and 98 in eighth grade. Of the 616 program participants, 426 were successfully matched with pre/post test data for EOG scores in the spring of 2003.

Overbay and Baenan's (2003) evaluation occurred in two phases. First, an ANOVA analysis used gain scores on EOG tests from Year 1 to Year 2 to compare scores between FFW participants and all other students in the school district. Statistically

significant differences in gain scores ($p < .05$) on the EOG were found, favoring FFW

participants for each grade level except sixth. However, when a control group was

established matching individuals based upon English proficiency status, special education

participation, free and reduced meal status, and grade level, there was no evidence of the

effect at any grade.

The second part of the Overbay and Baenan (2003) evaluation was a within-group

comparison of FFW participants that separated participants into two groups: those who

met EOG expectations the year after treatment, and those who did not. Using a

standardized least squares regression to estimate parameters for the full model and a

stepwise regression model, Overbay and Baenan sought to identify factors that were

significantly different within the treatment group. The only variable that predicted a

change in EOG testing results from *not-pass* to *pass* was the pretest score. Unfortunately,

the authors of the study did not seek to compare actual score differences between the

pass/not pass groups, which may have yielded important information related to treatment

decisions in the future. Instead, the authors simply reported that there were no

demographic identifiers associated with performance outcomes.

Included in Overbay and Baenan's (2003) evaluation was a survey of behavior

completed by the classroom teacher prior to student involvement with the FFW program.

Total scores on this instrument were not influential for explaining differences in

outcomes either, again calling into question the usefulness of teacher impression in

making placement decisions for FFW treatment.

The results from both Hall (2002) and Overbay and Baenan (2003) failed to implicate low achievement on state-required tests as helpful indicators for deciding to place a student into FFW. However, when data are tracked longitudinally, there appears to be a delayed effect for FFW with regard to reading comprehension, as explained by Divine and Botkin (2008).

Divine and Botkin (2008) used a longitudinal data analysis that measured program impact by comparing gain scores on the Florida Comprehensive Achievement Test (FCAT) from year to year for four years. Students in Duvall County, Florida, who started FFW interventions during the 2005-2006 school year, were eligible to be in the statistical review. Three criteria were used to create the comparison group; students needed to be (a) continuously enrolled in Duval County Schools from 2005-2008, (b) have FCAT scores the same as grade-level peers in FFW for the baseline year of 2005, and (c) have FCAT scores for all years of the study. This left 5,010 students in the FFW training group and 5,219 in the comparison group. Those completing one or more FFW programs increased scores over a three-year period by 405.07 ($SD = 434.88$) points on average. Those with no FFW intervention improved an average of 221.84 ($SD = 430.37$) points over the same three-year interval. This yielded a statistically significant difference ($p < .0001$) favoring FFW participants.

No differential scores between grade levels were reported, as all gains were averaged regardless of age. This finding supports conclusions about FFW effects being delayed and, with a 100% difference in gain scores between the groups, the results may provide evidence for schools considering use of FFW programs. However, several other

37

factors could explain these score differences. There are no controls for school,

socioeconomic status or other types of interventions that these individuals may have been

involved with after FFW. The effect could be due to other programs adopted and have

nothing at all to do with FFW.

White paper reports vary in degree of utility for making generalizations about

results. The larger the number of participants, the more stable the statistical process, but

variables associated with outside influences—e.g., types of follow-up instruction for the

participants, or duration of additional interventions outside of FFW—go unaccounted for

in most of these documents. The peer-reviewed articles reporting evaluations of FFW and

effects on reading comprehension account for more of these unreported confounds.

Peer-Reviewed Publications

In Baltimore, Rouse et al. (2004) were hired by the school board to determine the

cost effectiveness of implementing FFW given the district's continuous investment in

Success for All (SFA) school reform efforts. To test the effectiveness of FFW on reading

skills, these researchers invited students who scored below the 20th percentile on the state

assessment to participate in a study that placed students into two conditions: (a) treatment

with FFW, or (b) a noncontact condition until the evaluation processes had ended.

Additional requirements for treatment eligibility included enrollment in third, fourth, or

fifth grade during the 2000-2001 school year in a school using SFA reform strategies.

Rouse et al. (2004) required parents of qualifying students to sign permission for

involvement in the study. Students with signed consent were randomly assigned to the

treatment or control condition in each of four schools involved. In other words,

randomization was at the student level, versus the school level. Each study participant

took several pre/post assessments, including (a) the Reading Edge test published by SLC,

(b) the CELF language battery, (c) *Success for All* assessments, and (d) the state-

mandated reading test. Data were collapsed across the four buildings to represent

treatment and control groups. There were data from 237 FFW-trained students and 217

from the control condition for the evaluation.

Student growth on reading comprehension measures was significant for FFW

participants by effect size estimates ($p < .05$), but not greater than the growth experienced

by the control group. Interestingly, the time spent training with FFW was outside of the

regular reading instruction time, meaning the FFW treatment group got substantially more

reading intervention time than the controls. Given additional instructional minutes in the

skills that support reading comprehension, it would be expected that FFW participants'

gain scores would be greater than those of their matched peers. This was not the case, and

Rouse et al. (2004) concluded that implementation of FFW was not worth the additional

investment required.

Most recently, Borman et al. (2009) published the results of a study involving

students in second and seventh grade within the Baltimore City Public School System in

Maryland. Participants were students who scored at or below the 16th percentile on the

Comprehensive Test of Basic Skills, Fifth Edition (CTBS/5), which was taken by all

students in the state. The experiment was conducted across eight schools, using a

sampling plan that placed students in either a treatment or control condition within each

building. At the time of data analysis, second grade had 107 total participants, with 51 in

FFW and 56 in the academic enrichment group. In the seventh grade ($n = 180$), 94

students were in the FFW program and 86 in the enrichment group. In a comparison of

scores and demographic considerations in a multivariate analysis of variance

(MANOVA), seventh-grade students responded to FFW with a statistically significant

effect for reading comprehension (Cohen's $d = .5, p < .01$) while the second-grade group

obtained no effect for FFW. These results provide a bit more evidence that FFW

improves reading comprehension for some students, adding an additional consideration

for extremely low reading performance on pre-intervention state testing to the matching

process.

Summary

Several researchers have suggested that FFW is an appropriate intervention for

some elementary students, but they also noted there was little guidance for the placement

of students in the curriculum (Gillam, Frome-Loeb, et al., 2001; Hook et al., 2001;

Temple et al., 2003; Trioa and Whitney, 2003; Turner & Pearson, 1999). Collectively,

they presented a call for research centering on the development of a profile of test scores

and age or grade considerations that could place students likely to respond into this

intervention. Common themes emerged from the evaluations of FFW in relationship to

potential pre-intervention characteristics that may inform the intervention team regarding

the use of FFW.

When eligibility for treatment with FFW was based upon substandard levels of performance on state tests, the completion of FFW improved reading skills in some experiments (Borman et al., 2009; Hall, 2002; Overbay & Baenan, 2003) but not all (Given et al., 2008; Rouse et al., 2004). This pattern repeated regardless of whether the dependent variable was another state test, phonics or phonemic awareness. Level of state test score performance alone did not appear to be a useful screening criterion for matching interventions with FFW candidates.

The combination of demographic characteristics and pretest scores failed to point to criteria that could be used as screening considerations (Hall, 2002; Overbay & Baenan, 2003). When combined, teacher recommendation for participation and failure to meet standards on the state test did not accurately indicate an improved reading skill following treatment for the FFW group. However, when additional screening criteria relating to a skill level on other measures were combined with having a test score below expected state targets, differential effects favoring FFW were reported (Troia, 2004; Troia & Whitney, 2003).

A history of language disability was indicative of improved reading-related subskills posttreatment for phonics (Cohen et al., 2005) and phonemic awareness (Hook et al., 2001; Pokorni et al., 2004). This was true regardless of instructional setting (summer school, clinical offices, or skill support class during the regular school year). Based on the FFW literature, the exit criteria for program completion stipulate that skills supporting fluency must be intact. In relationship to automaticity theory, this would indicate that the individual is ready to master fluency. Curiously, the literature failed to

41

yield any experiments that use oral fluency as a dependent variable. It was expected that this variable had been studied previously, as many intervention systems rely on this measure to monitor reading skill and make decisions related to instructional intensity. Also missing from the literature were experiments that used vocabulary as an indicator of program effectiveness. Vocabulary CBM that are reliable and valid for measuring vocabulary have become readily available and will likely be used as part of an evaluation of FFW in the near future.

Research Questions and Hypotheses

To date, studies of FFW that considered automaticity with word reading (ORF) as an outcome for program evaluation have not been published—perhaps with good cause, as oral reading fluency is not practiced in the program exercises. However, fluency is supported by automaticity with the pre-alphabetic and alphabetic aptitudes addressed in the curriculum (Logan, 1997; Schwaneflugel et al., 2006). Therefore, theoretical support for an improvement in fluency exists, but only if students continue to get reading support in the regular classroom, which should provide daily instruction in fluency. This study evaluates this theory.

Using an extant data set from a study completed during the 2002-2004 school years, this dissertation examines three questions. The first question addresses practical issues related to purchasing and implementation decisions regarding the effectiveness of FFW for changes in oral fluency.

1. Is there a significant difference between oral reading fluency scores over time for Fast ForWord intervention students [treatment] when compared with nonparticipating peers [comparison] matched on grade? For implementation decisions to be data driven, there is a need to determine if FFW is more or less effective than other programs at improving fluency rate, a traditionally explored question for program evaluations. This tests the hypothesis that FFW works as well or better than other programs for improving fluency.

However, the intervention specialist is not interested in averaged score comparisons, as they evaluate programs based on the impact the program has on academic risk status. The first question assists with making decisions with regard to purchasing the program. If FFW is not better than other programs for treating fluency skills, then it should not be purchased as a fluency intervention. The second question addresses the need of the intervention specialists who lead teams of teachers and other school professionals in making instructional decisions based on risk for reading difficulties in the three-tiered instructional system. The question tests the hypothesis that FFW works as well or better than other programs for shifting oral reading fluency performance from one tier of intervention to a less intense tier (a decrease in reading risk status).

2. Is there a significant difference between the numbers of students improving risk status among Fast ForWord intervention students when compared with nonparticipating peers?

The third question was based on the assumption that it is possible to match interventions to skill deficit patterns using CBM or other measures of reading subskills, specifically expressive and receptive language.

3. Which pre-intervention language skills predict a change of risk status for FFW participants over time based on oral reading fluency score changes?

Based on the assumption that some participants in the FFW group would exit a reading risk category and some would not, the third question targets the hypothesis that there will be significant differences in pre-intervention language skills between those who changed risk status and those who do not.

CHAPTER III

METHODS

This study was designed to inform school-based practitioners who are making intervention decisions for students who are not meeting grade-level expectations in reading and for those making policy decisions about whether to purchase or use the FFW program already available. The investigation evaluates the effects of FFW on reading fluency rate changes for poor readers, and investigates any pretreatment student characteristics, which may inform the intervention specialist about optimal placement decisions.

Study Setting

The data obtained for this current study came from a program evaluation that took place in a suburban school district in the northwest. The evaluation of FFW products occurred from 2002-2004. At that time, the district had a student enrollment of over 11,000 students; 12% were Hispanic and over 80% were White. Nearly 55% of all students qualified for the national school lunch program, a measure used to establish the poverty rate. There were 16 elementary schools, six middle schools, two high schools and two alternative schools at the time of evaluation. Performances on state assessments in reading for third and fifth grades have been equal to the state average or just slightly lower for the six previous years.

A single K-5 elementary school implemented FFW curriculum as an after-school reading intervention program. Members of the treatment group came from that school exclusively. The comparison group were students from all other schools in the district, making reports of building-based context difficult. However, the context of the school for the treatment recipients is worthy of description.

The treatment school was one of nine in the district receiving funding for reading interventions through Title I (NCLBA). These funds increased the staff membership in such a way that 15 additional adults had been available to work with struggling students for five years prior to the FFW pilot study. The building was one of three in the district that had a full-time teacher English Language Learner (ELL) specialist and full-time ELL educational assistant placed in the building to help with second language learners. Special Education staffing mirrored that of the ELL program.

While the school received additional federal funding, scores on district-required DIBELS assessments did not show commensurate increases. By description of the DIBELS performance standards, 45% of the kindergarten and first-grade students in the treatment building were at risk, with 27% designated at some risk at the end of the 2001-2002 school year. The district's overall performance showed 31% of kindergarten and first-grade students at risk and 23% at some risk. In 2001, the average Rausch Instructional Unit (RIT) score on the state reading assessment given in third grade was 206 for this building, while the district outscored the state average of 213 with a score of 215. A similar pattern of results was found for the fifth-grade reading assessment.

Design

The data for this study were obtained from a repeated-measures quasi-experimental program evaluation from a school district in the northwest, using FFW products for intervention beginning in the fall of 2002 and ending June 2004. Data were obtained at the pre-, post- and follow-up stages. The interval between pre- and postintervention was 3 months, and between postintervention and follow-up was 6 months. These data were provided as an Excel file copied to a CD-rom without any identifying information for the participant or schools involved in the study. For this investigation, only pre-intervention and follow-up scores were analyzed.

Participants

A school district in the Pacific Northwest piloted Fast ForWord in a single elementary school beginning in the fall of 2002. The purpose of the evaluation was to determine the effectiveness of FFW for improving reading skills in readers scoring below the 25th percentile on measures of oral reading fluency. In that district, all students with reading scores this low were expected to receive some sort of supplemental reading instruction. At the time of the study, electronic records of these interventions were not maintained and therefore were not accessible for those using the extant data set. If records were maintained at the building level, identifying information would not be easily removed. Therefore, it was assumed that the comparison group would have received intervention with an unspecified program during the course of the study.

47

To be eligible for the intervention program, students had to meet one of two criteria: (a) they needed ORF scores at or below the 25th percentile, or (b) they required a teacher's recommendation that they receive training. Once the pool of eligible students was established, parents were invited to an information night about the program. Invitations for participation were sent to 134 families representing 142 students. Of those 134 families, 104 attended information sessions. During the meeting, parents were given a consent document explaining the additional testing students would participate in, the way the results of the program would be reported, and the policies for maintaining eligibility to participate (attendance and behavioral expectations). Interpreters were available for Spanish-speaking parents, and all documents were available in English and Spanish.

Ninety-six students returned written parental consent and were allowed to participate in the intervention. During the evaluation period, nine students moved from the treatment setting, and seven were dismissed from the program due to behavioral concerns or lack of attendance. The number of students completing FFW was 84.

Comparison Group

The initial evaluation design included a matched noncontact group from four district elementary schools that had similar demographic characteristics in terms of poverty rate and school size. The pool of students considered eligible were those scoring below the 25th percentile on DORF. Initially, participants and nonparticipants were

matched based upon gender, grade, special programs participation and ethnicity. There

were a number of limitations to using a matched control, largely due to the multiple

points of agreement needed for the match. Though the comparison group pool from the

four schools seemed large ($n = 424$), the treatment group proved to be unique enough that

quality matches were not possible; therefore, criteria for matching were reduced to age

and test score. Even then, many of the initial pairs were lost due to student movement out

of the district. A second process was used to create a more stable comparison group.

Based upon the instability of the matched group membership and the initially poor

quality of the match, a secondary comparison group was constructed by identifying all

students throughout the district scoring below the 25th percentile on the fall 2002 DORF

measures ($n > 700$). This group had been tracked over time to include only those students

identified in the fall of 2002 who were still in the district in the winter of 2004, when the

study concluded ($n = 212$). At that time, a skill-matched comparison group was

established using grade at pretest and time of year at pretest (fall, winter, and spring) as

additional matching criteria. The data set provided for the comparison group included

only those students with a pretest score at the 25th percentile or below.

Within the FFW treatment group were 12 students admitted to the study based on

teacher recommendation only. These 12 students had ORF scores above the 25th

percentile, the threshold established for participation in the study. The data from these

cases were handled in two ways: (a) Cases were eliminated from any between-group

comparisons to assure that entrance criteria for the study were similar between groups

($n = 72$), and (b) these 12 treatment group members were matched to students with the

highest available percentile rank for the time and grade group in the comparison pool.

The decision to keep the FFW cases in the study was related to the within-FFW question

regarding pre-intervention identifiers that would help identify a good match between

FFW and an individual. Removing these data would have eliminated 13% of the cases

available for that analysis.

Measures

Usually, studies of this sort can pull from the banks of scores on required state

testing to compare the contact and noncontact groups. Unfortunately, this state required

reading testing for students in third and fifth grade only, limiting the use of this data in

measuring program effectiveness between the groups. The extant data available for

matching participant to comparison group were limited to raw ORF scores and grade.

These data were gathered three times a year (fall, winter, spring) throughout the district.

Oral reading fluency scores were converted to percentile rankings based on national

norms published in technical reports (Good, Wallin, Simmons, Kame'enui, & Kaminski,

2002) and the work of Hasbrouck and Tindal (2006). Both the percentile rank and the

normal curve equivalency (NCE) associated with the obtained percentile rank were

recorded on the master data file for use in this evaluation.

DIBELS Subtest of Oral Reading Fluency

Measures of oral reading fluency are accepted as a proxy for comprehension

measures and are highly correlated to results on standardized reading assessments

(Barger, 2003; Fuchs et al., 2001). These measures are also used to determine risk for reading disability in most three-tiered instructional models (Torgesen et al., 2003). Therefore, the use of ORF scores to evaluate treatment over time was determined to be acceptable by the program evaluation team. At the time of this study, there were no studies found that used ORF as a dependent variable for measuring treatment effect.

Each ORF probe requires reading three equivalent reading passages aloud. The assessor tracks errors and records the number of words read correctly in one minute. The median score is recorded in the tracking system. When these scores are used to work within the three-tiered intervention systems, percentile ranks are assigned to the score based upon time of year (fall, winter, or spring) and grade level then recorded as well. The alternate form reliability coefficient for this measure ranged from .90 through .96 ($p < .01$). Concurrent validity ranged from .91 to .96 depending on the passage (Schenck, Walker, Nagel, & Webb, 2005).

With program evaluations in the school setting it can be difficult to obtain assessment data outside of that which is typically collected by the district for all students. Human subjects' protections limit the purposes for which data may be obtained without first providing prior consent options for those involved with the study. The treatment group (FFW) had completed all informed consent requirements and the treatment school was granted permission to gather additional assessment information. Those in the comparison group did not go through any of these processes. The additional data collected by the district as a whole were limited to state test scores and DIBELS data.

Therefore, all other measures related to the subskills of reading comprehension were gathered for the treatment group exclusively.

Standard Language Assessment Batteries

According to automaticity theory and the progress of skill development suggested by the NRP (2001) report, the foundational skill for developing reading comprehension is basic language. Once these basic features of communication are mastered at a level of automaticity, then the next subskill is ready for development. Phonemic awareness (knowing words are made of sounds) is the next skill to be mastered, followed by phonics mastery. The computer games used in FFW are designed to treat all of these subskills and train auditory processing and memory as well. The outcome measure of ORF tests the effects of the programs on developing automaticity with the skill in the progression of reading comprehension development. The relationship between basic language ability at pretest and the impact of FFW on ORF at follow-up is tested to see if knowing language scores prior to intervention is useful for making decisions about strong candidates for FFW.

One of three language tests was administered to the treatment group, depending on the entry date of the student to the program. These measures are used to assess basic understanding of oral language abilities in the areas of receptive and expressive language. The participants in the FFW treatment group took the Oral and Written Language Scales (OWLS) for ages 8-21 (Carrow-Woolfolk, 1995), the Test of Oral Language Development (TOLD: P3; Newcomer & Hammill, 1997) or the Clinical Evaluation of

Language Fluency (CELF3; Semel et al., 1995). The first two groups of students

participating in training were given the OWLS or TOLD: P3 depending on age of the

student (*r-cronbach's alpha* = .70-.91 for internal consistency reliability and a range of

.79 to .91 for criterion-related validity). The last three cohorts of FFW participants were

administered the CELF3 (internal consistency reliability = .75-.91, with concurrent

validity of subtests ranging from .29 to .85).

Ideally, all students would have taken a single language assessment. These tests

were theorized to be equivalent at the construct level; however, no tests confirming or

refuting that assumption were completed. Data included from the TOLD P3 and CELF3

assessments were reported as age-based standard scores ($M = 100$, $SD = 15$). The scores

for OWLS were reported as grade-based scale scores ($M = 10$, SD, 3). For this analysis all

expressive and receptive language scores were converted to a scaled score for equating

manually matching standard score to the equivalent scaled score on a psychometric

conversion table (Pepperdine University, 2005).

Procedures

Two programs from the FFW line of products were available for use during the

treatment phase of the study. All students were involved with Fast ForWord Language

and Fast ForWord Language to Reading. Each program utilizes seven computer games

that adapt to individual student needs at each response. In this way, participants are

incrementally moved through these exercises until they have achieved 80% mastery for

each of the skills the game is designed to teach. Daily practice sessions were 100 minutes long and involved five games each day.

Implementation of Fast ForWord occurred during the last 45 minutes of the school day and extended 90 minutes after school for a total of two hours and 15 minutes daily. The training regimen was established so that students entered the computer lab and complete two 20-minute training sessions followed by a 20-minute nutrition and water break. Training resumed following the break and participants completed the remaining three exercises. Once students had finished all training exercises for the day, recreational activities were provided until the bus arrived to take them home. Teachers in the school agreed to suspend homework assignments for participants until they had completed the program.

Participants trained daily for four to eight weeks depending on when the student met the exit criteria for both FFW programs. Exit criteria used were those provided by SLC. Fast ForWord participants continued with the same reading instruction they had received prior to intervention. Five separate cohorts were trained from January 2002 through May 2004.

Notes taken by the program coordinator informed the staff which students and games to monitor and coach the next day. A certified teacher served as the lead reading coach with the other coaches being educational assistants. All coaches received corporate training. Notes accompanying the data file indicate that an implementation fidelity checklist provided by SLC was used to track issues related to standardization of implementation. However, none of these data were contained in the data file.

The program-evaluation data obtained for the current study did not include records that describe the interventions being used with the comparison group. However, the district had a practice of always providing supplemental instruction for individuals with ORF scores below the 25th percentile and it is assumed that some additional instruction was provided. Data were gathered at three points—pre-intervention, postintervention, and follow up—with follow-up testing conducted six months past treatment.

While this study focuses on the effects of FFW on oral reading fluency (ORF), it also investigates possible student characteristics that assist practitioners with making decisions regarding the use of FFW. The research questions were based on the assumption that FFW would improve oral reading fluency for some individuals and that knowing information related to individual skill patterns is helpful for making placement decisions.

Data Analyses

To approximate the effects of FFW on fluency a repeated-measures 2 x 2 ANOVA was estimated to test if the NCE scores on ORF for the treatment and comparison groups differ when controlling for time. The between-group factor was the group with two levels (treatment and comparison). Time (pre/follow-up) was the within-subjects factor. The analyses estimated the main effects of (a) treatment, and (b) time, and (c) the interaction effect of treatment by time. The interval between pretest and follow-up was six months. The Fast ForWord products used in this evaluation are designed to create

55

a neurological foundation for mapping phonemic and phonetic relationships, as well as to treat language skill deficits that have been proven to interfere with comprehension. The decision to use follow-up data, as opposed to posttest scores, was related to automaticity theory. If FFW did build the foundational skills of fluency, but did not provide instruction with fluency, then time for additional instruction with fluency would be needed. Immediate posttest scores would not provide a valid measure of the effect of FFW on fluency.

The second test of treatment effectiveness considered the effects of FFW on individual risk status. The question explored whether one method of intervention was better than the other in reducing academic risk. To make this determination, participants were categorized in one of two groups: (a) those improving risk status (coded 1), and (b) those who did not change category or changed to a more intensive risk status (coded 0). Improvement in risk status was indicated by a shift in percentile rank, based on decision rules used in three-tier intervention models. Moving from under the 10th percentile to above it was considered an improvement in risk status, as was a change from the 10th to 25th percentile to above the 25th percentile. A chi-square contingency table analysis of treatment/control compared with the characteristic of reduction of risk/no reduction of risk was conducted to validate the overall impact of treatment. Although those reporting effects of FFW have specified the need to build a study that tests pre-intervention skills associated with a significant change in reading performance following treatment (Gillam, Frome-Loeb, et al., 2001; Hook et al., 2001; Temple et al., 2003; Troia and Whitney, 2003; Turner & Pearson, 1999), few have taken the challenge. In answer to

56

this call for research, Question 3 asked which pretest skill levels [Expressive Language or Receptive Language] predicted a change of risk status for FFW participants over time based on oral reading fluency score changes.

In order to recommend FFW as an intervention, educators need to know if there are unique attributes related to language scores that accurately classify student response to treatment. A logistic regression model was estimated to determine if knowing language scores would be useful for making recommendations with regard to FFW treatment. Scores from the expressive and receptive language assessments were the independent variables entered into the logistic regression model to determine if there is a relationship between language scores and changes in risk status (coded 0,1) following treatment for the treatment group exclusively.

CHAPTER IV

RESULTS

The purpose of this study was to determine (a) whether students with oral reading fluency deficits respond to FFW intervention differentially from skill-matched peers who did not receive FFW intervention, and (b) if there were specific characteristics that could be useful for making decisions about using FFW in the school setting. Three questions were explored. The first asked if FFW worked better than other instructional programs for changing oral fluency over time. This information is useful for making decisions about purchasing FFW as a fluency intervention. However, the first question does not determine FFW impact on individual reading risk categories changes, a piece of information that informs placement decisions made by the school-based intervention teams.

It is possible that FFW may not be differentially advantageous for improving reading fluency and yet demonstrate differential effects for reduction in reading risk. Therefore, a second question asked if FFW worked better than other programs for reducing risk of reading difficulties associated with poor oral reading fluency. A final within-treatment group analysis answered the question regarding whether knowing language scores prior to intervention was useful information for decision rules used to match students with FFW as a fluency intervention. The results of the data analyses used to address these questions are presented here.

Fast ForWord and Changes in Oral Reading Fluency NCE Scores

Scientific Learning Corporation claimed FFW products address the foundations of language, phonemic awareness, and phonics, but made no claims regarding effects on fluency, which is used as the basis for identifying changes in risk-status in this study. It was theorized that establishment of the foundational skills of fluency addressed by FFW would cause changes in fluency rate. Importantly, FFW does not provide instruction with oral reading fluency directly. While measures of fluency were obtained pre- and postintervention, and again at follow-up, the research analysis used pre-intervention and follow-up data only. This design served as a statistical control for time that would allow opportunity for instruction on fluency to occur.

The first statistical analysis tests the hypothesis that FFW will change oral reading scores from pre-intervention to follow-up in a statistically significant way, when compared with other programs. To test this theory, NCE scores were the dependent variable in a 2 x 2 repeated mixed measures ANOVA to determine whether there was an FFW effect (treatment versus control) or a time effect (pre-intervention to follow-up) on oral fluency NCE scores.

Table 1 provides summary statistics for both groups, showing that the comparison group has higher gains in pre-intervention to follow-up mean scores (+ 4.62 points) than the FFW group who gained 2.31 points on average. Group membership between treatment and control was different in size for two reasons: (a) missing fields in the data file for the comparison group at follow-up ($n = 3$), and (b) elimination of 12 students from the treatment group for having ORF scores above the 25th percentile threshold.

59

These data for the treatment group were removed in order to reduce perceived advantages

for the FFW group. Figure 3 illustrates the means and corresponding standard deviations

for each group over time. These scores were used in a repeated mixed measure ANOVA.

TABLE 1. Mean of Pre-Intervention and Postintervention
Reading Fluency NCE Scores

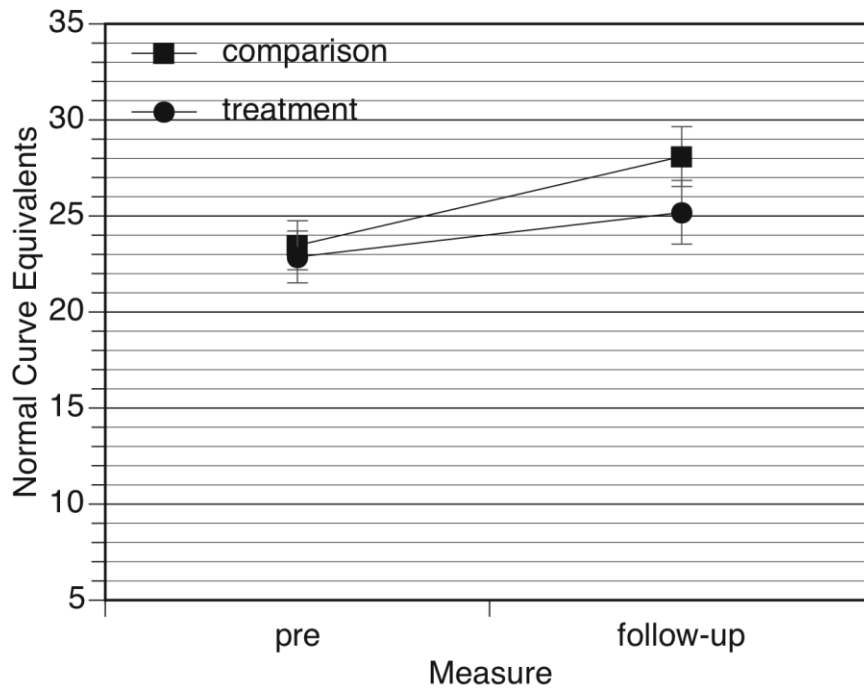| | | NCE scores for ORF | | | |
| | | Pre-intervention | | Follow-up | |
Group	*n*	Mean	*SD*	Mean	*SD*
Treatment	72	22.88	12.31	25.19	13.67
Comparison	81	23.48	10.65	28.10	14.43

FIGURE 3. Treatment and comparison group mean NCE scores and 95%
confidence bands at pre-intervention and follow-up.

The power analysis value for the ANOVA used to answer the first question was

0.98 for an alpha of 0.05 in relationship to time (pre/follow-up). If an effect for time does

exist, the probability of observing the effect is 0.98, while the probability of not observing

the effect is (1-0.98) or 0.02.

In relationship to effects for group membership, the power analysis value was 0.06

at the alpha level of 0.05. In other words, if there is an effect for group membership

(treatment/comparison), the probability of observing the effect is 0.06 and the probability

of missing it is 0.94. A very low value for observed power. If there is an interaction effect

between group membership and score changes, the probability of observing the effect is

0.25 at the alpha level of 0.05, a low value for observed power.

Table 2 shows the repeated mixed measures ANOVA summary statistics. As

indicated, the group-by-time interaction was not significant statistically, while a

statistically significant main effect for time was obtained, $F(1,151) = 15.52$, $p < .001$,

explaining 93% of the variance between scores. Increases in NCE scores for ORF were

TABLE 2. Repeated Mixed Measures ANOVA Summary Table

Source	Sums squares	df	F	Partial eta square
Between intercept	189255.280	1	700.118**	.823
Group	234.914	1	.869	.006
Residual (or error)	40818.191	151		
Time	917.076	1	15.520**	.93
Group by time	100.632	1	1.703	.01
Residual (or error)	8922.394	151		

**$p < 0.01$.

significant for both groups at follow-up, which was six months after the completion of interventions. There was no significant group effect, F (1,151) = .869, n.s., as less than 1% of variance in NCE scores was explained by group membership.

Fast ForWord and Changes in Risk Status

Those responsible for correcting reading difficulties in the school setting ask different questions about program effectiveness that are equally important to consider when deciding to purchase or use an existing intervention program. The school-based intervention team will want to know if the program is better at reducing academic risk than the products currently used. Thus, a contingency analysis examined group membership (treatment/comparison) by change in risk status (reduced risk/no reduced risk). Change in risk was based on comparisons of pre-intervention and follow-up ORF percentile ranks. Using decision rules associated with three-tiered interventions, *reduced risk* was indicated by any reader moving from one risk category to a less severe rating pre-intervention to follow-up (coded 1). All others were in the *no reduced risk* group (coded 0). Those in the treatment group who had initial ORF percentile rankings above the 25th percentile were excluded from this analysis.

With an alpha level of 0.05, the observed power value for the second question, which is analyzed with a chi-square statistic, is 0.090. In other words, if there is a significant difference in the portion of individuals reducing risk status (reduced risk/no reduced risk) between groups (treatment/comparison) at the conclusion of the study, the

62

probability of observing that significant difference is 0.09, while the probability that the observation will be missed is 0.91.

Table 3 shows that the number of participants with a reduction in risk for the FFW group ($n = 18$ of 72, 25%) is fewer proportionally than the comparison group ($n = 31$ of 82, 38%). The chi-square contingency analysis found the difference between groups was significant, $\chi^2 = 3.81$, (1), $p = .05$, favoring the comparison group.

TABLE 3. Percentage of Students Reducing Academic Risk Following Treatment

	Change in risk status					
	No reduction		Reduction		Totals	
Groups	*n*	%	*n*	%	*n*	%
Treatment	54	75	18	25	72	100
Comparison	52	62	32	38	84	100
Totals	106	69	50	31	156	100

As an effect size indicating the strength of association between experimental group and change in risk status, the odds ratio is 0.51. This indicates that the odds of risk reduction for an FFW treatment participant is approximately one half that of a comparison group participant. An alternative report of this effect size is that the comparison group participant is approximately twice as likely as an FFW treatment participant to experience a reduction in risk status. The 95% confidence bounds for this odds ratio is 0.26 (lower bound) to 1.01 (upper bound).

Do Pre-Intervention Language Scores Predict Change?

The first and second questions are between-group comparisons ($n = 156$). The third question is a within-treatment group of data that is used to determine whether knowing about language scores prior to treatment with FFW is useful for making decisions regarding the use of the program for improving oral reading fluency. This is a within-group binary regression model. The logistic regression used to consider the usefulness of language scores in predicting outcomes related to change of risk status was a within-group comparison ($n = 83$). The power analysis value for Question 3 is 0.24 for the alpha value of 0.05. In other words, if there is an effect for language scores on reducing risk status, the probability of observing that effect is 0.24, and the probability the observation would be missed is 0.76.

The third research question focused on pre-intervention language scores as predictors of treatment outcomes. The data file used for this study includes measures of pre-intervention basic oral language skill for those who were in the FFW treatment condition exclusively. A binary logistic regression model was used to determine if expressive and receptive language scores predict treatment outcomes (change in risk status) for the treatment group exclusively.

In this analysis, all FFW members with both receptive language scores and expressive language scores were included ($n = 83$); in this case, one participant with a receptive language score, but no expressive language score, was eliminated from the analysis. The categorical change for these individuals, who had ORF percentile ranks above the 25th percentile at pretest, used the following decision rule: (a) Individuals who

changed percentile ranks to above the 50th percentile were considered to have reduced

risk (coded 1), and (b) those who did not move above the 50th percentile were in the no-

reduction-in-risk category (coded 0).

Table 4 shows the average scores for both basic language assessments are within

the normal range (M = 10, SD =3) regardless of group membership. The differences in

expressive language scores are even smaller, showing an average score of 8.1 (SD = 2.3)

for the reduced-risk group and 8.2 (SD = 2.6) for the no-reduction in risk group. There are

slight differences between groups for receptive language; the mean score of the reduced-

risk group (M = 9.1, SD 2.3) is slightly greater than that of the no-reduction in risk group

(M = 8.4, SD = 2.3). Table 5 summarizes the logistic regression data, showing the

changes in risk category as a function of language scores. The correlation between

expressive and receptive language scores for this sample (r = .65, p < .05) indicates that

TABLE 4. Mean Language Scores for FFW Group by Change in Risk Status

Subtest and risk	Language Scores		
	n	Mean	SD
Expressive Language			
No Reduction in Risk	63	8.2	2.3
Reduction in Risk	20	9.1	2.3
Total	83	8.6	2.3
Receptive Language			
No Reduction in Risk	63	8.2	2.6
Reduction in Risk	20	8.1	2.3
Total	83	8.2	2.6

Note. Standardized mean scores for these subtests were $M = 10$, $SD = 3$.

knowing a person's score on one of the language tests will explain 42.25% of the score on the other test. This size correlation is indicative of a strong relationship between language test scores for this sample ($n = 83$).

TABLE 5. Summary of Logistic Regression Analysis for Language Skills Predicting Change in Risk Status

Predictor	Score range	β	SE	Wald	df	Odds ratio	95% confidence bounds	
							Lower	Upper
Receptive language								
X_1	3.3-13.3	0.13	0.12	1.29	1	1.14	0.91	1.44
Constant		-2.31	1.08	4.61	1	0.01		
Expressive language								
X_1	1.0-14.3	-0.02	0.10	0.04	1	0.98	0.80	1.20
Constant		-0.99	0.86	1.32	1	0.37		

Note. Model fit $\chi^2 = 4.49$, $df = 1$, n.s. Percent of *no risk status change* ($n = 63$) classified correctly = 100%. Percent of *decreased risk status* ($n = 20$) classified correctly = 0%. Percent of overall classification correct = 75.9%.

The summary of results in Table 5 indicates nonsignificant findings for both independent variables. Had these results been significant, holding all other factors constant, a one-point change in receptive language score would increase the odds of reducing risk by a factor of 1.14. When the variable of expressive language increases by one unit, and all other factors remain unchanged, the odds of improved risk following treatment with FFW decreased by a factor of 0.98, meaning a one-unit change in the expressive language variable does not affect the odds of reducing risk. However, the results were found to be nonsignificant for both receptive and expressive pre-intervention language skills, meaning neither were statistically significant predictors of positive pre-to-follow-up change in risk status.

CHAPTER V

DISCUSSION

As measured by the NAEP (NCES, 2009), the low literacy rate in the U.S. has led to the funding of several policy decisions at the national level with regard to reading instruction. Federal funding has been directed toward understanding the issues, and legislation has tied the use of federal funds to specific educational practices in schools. Not only are these funds tied to specific practices; results of test scores must be published annually, and there are sanctions with regard to school funding and structures if the schools do not improve their rate of literacy. With a sense of urgency and little time to wait, schools are turning to research-based commercial programs that the literature has not yet proven to be effective. This is particularly true for students who are nonresponsive to traditional interventions.

Most school districts have approached using FFW without knowing which students will be successfully targeted or whether the program will make a difference in reading skill levels. While program evaluations to date have attempted to discern predictive pre-intervention characteristics that forecast improved reading or language abilities following FFW treatment, the conclusions are too simplistic to be of use in selecting students for participation.

The primary purpose of this study was to investigate the value of FFW as a remediation program for students with significant reading deficits. Theoretically, FFW should have diminished the reading deficits characteristic of high-risk students

performing below expectations. This research tested hypotheses pertaining to the effects of the program on oral reading fluency and changes in risk status. Also, characteristics that may be associated with prereading skills were examined to assist the intervention team regarding the use of FFW.

Results are interpretable in terms of the FFW intervention program compared to the business-as-usual reading intervention programs used in other schools. The theoretical underpinnings of FFW substantiate the hypothesis that students using FFW should experience positive reading growth relative to the effects of other programs. Significant effects were not obtained. Explanations for these null findings will focus on the possible inaccuracies of the FFW program's theoretical model or the possible flaws of this research. Discussion of these results may facilitate some understanding about the FFW programs and the implications for language and reading theorists, researchers, and practitioners.

Interpretation of Results and Alternate Explanations

Three hypotheses were tested in this evaluation. The first was that FFW would increase oral reading fluency. The second was that these increases would result in a reduction of risk status for reading failure associated with percentile ranks and decision rules of intervention systems. The third hypothesis stated that student language skills would predict treatment outcome.

These hypotheses were tested with data from a program evaluation that used a treatment group in a single school and a noncontact comparison group sampled from all

68

other schools in the district. Students were considered eligible for the evaluation if they had an oral reading fluency score that was below the 25th percentile. It was the practice of the district to provide some sort of supplemental reading instruction to students with extremely low reading scores such as these. However, data systems that track supplemental instruction programs were not in use at the time of the study. Therefore, it is assumed that all participants in the comparison had some type of unspecified supplemental intervention in addition to the regular reading instruction program.

Hypothesis 1: Fast ForWord Will Improve Reading Skills

Some students have failed to learn to read even though research-based curricula and instructional practices have been employed (Al Otaiba & Fuchs, 2002; Coyne et al., 2004). Schools are responding to this issue by investing money in programs to help struggling readers, often without evidence of treatment effects in the literature. These data come from a program evaluation designed to determine if the effects of FFW warrant the adoption of the program as a district-wide intervention option. The first question investigated the basic question, does FFW work?

Although a straightforward question, theoretical frameworks for designing reading interventions and assigning individuals to them are important considerations for evaluations intended to answer it. First, selection of the dependent variable should be a contextually embedded decision, according to theories of reading development. In the case of FFW, the program was designed based on the assumptions of automaticity and the

medical explanation of reading skill attainment, which are reviewed here for the purposes of providing background.

Automaticity theory is not a specific reading theory, but rather a framework for understanding the development of any complex skill, including walking, driving or conducting a symphony. Using the reading development theory reported by the National Reading Panel (2001) in the context of automaticity theory yields an explanation of reading development such as this. Learning the complex skill of reading is based on mastering a series of less complex skills (language, phonemic awareness, phonics, fluency, vocabulary). These skills need to be executed without conscious thought so that limited attentional abilities for procedural tasks can be focused on executing the complex skill (reading and comprehending print).

The reading development framework established in the NRP (2001) report is not the only perspective from which reading interventions are being developed; medical practitioners are contributing knowledge to the design of reading programs and interventions as well. The medical view of reading, which still considers automaticity theory in explaining reading development, details the constructs in the NRP report in a more complex fashion. Instead of assuming that automaticity with a skill leads to automatic use of that skill during a reading task, the systems responsible for attention/arousal, working memory and executive function are implicated for coordinating skill application with the model.

For example, automatic use of phonemic awareness is dependent on executive functioning and sensory-processing skills that support the neural activation patterns

associated with hearing sounds in words, holding that knowledge in working memory and then applying the skill to the task at hand. Designing interventions by the medical model of reading requires strengthening of one or all of these processes (working memory, attentional/arousal coordination of sensory input and executive function) in addition to directly instructing the reading skill.

The dependent variable selection for a program evaluation of FFW should not only be based on assumptions of automaticity, but also include measurement of the processing systems. Finally, the dependent variable should be sensitive to change, based on the intervention content, which was not the case for this evaluation. The content addressed by FFW comprises the underlying subskills associated with developing fluency. These skills include auditory memory, sensory processing of auditory stimuli, as well as phonemic awareness and phonics.

It was assumed that treating the underlying causes associated with poor fluency would provide the necessary foundation for fluency skill to improve. It was further assumed that the treatment school would be providing daily instruction in fluency during the intervention and into the follow-up period, as that was the expected practice for all schools. Therefore, ORF was selected as the outcome measure. This was a fairly major design flaw for two reasons: (a) FFW does not directly treat fluency, and (b) there was no assurance that fluency instruction occurred as expected in the general reading program. Therefore, the measure of ORF was too distally related to the constructs addressed by the intervention. The CBM probes measuring automaticity with phonics application would have likely yielded more reliable results for discerning the impact of the FFW programs.

71

Results From Question 1

The first hypothesis, that FFW works better than the other unspecific programs improving fluency, was evaluated using a repeated mixed measures ANOVA. The model compared group membership (treatment/comparison) to changes in ORF NCE scores (pre-intervention to follow-up), controlling for time. Both groups experienced an effect for time, and no interaction effect was found, making these results interpretable. There was no group effect. In the context of this study, FFW did not work better than other programs with regard to improved oral reading fluency. This finding alone adds to the literature regarding impact of FFW, as no other study has reported oral fluency as a dependent variable.

Interpretation of Results From Question 1

A critique of the study design will serve to interpret the results from Question 1 in light of several unresolved threats to validity. Already considered were construct validity issues associated with the alignment of the dependent variable to the theoretical frameworks involved with the intervention design. Issues related to the measure itself will be deliberated later in this discussion. The organization of the discussion includes consideration of validity as related to (a) the sampling plan, (b) participants, (c) implementation, (d) measures, and (e) statistical conclusion validity.

Sampling Plan

As with many program evaluations conducted in the school setting, there are few, if any, random-assignment strategies employed for determining membership in the treatment or control conditions for this study. Students were selected from a single school for treatment, and anyone who qualified or was recommended was enrolled in the program after parents provided written permission.

The likelihood that this sample was representative of all student populations across the district was diminished by two conditions: (a) The requirement of parental permission necessitates a degree of parent involvement that may not be found in the comparison group; and (b) this school had the distinction of having more students who live in poverty, speak English as a second language, and who possess lower test score averages than all but one other school in the district. The original evaluation model used students from schools with the most similar contextual considerations for creating the noncontact comparison group, but mobility and unique characteristics associated with the treatment participants interfered with this attempt to control for school context by design. Therefore, validity concerns associated with selection bias cannot be ruled out in this study.

The comparison group was formed after the fact by using data from all other schools in the district, matching grade level and percentile rank on ORF to a member in the treatment condition. Variables in other primary areas of aptitude (basic language, phonemic awareness, phonics application, vocabulary, and comprehension) were not

included in the matching process. This could explain differences in the outcomes that are unaccounted for currently.

Participants

Although the unit of study was indicated to be the individual, in actuality it may have been the school. Students from the treatment school had poor performance on state and local measures of reading, despite having a number of additional staff to support the reading development process. Perhaps the school was not providing sound reading instruction to the entire population and the opportunity to learn was different for the treatment participant than it was in the comparison schools. This could mean that deficits in all reading skills were greater for the treatment group than for the comparison group due to school-based confounds, and results from this study do not evaluate FFW at all.

Controlling for the type of reading intervention used with the comparison group in other schools was not possible. This is an important limitation to the study, especially in light of the intervention trends at the time, which included fluency development practice as a supplemental program for the regular reading curriculum. It is possible that the comparison group represents students that have been exposed to more intensive instruction with fluency or other reading skills, and that they continue to receive intervention related to reading skill for longer or shorter periods of time. If that is the case, then FFW participants may not have improved fluency at the same rate as their peers due to differences in interventions. Reading skills develop predictably, with automaticity of the first building block (basic language) supporting the development of

the next skill set (phonemic awareness). To go from developing a skill to being able to perform it automatically requires targeted instruction and multiple opportunities to practice the skill.

If the students coming into the FFW treatment condition had attained automaticity with phonemic awareness skills, but not with phonics, they may have benefited from the phonics application games, but would likely need more practice with phonics skills in order to push those skills to an automatic level. If the students were fortunate enough to have phonemic awareness and phonics application facility, then perhaps they would have been ready to develop oral fluency. In either case, FFW would not be the appropriate intervention, as it does not teach word reading fluency within the program. Failing to perform diagnostic screening prior to placement in the study may have led interventionists to place students into FFW who were better suited for a fluency intervention alone.

Score differences are not likely to be related to maturation threats to validity, as an aged-matched control group was used to reduce such concern. However, this assumes that age is an indicator of reading skill proficiency and does not account for differences in skills based on the expectation associated with those skills. For example, being unable to use phonemic awareness to assist with reading is not unusual for first graders, but highly unusual for fourth graders. If the untested skills (phonics and phonemic awareness) were not at grade-level expectations for members of the treatment or comparison group, maturation-related concerns cannot be ruled out in terms of how the individual would

likely improve the fluency rate. This application of thought regarding maturation would essentially leave the selection maturation interaction without an investigative control.

The treatment group included only participants who completed treatment and did not include scores from individuals who moved or exited the program due to lack of attendance or behavioral issues during training. Generalizations about the treatment group must consider this information, as those who exited the study are likely to have unique characteristics compared to those who remained. This is one of several confounding threats to validity with regard to participants.

Implementation

Implementation of FFW occurred after school for 90 minutes each day, until the individual exited the program based on meeting SLC criteria provided with the materials. Those conducting the program evaluation reported the use of an implementation protocol checklist to assure proper implementation of FFW. These data were not presented as part of the data set. The computer programs themselves have a fidelity indicator that tracks student progress over time, in relation to minutes trained per day. These data were not provided either. It is possible that FFW was not implemented as intended and that was the reason there were no group effects for fluency. Or perhaps the circumstance of training in the after-school environment made participants resentful of the program, resulting in a lack of effort with regard to training. Equally plausible is the conclusion that the program was used properly and it does not work as a fluency intervention. Although treatment diffusions are not likely, as the program design of FFW is fairly unique, issues related to

76

fluency interventions would be important considerations here as well. It is not known

whether either group received any formal intervention for fluency during the evaluation

period, a major flaw in the study design.

Measures

Several validity concerns related to the measure used for determining whether

FFW works may sufficiently limit the conclusions that can be drawn from the study.

Already mentioned were concerns related to construct validity and program match,

including the distal relationship between the skills treated by the intervention and the use

of oral fluency. These concerns are significant with regard to the dependent variable.

The measure of ORF is considered valid and reliable for measuring fluency

(Hasbrouck & Tindal, 2006), and measures of fluency are found to a reliable proxy for

reading comprehension (Fuchs et al., 2001; Good et al., 2001). However, this was the

lone criterion on which the study participants were compared. Technical adequacy of the

measures is of less concern than the validity issues associated with how the test was used.

Participants were selected based on very low scores for ORF (25th percentile or lower).

The first concern is that any changes in reading score are going to be related to the

tendency of extreme scores to gravitate toward the mean, regardless of intervention. If

this is true, then differences in scores are not attributable to program influences. In this

case, the comparison group just moved scores closer to the mean than the treatment group

as a matter of luck. Concerns associated with the testing protocols may interfere with the

validity of assessment outcomes as well. Testing protocol requires repeated testing of the

same skill with different versions of reading passages, a feature that could possibly threaten validity. However, testing of ORF is common for all students outside of the study situation, and the passages read aloud change each time.

The DORF measures include several forms of reading protocols for each grade level to assure that the reading passage presented is novel to the learner. It is difficult to assure passage equivalency, as familiarity with a topic may influence the rate of reading, and these factors cannot be controlled for. The ORF probe may actually be measuring student background experiences, rather than reading rate, making differences in this study attributable to differences in background experiences rather than differences in the program. To overcome concerns associated with the equivalency of the probes and biases associated with prior learning, assessors used three reading passages and recorded the median score. The interaction between student background and the probes presented is still somewhat of a concern in this study. There were a great number of LEP students in the treatment group who may have been unfamiliar with the topics presented.

Statistical Conclusion Validity

Concerns about statistical conclusion validity relative to Question 1 include (a) the decision to insert a six-month delay between pre- and follow-up testing of ORF, and (b) using NCE scores as opposed to raw scores in the analysis.

Whenever a delay between the end of a program and the measurement of the dependent variable exists there is danger that other circumstances may influence the results. In the case of this study, 6 months elapsed between treatment and follow-up

measures, and the intervention lasted only two months. Outside influences were not controlled for during the four-month time gap. Yet the rationale for a delayed-measure model was theoretically supported, as FFW does not directly teach fluency skills and the delay was an attempt to control for that fact by design. This aspect of the study design seriously threatens the validity of conclusions one can draw with regard to the effects of FFW on fluency. At best, the study measures an unknown variable related to other instruction following treatment.

The use of standardized scores for representing changes in oral fluency was deemed preferable to using raw scores, as a 5-point change in raw score at first grade is not interpreted the same way as a 5-point change in the fifth grade. Scores are interpreted based on grade of the learner and time of year. An outstanding reading rate for a first-grade student in the fall would be 45 *wcpm.* The same score during the same testing period for the fourth grader is considered extremely low. Yet, the conversion of scores may interfere with the precision of statements that can be made about these scores when interpreting the results.

Fast ForWord program completion does not appear to improve reading fluency for students from the school used in this study. This finding cannot be explained away by issues related to the assumptions for the ANVOA estimate, as all conditions for normal distribution of the dependent variable, sphericity, and equal covariance were met, requiring no further statistical corrections for the model.

Analysis of the power value for the ANOVA was unbalanced with respect to finding effects for time and for groups. If there was an effect for time, the probability of

observing that effect was 0.98 at an alpha of 0.05. These are expected findings, as time should improve scores for both groups regardless of intervention type. This ANOVA estimated a significant effect for time. If an effect for group had been obtained, the probability of observing the effects would have been 0.02. Given these values, the results from this study are to be cautiously interpreted.

Implications

Evaluation of the treatment effect of FFW with regard to fluency indicates that in this particular school, with the reading program as it was being implemented during the regular school day for the years 2002-2004, FFW most likely will not change fluency scores in a significant way. FFW purchasing decisions are being made despite scarce funds for purchase, multiple demands on instructional time, lack of understanding about which students are likely to benefit from the program, and limitation of staffing capacity regarding time for use of the program during the school day. Given these constraints and the ANOVA data, it may not be wise to purchase FFW as an oral reading fluency intervention.

Nonetheless, the cost of providing intensive instruction that treatment resistors seem to warrant is going to be high anyway, and schools have institutional and ethical obligations to maximize each student's progress. Tests used to evaluate the effectiveness of FFW in this evaluation are not sensitive enough to capture program impact on other reading skills. So while FFW may not be a useful intervention for ORF as a whole, the instructional delivery model may hold promise for future use, as it is standardized and

individually responsive to student need, eliminates human implementation error, and increases the number of students who receive intensive instruction based on the availability of computers. This study could not evaluate the impact of FFW on treatment resistors, as there was no history of individual intervention included with the data set. If it were found that those with a history of being nonresponsive to alphabet treatments were positively impacted by FFW, then the purchase of the program may be warranted.

Hypothesis 2: FFW Will Help Reduce Reading Risk

The impact of federal legislation relating to new identification processes for specific learning disabilities has been far reaching. The IDEIA (2004) reauthorization responded to complaints that the system for identifying learning disabilities required students to fail before getting extra help with reading. The latest authorization of IDEIA allows schools to track individual student response to intervention as a method for identifying the disability. Data-based decision-making teams of reading and learning experts in the school evaluate universal screening data three times a year in order to identify students not on track for the requisite reading, writing, math or behavior improvements.

Students are sorted into three groups: (a) those needing very little additional support to develop the skill, (b) those needing strategic support in the form of supplemental instruction or a decrease in group size to allow more opportunities to respond, and (c) those in need of intensive intervention. Percentile ranks associated with scores on the screening are the primary identifiers for risk association. In the case of

reading, being below the 25th percentile on a screening measure is associated with a need for strategic or intensive support. Placing results of FFW in the context a school intervention team would use to assign treatment adds an additional dimension for understanding the effects of the program.

Results From Question 2

The second hypothesis was that FFW works to change risk status, associated with ORF percentile rankings in three-tiered instructional models, better than the unspecified programs used with the comparison group. A chi-square contingency analysis of group by risk change failed to confirm this hypothesis. In fact, the odds of reducing risk from pre-intervention to follow-up were 1.41, favoring the noncontact comparison group.

Interpretation of Results From Question 2

In this study the process for assigning risk status may confound interpretation of results. Determining whether an individual reduced risk by the end of the study cycle was accomplished by very strictly applying decision rules related to three-tiered instructional models. To do this, I converted raw scores in oral reading fluency to percentile ranks by using fairly well-substantiated norms for this CBM (Good et al., 2002; Hasbrouck & Tindal, 2006). However, there were no allowances for the possibility of measurement error when recording the association between the raw score and the percentile rank. Those near the 10th- or 25th-percentile thresholds may have been misclassified due to unaccounted-for measurement error; i.e., changes in risk were not actually changes at all.

These cut points are not arbitrary but neither are they precise. If there were measurement errors associated more with one group than the other, then evaluation of these data would yield inconclusive findings.

The chi-square analysis obtained significant differences for changes in risk category between groups (treatment/comparison), favoring the comparison group. Those in the comparison group were 1.5 times more likely to change risk category than the treatment group at follow-up. Not only was FFW not effective for reducing risk overall, students in the treatment group might have been much better off if they had been involved with the nonspecified interventions. Additional consideration of the context of the schools from which the comparison group came would help explain whether differences in treatment outcome were related to FFW or to the reading programs in the schools overall.

Implications

Reading fluency is related to risk status, but improvement in fluency may not actually reduce risk for reading failure. There are other considerations that must be measured (grade of the student, history of intervention, etc.) before determining if, in reality, the student is out of the risk zone. Further, a change from needing intensive support to needing strategic support in no way indicates that the individual has been cured of all reading problems. It is very easy to confuse improved fluency with decreased risk, which may or may not be an accurate conclusion. Construct validity matches between the

dependent variable and the associated interpretation of the variable (risk change) confound the ability of the evaluator to make decisions regarding purchasing.

Those who use data-based decision-making models in schools are trained to view these data in a very specific context. For that group of individuals, findings of reduced risk would only mean that something was working in the instructional plan. Then decisions for the individual would follow and likely involve more diagnostic testing to determine which skill needs attention next.

Hypothesis 3: Specific Characteristics Predict Treatment Outcomes

In the literature there is a collective call for research centering on the development of a profile of test scores and age or grade considerations that could place students likely to respond into this intervention. These recommendations for more research come from (a) the field of education (Hook et al., 2001; Troia and Whitney, 2003); (b) the medical treatment fields such as speech and language pathology (Gillam, Frome-Loeb, et al., 2001; Turner & Pearson, 1999); and (c) those studying dyslexia from a neurobiological perspective (Temple et al., 2003).

The school-based intervention team would be interested in knowing if characteristics associated with individual learning profiles are useful for making placement decisions with regard to the use of the FFW curriculum. Within the context of this study, the question would relate to the use of FFW for a fluency intervention.

Results From Question 3

The third hypothesis, that specific characteristics predict reduction in risk, was not supported with the language variables tested. In other words, knowing something about individual language ability prior to intervention is not helpful for making placement decisions with regard to the use of FFW for reducing reading risk. The odds ratios associated with each language subtest were close to one, nonsignificant, and oppositely related. The odds of exiting a risk category were more favorable based on receptive language scores (odds ratio = 1.14), while there was a negative relationship between exiting a risk category and expressive language scores (odds ratio = .98).

Interpretation of Results From Question 3

These limitations could be related to internal validity issues with regard to (a) the group evaluated, (b) the measures used, and (c) the lack of power associated with the statistical model.

Group

A substantially limiting factor in designing the logistic regression model is related to the lack of data from the comparison group for the independent variables tested. The design of the study precluded obtaining evaluation data outside of the information banks encompassing all students in the district, DIBELS and state testing results. The question written to investigate the hypothesis that knowing pre-intervention skill level would be useful for making treatment decisions required that there be a comparison of some sort

between groups. Had there been independent variable data available from the comparison group for language, as there were for the treatment group, then the two groups would have been treatment versus comparison. However, a different approach was necessary, and the regression was estimated with data from the treatment group only. The treatment group was separated into two groups (reduction in risk/no reduction in risk). Issues relating to the accuracy of these groupings and contextual definition for risk change have been previously discussed. Not discussed, though, were issues related to group size that directly impact the logistic regression.

The groups are not only small in terms of providing robust conclusions, but also sufficiently disproportionate from one another, with the change-in-risk group containing only 18 members and the no-change-in-risk group having 65 cases. These differences in group size at the percentage level are not the concern. Rather, the concern is the groups are too small to make statistically sound assumptions with regard to the results of the regression model tested, a concern supported by the power value associated with the regression model (0.24, alpha = 0.05).

Measures

The data provided for the current investigation had expressive and receptive language scores for the treatment group. These data came from three different standardized language batteries. Although scores were recorded for these subtests on two different scales (standard scores and scale scores), the act of converting them to one scale was fairly simple and necessary if these data were to be useful in the regression model.

Nonetheless, concurrent validity between the three measures used was not established through a linking study or other means of association, making the use of these data suspect. Had the model been significant, the odds ratio associated with these scores would have been very difficult to interpret accurately; ideally, all students would have had language scores for a single assessment. Issues related to statistical assumptions must be considered when determining the implications of results related to Question 3.

Statistical Conclusion Validity

The use of a logistic regression model was appropriate for the research question focusing on using language as a predictor of change coded 0,1. The model provided estimates of the odds that positive changed would occur given incremental increases in the language measures. However, use of this model with both language variables may be inappropriate. The relatively high correlation between the independent variables ($r = .67$) indicated that the use of both measures may be uninformative, as the effect of one is likely to be related to the effect of the other. The Hosmer and Lemeshow goodness-of-fit test indicated that the logistic model was not a good fit, and nonsignificant findings with the Wald statistic for each parameter indicated that knowing receptive and/or expressive language scores is not useful information for making decisions about the use of FFW to change reading risk. In addition, the estimated logistic regression models inaccurately classified all students who did experience a decrease in risk. This poor classification, however, may be attributable to the relatively large proportion of students who did not

change compared to those who experienced a decreased risk (76% and 24%, respectively).

Implications

The findings in the literature review regarding the effects of FFW would lead one to think there may be utility in knowing language scores prior to assigning students to FFW. These larger categorical scores of expressive and receptive language have been useful in examining results for students in FFW treatment versus noncontact treatment when the dependent measure is a reading assessment, but only when comparing students with language scores below the 25th percentile for the broad language measure (Troia, 2004; Troia & Whitney, 2003). The relationship between reading fluency and language scores has not been tested in any other study. However, there have been effects for phonics skills (Cohen et al., 2005) and phonemic awareness (Hook et al., 2001; Pokorni et al., 2004) in groups that were admitted to studies due to poor language skill. Whether treating these language deficits has a cross-generalizing effect on reading cannot be directly assessed in this study, as the reading measure is not sufficiently robust to confirm or disallow this hypothesis. However, a significant body of research supports treatment of language skills as an intervention for poor reading (Shaywitz et al., 1999; Vellutino et al., 2004), especially for students in grades K-3 (Al Otaiba & Fuchs, 2002; Coyne et al., 2004; Denton et al., 2003).

While it would be accurate to conclude that knowing language abilities is not useful for making judgments regarding change in risk status, concluding that pre-

intervention language skills are informative for predicting changes in ORF is not valid. The regression model did not include fluency measures as a dependent variable. The status of changed risk was used instead. The categorical dependent variable was change in percentile rank associated with ORF scores. It is possible that two individuals made the same gain on NCE points from pre-intervention to follow-up, but one changed risk category and the other did not. The dependent variable only represents changes suggesting to a reading specialist team that it may be valuable to review a student's reading intervention plan. A more precise conclusion regarding the logistic regression analysis would be that knowing language scores is not helpful for predicting whether reading risk status will be reduced after intervention.

Finally, because the language variables were uninformative, the model predicted all students would not actually change. This, of course, does not imply that there is no value in identifying important predictors of change. Rather, language skills is simply not one of the valuable predictors. Both Troia (2004) and Borman et al. (2009) suggest other possible valuable pre-intervention predictors, noting those that can be reliably obtained by specialists and those which may be valuable, but not readily available.

Relevance of Findings and Implications for Practice

Data-Based Decision-Making

School based reading intervention systems have been created based on federal guidelines that require the use of data-based decision-making models for academic and behavioral intervention. If schools want to access federal funding associated with special

education, they must adopt a data-based decision model aligning curriculum and resources with student needs. Data are used to monitor intervention effectiveness as frequently as twice a week, necessitating a quick, reliable measure sensitive to growth. Curriculum-based measures (automaticity probes) are used as for these purposes.

Theoretically, the measurement of automaticity with each skill indicates the true automatic level with which the individual can execute the skill during the reading process. This one-to-one correspondence is not necessarily accurate, as it is possible to pass the state reading test and still be dysfluent with passage reading by standards of automaticity. However, this is the simplified theory of reading skill development used by most school systems for designing interventions and measuring treatment effect.

In this study, the universal screening data (ORF) used to determine the need for an intervention were not sufficient to measure program impact. Had the intervention team used the data to determine that a student may need diagnostic testing, the data would have been useful. Instead, the screening measure was used as the diagnostic measure and the match between FFW and student was not strong. Implications for this design are twofold. First, the intervention teams need to use screening data as intended, which implies a need for training in this regard. Had they tested the individuals for automaticity with phonics, and phonemic awareness following the review of screening data, a more accurate progress-monitoring CBM could have been selected to test effectiveness of the program for the individual.

Secondly, it may be that automaticity measures are not necessarily measuring everything we originally hoped they would. Implied in the design of the CBM probes is a

theory that automatic retrieval of one skill is indicative of having the processing skills needed to apply the skill during reading. This may or may not be a correct assumption. Without burdening the system with more testing, a simplified approach to measuring effectiveness of an intervention may need to rely on at least two measures of automaticity for detecting next levels of intervention and for reporting treatment effectiveness.

In the practical sense, and relative to automaticity theory, the implications of the results from this study are clear: The use of ORF to measure treatment effect of FFW is not likely to yield meaningful results regarding true treatment impact, and, it is not likely that oral reading fluency will change after FFW unless some other intervention is used as a follow-up. In the more global sense, there are implications for the use of CBM in differing contexts. Currently these data are used as a tool for monitoring treatment effectiveness, as a tool for identifying topics to be covered in treatment, as a diagnostic tool, and as a tool to flag a likely need for more help. Considerations of these contexts were ignored in the design of the current evaluation of FFW products.

The Interaction of Medical and Educational Research

Fast ForWord products were designed by members of the medical profession who used clinically significant standards to evaluate evidence of treatment effectiveness. Clinical significance is defined by the amount of individual growth experienced by the individual pre- to posttesting. The use of three-tiered instructional models in schools mirrors some of the features associated with the practice of medicine: screen everyone, watch for flags of poor reading health, provide treatment, and then monitor and adjust

based on symptom changes. In terms of measurement, the uses of CBM have made this process relatively painless in schools; however, association of treatment with symptoms has yet to be perfected with regard to intervention selection.

When schools first implemented the model used for three-tiered instruction, interventions were associated with the screening measure (ORF or other rapid-naming probes) and then nearly any intervention was provided. Most often there was an assumption that fluency was low and that direct practice with passage reading would therefore correct the problem. This was true for some and not others. As time has progressed, diagnostic assessments with CBM have connected troubled readers with more accurately matched intervention programs.

The IDEIA legislation, which allows identification of learning disabilities based on an individual's response to intervention, has changed the way services are delivered to all students and eliminated the odds of students being labeled learning disabled when they are truly instructionally disabled, meaning they lacked instruction that would help them master the skill. Yet, we are left with a number of students who do not respond to instruction and become perplexed with regard to treatment. Perhaps this is where medical models of intervention plans become necessary. Not too long ago, medical models of reading were merely interesting to consider, but not often used for creating intervention programs in the field of reading. For students who are nonresponsive to well-designed interventions, medical models of reading are appropriate for informing instructional choice.

Although this is a far-reaching implication given the study conducted, the fact that the program emerged from the medical science for explaining reading, and not a traditional educational model, means that there are practical implications for local use of the product. It will be important for leaders of reading research in both fields (education and medicine) to work together in developing these intensive intervention systems for the most challenged readers. Otherwise, schools may prescribe intensive, resource-consuming treatment for a problem that may require only a few stitches. This current study used such an approach to implementing FFW. Everyone who had a fever received intensive care when a few aspirin and a change in instructional diet may have been the only treatment needed. The implication for schools currently using these products would be to make sure that the students' intervention history and instructional needs warrant intensive treatment with FFW.

Recommendations

The recommendation for the use and adoption of FFW as an intervention system must consider more than the findings from any single study. This study, particularly with regard to the number of substantial threats to validity embodied in the design, should not be considered as a sole source for making adoption decisions. The complexity of reading skill development, the systems schools used to implement interventions, and the decision-making strategies used for placing students in these intervention systems make recommendations more complex than a simple declaration to use or not use this program. Adding to these typical contextual clarifications are foundational FFW research matters

and the newness of these theories as applications for treating reading challenges. With these influences in mind, however, specific recommendations regarding the use of FFW can be made.

Previously mentioned were recommendations related to directly associating the need for oral reading fluency intervention with the use of FFW. There does not appear to be a direct effect for FFW on reading fluency. Therefore, purchase and/or use of FFW as an exclusive treatment for fluency skill deficits is not recommended. Further, for those with expressive language or receptive language in the normal range of development, FFW may not assist with reading fluency improvement.

The results of this study did lead to recommendations that are useful to those already using FFW. The first of these would be to use something other than ORF to place students into the program. If a fluency intervention is all that is needed for the individual, there are several excellent programs available that are far less intensive in terms of time and other resources. Consider intervention history of the individual prior to using the FFW programs. If other phonics-based reading programs have failed and intervention with phonemic awareness has been limited, consider the use of FFW. Test down the skill ladder (phonics, phonemic awareness) and determine which measure would be most useful for monitoring progress as the intervention is taking place. Once the program is completed, use a phonics-based reading system to reteach the phonics skills, and work on improving automaticity in these skills, prior to moving on to oral fluency intervention.

For those studying FFW, a recommendation for future research in the form of school-based program evaluations will be succinctly outlined later in this chapter, but

here are topical considerations for other types of research design: (a) decision rules for

placing students into FFW, (b) CBM measures not currently used in schools that would

be useful for placement decisions with regard to FFW, and (c) the best outcome measure

for use in progress monitoring with FFW.

Limitations

Several limitations to interpreting results and reaching valid conclusions regarding

the use of FFW as a tier-two or tier-three intervention are worthy of mention. Problems of

internal-validity-related quasi-experimental design and numerous possible uncontrolled

variables interfere with making generalizations based on the findings reported in this

study. These limitations are examined in relationship to (a) participants and the sampling

plan, and (b) concerns about the data available for use.

Participants and Sampling Plan

The sampling plan for this study was flawed to the point that the evaluation could

only be classified as quasi-experimental. Although there were score criteria established

for membership in the study (ORF below the 25th percentile), participants in the

treatment group came from a single school while the comparison group was from all

other schools in the district ($n = 16$). Parent permission for participation was not required

for the control group but was for the treatment condition. Use of this method for creating

a treatment group raises concern, as random selection of participants within a qualifying

pool is not possible. Even if a single school could be considered a representative sample

of all schools, the requirement of permission adds an element of self-selection to the construction of the treatment group. It is possible that the school would no longer represent the population of students who qualify for the study.

Even more limitations exist as the lack of a true comparison group jeopardizes the internal validity of the study. The comparison group was formed after the fact, using data from all other schools in the district, and matching grade level and percentile rank on ORF to a member in the treatment condition. Variables in other primary areas of aptitude (basic language, phonemic awareness, phonics application, vocabulary, and comprehension) were not included in the matching process. This could explain differences in the outcomes that are currently unaccounted for but could be easily overcome in the design of future studies.

Substantially threatening the validity of this study is the fact that there are no controls for other interventions in place by design, and that this remains true for the treatment group as well during the period of time between program completion and follow-up assessment. Conclusions about the effects of FFW on ORF will need to be considered in light of this major design flaw.

<center>Measures</center>

In addition to the weak design, oral reading fluency was the only reading outcome measured for both groups. One measure of automaticity is not likely to give a robust picture of program effectiveness. Further, construct validity is questionable in terms of the relationship between what was taught and what was measured, as fluency was not a

<center>96</center>

part of the FFW intervention. Oral reading fluency is a skill that is distally related to the intended treatment outcome: language, phonemic awareness and phonics. It would have been more beneficial to test phonics application with CBM, as the program begins training phonics at the end of the series. Variables were selected knowing these considerations, but an assumption that fluency development was a part of the general curriculum, and therefore the FFW treatment would be provided for fluency instruction, led to the selection of ORF. Rather than measuring treatment effect, ORF may have been measuring implementation of the general curriculum, or lack thereof. Additional concerns regarding the use of ORF are related to the two contexts in which the scores were used. One was to measure change in oral fluency, the other to determine change in reading risk status.

The second application of ORF scores in statistical models was to use the ORF percentile rank to determine level of risk related to reading skill. This is an application of intervention decision rules used in the instructional models supporting RTI. Measurement error associated with the ORF raw score was not considered when mapping percentile rank to raw score, leaving risk status assignment as a potentially inaccurate accounting of risk. The context of these risk associations is very specific and simply means that the reading intervention team monitoring student progress needs to do some follow-up assessment to determine the next steps in an intervention. It does not mean a student is cured of reading issues. Validity of results will be jeopardized if these contexts are not considered. Further, a change in risk is not necessarily the same as a change in fluency. Statements using risk change data to describe outcomes must be carefully crafted.

97

The scores from the independent variables of expressive and receptive language came from three separate measures. These measures were easily converted to a similar scale with a conversion chart, but the constructs of expressive or receptive language between the three tests were not tested for concurrent validity. In other words, if there was an effect for language, its significance would need to be regarded with caution.

Summary of Limitations

Despite these weaknesses, the results presented here are still interpretable, if due consideration of the limitations is given. The hypothesized effect of FFW was found null in all cases. The meaning of these results is interpreted in relationship to theory and the literature that led to the hypothesis. Lack of effect was likely related to poor alignment between measurement constructs and content of the program assessed.

Suggestions for Future Research

Understanding the effect of FFW in view of the progression of reading skills will inform decisions about future research. To review, the NRP (2001) report describes the steps to reading as incrementally associated with age. In terms of developing reading skill, the focus of early childhood is on learning to use and understand language. Once a student is school aged, the instructional focus evolves to learning that words are made up of sounds (phonemic awareness) and that these sounds are represented by letters that can then be used to make words (phonetic understanding). Once these understandings are in place, typically at the end of first grade or the beginning of second grade, curriculum

focuses on learning to read and recognize words fluently. All the while, vocabulary and comprehension skills are taught through discussion and application of meaning (in ways that are not print dependent). Fast ForWord products aim to improve non-print-specific skills up until the end of training with FFW to Reading, the second program that all study participants completed in this study, during which time phonics instruction occurs.

Knowing the skill level of individuals prior to treatment with FFW will be useful for making placement decisions. As attested in the literature review, students who responded to FFW with gains in phonemic awareness were recommended for treatment based on language skill (Cohen et al., 2005); this was true for those with gains in phonics as well (Troia, 2004). Phonemic awareness is considered a part of the expressive language skill, and has been indirectly tested as a criterion for entrance into a study, but not as an isolated skill. Using low phonics skills as a screening criterion has not been tried either. This study attempted to discern if knowing a student's language skills is useful for making decisions regarding the use of FFW as an intervention for fluency. Level of language skill was not useful in the context of improving reading fluency for those scoring below the 25th percentile on an ORF probe.

This is the first study to use fluency as screening criteria, and as the dependent variable as well. Poor fluency at pretest was not associated with improved fluency at follow-up. Studies that used vocabulary tests as screening criteria for treatment were not found in the literature, but several studies reported using performance on state reading tests as criteria for participation in FFW. If simplistic decision rules were to evolve from these studies, then a few recommendations for future research would follow.

Understanding whether information regarding phonemic awareness, phonics or language skill is useful for making FFW placement decisions will require research.

First, studies in the school setting need to be conducted using poor phonics as a participation criterion; the same is true for studies that use poor phonemic awareness or poor vocabulary as participation criteria. However, this recommendation must be tempered with consideration for the likelihood that there are multiple symptoms that could be associated with treatment recommendations for the use of FFW. Rather than test these skills independently, an evaluation of FFW that included pre- to post-CBM measures of each of those skills would be recommended. Curriculum-based measures are regularly used as part of the intervention systems within schools, and after this evaluation of FFW was begun, data systems have evolved that track those scores at a district level.

Using CBM gives the advantage of being able to use noncontact comparison groups for program evaluation, reducing costs and permissions required to gather data. Issues related to tracking other interventions for the control group will not be resolved in this model. Regression models of these data would begin to provide an understanding of the characteristics associated with improved reading skill following treatment with FFW. It is recommended that the dependent variable in these cases be a phonics application CBM. This is in contrast to the current study, which used a CBM distally related to the constructs addressed in the intervention. Only at the end of the FFW program does phonics begin to be treated directly with print-based exercises. Using CBM of phonics skill as the dependent variable will likely increase the possibility of observing positive program effects.

Age of student at intervention with FFW may still be important. However, FFW is not recommended for the very young, but rather for late second grade at the earliest. The skills trained by FFW are the substantive focus of the curriculum up until that point and would likely develop anyway. Third-grade reading curriculum begins to take the learner from learning to read to reading to learn. If curriculum focuses on decoding multisyllabic words at the phonics level, phonemic awareness is not directly taught except as the ability to hear the chunks of sounds in syllables as a part of the initial decoding lessons. Focus on the syllables means specific forms of speech such as suffix, prefix and common Latin or Greek roots rather than decoding the letter combinations. If fluency standards indicate that a reader is still at risk in this grade, in combination with an inability to meet assessment standards for comprehension outside of the risk proxy ORF, then FFW may be helpful. Even then, diagnostic probes should be used to make sure that the individual exhibits lack of automaticity with phonics and difficulty noticing the sounds in words.

Conclusions

Learning to read is difficult for some students, and many programs have been created to help mitigate this issue. At this time, there are two primary sources of expertise informing the field with regard to reading intervention: educational researchers and medically trained doctors. The coordination of information from these fields has influenced the development and use of complex reading intervention systems. This study evaluated one of those interventions, Fast ForWord computer programs, in an effort to determine the impact of FFW on reading fluency. The dependent variable in the study

was oral reading fluency rate. The variable was selected for two reasons: (a) The evaluation team believed that the constructs taught in FFW would lead to improved fluency, and (b) oral reading fluency rate is the primary measure for placing students in interventions.

The effect of FFW on fluency was determined to be inconclusive, as the design of the study could not rule out other substantial threats to validity regarding interpretation of the data. Primary concerns centered on the influence of unknown programs on the comparison group and the treatment group. However, it is not likely that FFW directly impacts oral fluency rates for many people.

An attempt to discern the value of knowing language scores prior to intervention led to the conclusion that this knowledge did not help a school intervention team decide to place students in an intervention whose anticipated outcome was a change in risk status. The notion that individual attributes can predict intended outcomes was not supported in this context.

REFERENCES CITED

Alexander, A., & Slinger-Constant, A. (2004). Current status of treatments for dyslexia: Critical review. *Journal of Child Neurology, 19*(10), 744–768.

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention. *Remedial and Special Education, 23*(5), 300–317.

Anthrop, H. (2002). *School practices for helping children meet language arts standards: Preliminary findings from McREL's study of high-performing, high-needs schools*. Washington, DC: Regional Education Laboratory for the Institute of Education Sciences.

Aylward, E., Richards, T., Berninger, V., Nagy, V., Field, K., Grimme, A., . . . Cramer, S. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology, 61*(2), 212–219.

Bali, V., & Alvarez, M. (2004). The race gap in student achievement scores: Longitudinal evidence from a racially diverse school district. *Policy Studies Journal, 32*(3), 393-415.

Barger, J. (2003). *Comparing the DIBELS Oral Reading Fluency indicator and North Carolina end of grade reading assessment* (Technical Report). Asheville, NC: North Carolina Teacher Academy.

Borman, G., Benson, J., & Overman, L. (2009). A randomized field trial of the Fast ForWord Language computer-based training program. *Educational Evaluation and Policy Analysis, 31*(1), 82–106.

Breier, J., Simos, P., Fletcher, J., Castillo, E., Zhang, W., & Papanicolaou, A. (2003). Abnormal activation of temproparietal language areas during phonetic analysis in children with dyslexia. *Neuropsychology, 17*(4), 610–621.

Brown, W., Eliez, S., Menon, V., Rumsey, J., White, C., & Reiss, L. (2001). Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology, 56*, 781–783.

Campbell, J., Hombo, C., & Mazzeo, J. (2000). *Trends in academic progress: Three decades of student performance* (NCES Report No. 2000-469). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Carrow-Woolfolk, E. (1995). *Oral and Written Language Scales (OWLS)*. Circle Pine, MN: American Guidance Service.

Cognitive Concepts. (1999). *The scientific foundations of Earobics: How proven techniques for developing phonological awareness and auditory processing skills are built into Earobics* [Brochure]. Cambridge, MA: Author.

Cohen, W., Hodson, A., O'Hare, A., Boyle, J., Durrani, T., McCartney, E., . . . Watson, J. (2005). Effects of computer-based intervention through acoustically modified speech (Fast ForWord) in severe mixed receptive-expressive language impairment: Outcomes from a randomized controlled trial. *Journal of Speech, Language and Hearing Research, 48*(3), 715–729.

Coyne, M., Kame'enui, E., Simmons, D., & Harn, B. (2004). Beginning reading intervention as inoculation or insulin: First-grade reading performance of strong responders to kindergarten intervention. *Journal of Learning Disabilities, 37*(2), 90–104.

Denton, C., Vaughn, S., & Fletcher, J. (2003). Bringing research-based practice in reading intervention to scale. *Learning Disabilities Research and Practice, 18*(3), 201–211.

Deppeler, J., Taranto, A., & Bench, J. (2004). Language and auditory processing changes following *Fast ForWord. Australian and New Zealand Journal of Audiology, 26*(2), 94–109.

Divine, K., & Botkin, D. (2008). *A study of the longitudinal effects of Fast ForWord on student performance in Duval County*. Jacksonville, FL: Duval County Public Schools.

Eden, G. (2002). The role of neuroscience in the remediation of students with dyslexia. *Nature Neuroscience Supplement, 5*, 1080–1084.

The Education Trust. (2009). *Education watch: Achievement gap summary tables*. Retrieved from www2.edtrust.org/edtrust/summaries2009/ 2009AchievementGapandSummary Tables.pdf

Engel-Edlar, R., & Rosenhouse, J. (2000). Reading difficulty characteristics in dyslexic and hearing impaired children. *Educational Psychology, 20*(4), 459–482.

Faeoetti, A., Lorusso, M., Paganoni, P., Cattaneo, C., Gaill, R., Umilta, C., & Mascetti, G. (2003). Auditory and visual attention deficits in developmental dyslexia. *Cognitive Brain Research, 16*(2), 185–191.

Fawcett, A., Nicholson, R., & Dean, P. (1996). Impaired performance of children with dyslexia on a range of cerebellar tests. *Annals of Dyslexia, 46*, 259–283.

Frome-Loeb, D., Gillam, R., Hoffman, L., & Marquis, J. (2009). The effects of Fast ForWord Language on phonemic awareness and reading skills of school-age children with language impairments and poor reading skills. *American Journal of Speech-Language Pathology, 18,* 376–387.

Fuchs, L. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disability Research and Practice, 18*(3), 172–186.

Fuchs, L., Fuchs, D., Hosp, M., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239–256.

Gaab, N., Gabrieli, J., Deutsch, G., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: An *f*MRI study. *Restorative Neurology and Neuroscience, 25*, 295–310.

Gillam, R. (1999). Treatment for temporal processing deficits computer-assisted language intervention using Fast ForWord: Theoretical and empirical considerations for clinical decision-making. *Language, Speech, and Hearing Services in Schools, 30*, 363–370.

Gillam, R., Crofford, J., Gale, M., & Hoffman, L. (2001). Language changes following computer-assisted language instruction with Fast ForWord or Laureate Learning Systems Software. *American Journal of Speech-Language Pathology, 10*(3), 231–247.

Gillam, R., Frome-Loeb, D., & Friel-Patti, S. (2001). Looking back: A summary of five exploratory studies of Fast ForWord. *American Journal of Speech-Language Pathology, 10*, 269–273.

Gillam, R., Frome-Loeb, D., Hoffman, L., Bohman, T., Champlin, C., Thibodeau, L., . . . Friel-Patti, S. (2008). The efficacy of Fast ForWord-Language intervention in school-age children with language impairment: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research 51*(1)*,* 97–119. .

Given, B., Wasserman, J., Chari, S., Beattie, K., & Eden, G. (2008). A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain and Language, 106,* 83–97.

Good, R., & Kaminski, R. (2002). *Dynamic Indicators of Basic Early Literacy Skill (DIBELS)*. Eugene, OR: University of Oregon.

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

Good, R., Wallin, J., Simmons, D., Kame'enui, E., & Kaminski, R. (2002). *System-wide percentile ranks for DIBELS benchmark assessment* (Technical Report No. 9). Eugene, OR: University of Oregon.

Habib, M. (2000). Review article: The neurological basis of developmental dyslexia—An overview and working hypothesis. *Brain, 123*, 2373–2399.

Habib, M., Rey, V., Daffaure, V., Camps, R., Espesser, R., Joly-Pottuz, B., & Demonet, J. (2002). Phonological training in children with dyslexia using temporally modified speech: A three-step pilot investigation. *International Journal of Language and Communication Disorders, 37*(3), 289–308.
.
Hall, S. (2002). *Final report: Scientific Learning/Fast ForWord program: 2001-2002* (Technical Report). Dallas, TX: Dallas Independent School District.

Hasbrouck, J., & Tindal, G. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*(7), 636–644.

Helenius, P., Salmelin, R., Richardson, U., Leinonen, S., & Lyytinen, H. (2002). Abnormal auditory cortical activation in dyslexia 100 msec after speech onset. *Journal of Cognitive Neuroscience, 14*(4), 603–617.

Hook, P., Macaruso, P., & Jones, S. (2001). Efficacy of Fast ForWord training on facilitating acquisition of reading skills by children with reading difficulties: A longitudinal study. *Annals of Dyslexia, 51*, 75–96.

Individuals with Disabilities Education Improvement Act. (2004). Public Law 108-446 (20 U.S.C. 1400 *et seq.*).

Joanisse, M., Manis, R., Keating, P., & Seidenber, M. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology, 77*(1), 30–60.

Kujala, T., Karma, K., Myllyviita, K., Tervaniemi, M., Alho, K., Kallio, J., & Naatanen, R. (2000). Basic auditory dysfunction in dyslexia as demonstrated by brain activity measurements. *Psychophysiology, 37*, 262–266.

LaBerge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.

Loeb, D., Stoke, C., & Fey, M. (2001). Language changes associated with Fast ForWord-language: Evidence from case studies. *American Journal of Speech-Language Pathology, 10*(3), 216–230.

Logan, G. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly, Overcoming Learning Difficulties, 13*, 123–146.

Marler, J., Champlin, C., & Gillam, R. (2001). Backward and simultaneous masking measured in children with language-learning impairments who received intervention with Fast ForWord or Laureate Learning Systems software. *American Journal of Speech-Language Pathology, 10*(3), 258–268.

Merzenich, M., Jenkins, W., Johnson, P., Scheiner, C., Miller, S., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271*, 77–81.

National Center for Education Statistics. (2009). *The nation's report card: Reading 2009*. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/stt2009/2010460NP4.pdf

National Reading Panel. (2001). *Report of the National Reading Panel: Teaching children to read*. Bethesda, MD: National Institute of Child Health and Human Development, National Reading Panel.

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

Newcomer, P., & Hammill, D. (1997). *Test of Language Development, Primary* (TOLD P: 3, 3rd ed.). Austin, TX: Pro-Ed.

No Child Left Behind Act (NCLBA) Act of 2001, Pub. L. No. 107-110, §115, Stat. 1425 (2002).

O'Connor, R., & Jenkins, J. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies in Reading, 3*, 159–197.

Overbay, A., & Baenan, N. (2003). *Fast ForWord evaluation* (Report). Raleigh, NC: Wake County Public School System.

Pepperdine University. (2005). *Psychometric conversion table*. Retrieved from http://faculty.pepperdine.edu/shimels/Courses/Files/ConvTable.pdf

Pokorni, J., Worthington, C., & Jamison, P. (2004). Phonological awareness intervention: A comparison of Fast ForWord, Earobics, and LiPS. *Journal of Educational Research, 97*(3), 147–157.

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*, 1–7.

Ramus, F., Rosen, S., Dakin, S., Day, B., Castellote, J., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*(4), 841–865.

Rashotte, C., & Torgesen, J. (1985). Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly, 20*, 180–188.

Rashotte, C., & Torgesen, J. (1985). Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly, 20,* 180–188.

Ross-Kidder, K. (2003, March 15, 2004). "Reading disability" or "learning disability": The debate, models of dyslexia and review of research-validated reading programs. Retrieved from http://www.ldonline.org/ld_indepth/reading/reading_approaches.html

Rouse, C., Krueger, B., & Markman, L. (2004). *Putting computerized instruction to the test: A randomized evaluation of a "scientifically-based" reading program.* Retrieved from http://www.nber.org/papers/w10315

Schenck, A., Walker, D., Nagel, C., & Webb, L. (2005). *Analysis of State K-3 Reading Standards and Assessments.* Retrieved from http://www2.ed.gov/rschstat/eval/other/reading/state-k3-reading.pdf

Schopmeyer, B., Mellon, N., Dobaj, H., Grant, G., & Niparko, J. (2000). Use of Fast ForWord to enhance language development in children with cochlear implants. *Annals of Otology, Rhinology and Laryngology, Supplement, 185*(109), Part 2: 95–08.

Schwaneflugel, P., Meisinger, E., & Wisenbaker, J. (2006). Becoming a fluent and automatic reader in the early elementary school years. *Reading Research Quarterly, 41*(4), 496–522.

Scientific Learning. (2010). *Fast ForWord Language v2*. Retrieved from http://www.scilearn.com/prod2/ffwd_l/ main=howitworks

Semel, F., Wiig, F., & Secord, W. (1995). *Clinical Evaluation of Language Fundamentals* (CELF-3, 3rd ed.). San Antonio, TX: Psychological Corporation.

Share, D., & Leikin, M. (2004). Language impairment at school entry and later reading disability: Connections at lexical versus supralexical levels of reading. *Scientific Studies of Reading, 8*(1), 87–110.

Shaywitz, B., Shaywitz, S., Pugh, K., Fulbright, R., Menci, W., Constable, R., . . . Gore, J. (2001). The neurobiology of dyslexia. *Clinical Neuroscience Research, 1*(4), 291–299.

Shaywitz, S., Fletcher, J., Holahan, J., Shneider, A., Marchione, K., Stuebing, K., . . . Shaywitz, B. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at adolescence. *Pediatrics, 104*(6), 1351–1359.

Simos, P., Fletcher, J., Bergman, E., Breier, J., Foorman, B., Castillo, E., . . . Papanicolaou, A. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58*, 1203–1212.

Speece, D., & Case, L. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93*, 735–749.

Stevens, C., Fanning, J., Coch, D., Sanders, L., & Neville, H. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: Electrophysiological evidence from language-impaired and typically developing children. *Brain Research, 1205*, 55–69.

SuccessMaker® (Version 4.0) [Computer software]. Upper Saddle River, NJ: Pearson Digital Learning.

Talcott, J., Witton, C., Hebb, G., Stoodley, C., Westwood, E., France, S., . . . Stein, J. (2002). On the relationship between dynamic visual and auditory processing and literacy skills: Results from a large primary-school study. *Dyslexia, 8*(4), 204–225.

Tallal, P. (2003). Language learning disabilities: Integrating research approaches. *Current Directions in Psychological Science, 12*(6), 206.
.
Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Nagarajan, S., . . . Merzenich, M. (1996). Language comprehension in language learning impaired children improved with acoustically modified speech. *Science*, 271, 81–84.

Tallal, P., Miller, S., Bedi, G., Wang, X., Nagarajan, S., Schreiner, S., . . . Merzenich, M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Nature, 271*, 81–84.

Temple, E., Deutsch, G., Poldrack, R., Miller, S., Tallal, P., Merzenich, M., & Gabrieli, D. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *PNAS-Proceedings of the National Academy of Sciences, 100*(5), 2860–2865.

Temple, E., Poldrack, R., Protopapas, A., Nagarajan, S., Salz, T., Tallal, P., . . . Gabrieli, J. D. E. (2000). Disruption of the neural response to rapid acoustic stimuli in dyslexia: Evidence from functional MRI. *PNAS-Proceedings of the National Academy of Sciences, 97*(25), 13907–13912.

Thibodeau, L., Friel-Patti, S., & Britt, L. (2001). Psychoacoustic performance in children completing Fast ForWord training. *American Journal of Speech- Language Pathology, 10,* 248–257.

Torgesen, J., Rashotte, A., Alexander, J., & MacPhee, K. (2003). Progress towards understanding the instructional conditions necessary for remediating reading difficulties in older children. In B. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 275–298). Parkton, MD: York Press.

Troia, G. (2004). Migrant students with limited English proficiency: Can Fast ForWord Language(TM) make a difference in their language skills and academic achievement? *Remedial and Special Education, 25*(6), 353–366.
.
Troia, G., & Whitney, S. (2003). A close look at the efficacy of Fast ForWord Language for children with academic weaknesses. *Contemporary Educational Psychology, 28*, 465–494.

Turner, S., & Pearson, D. (1999, Spring/Summer). Fast ForWord Language intervention program: Four case studies. *Texas Journal of Audiology and Speech Pathology, 13*, 1–3.

Vaughn, S., & Chard, D. (2006, Winter). Three-tier intervention research studies: Descriptions of two related projects. *Perspectives, 32*(1), 29–34.
.
Vaughn, S., Moody, S., & Shuman, J. (1998). Broken promises: Reading instruction in the resource room. *Exceptional Children, 64*, 211–225.

Veale, T. (1999). Targeting temporal processing deficits through Fast ForWord: Language therapy with a new twist. *Language, Speech, and Hearing Services in Schools, 30*, 353–362.

Vellutino, F., Fletcher, J., Snowling, M., & Scanlon, D. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2–40.

Walczyk, J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly, 35*(4), 554–566.

The What Works Clearinghouse. (2002). *Fast ForWord*. Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from http://ies.ed.gov/ncee/wwc/search/index.asp?q=Fast+ForWord&btnG=Search&output=xml_no_dtd&client=wwc&sort=date%3AD%3AL%3Ad1&entqr=3&site=wwc

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading, 5*(3), 211–239.