JEAN STOCKARD, National Institute for Direct Instruction and University of Oregon

Promoting Early Literacy of Preschool Children: A Study of the Effectiveness of Funnix Beginning Reading

Abstract: Thirty-seven students from a suburban community in the southern United States were randomly assigned to receive 30 minutes of additional instruction each day in their usual language arts curriculum or 30 minutes of instruction with the Direct Instruction program, *Funnix Beginning Reading*. Instruction for the *Funnix* group was provided by high school-aged tutors, trained and supervised by an experienced teacher. Pretesting before instruction began indicated no significant differences between the two groups in letter naming fluency or initial sound fluency. However, by winter and spring the students in the *Funnix* group had significantly higher scores on numerous measures of beginning literacy. These results occurred with simple comparison of means, comparisons of scores to established benchmarks, and multivariate analyses that controlled for initial levels of skill and minority status. The results also appeared when a reduced sample that individually matched children on their pretest scores was used.

A large body of literature has documented the relationship of early reading achievement to later academic accomplishments and economic and social well being. Students who are poor readers in first grade have substantially higher probabilities of later academic, economic, and social problems than students who achieve at grade level at that time (e.g., Francis, Stuebing, Shaywitz, Shaywitz, & Fletcher, 1996; Lipson & Wixson, 1997; Wharton-McDonald, Pressley, & Hampston, 1998). These consistent and strong research findings have prompted extensive policy attention to promoting early reading, especially for populations judged to be at risk based on the poverty status of their families.

The federally funded Head Start program has long been part of the policy approach to enhance the academic preparedness of children in poverty. Controversy exists over the extent to which Head Start has provided academic advantages for children from disadvantaged backgrounds (e.g., Currie & Thomas, 1995, 2000; Lee & Loeb, 1995), and recent changes to the legislation have enhanced requirements for academic instruction in the program. Specifically, the "Improving Head Start for School Readiness Act of 2007" requires that Head Start agencies "implement a research-based early childhood curriculum that promotes young children's...early reading and mathematics skills... is based on scientifically valid research, and has standardized training procedures and curriculum materials to support implementation" (U.S. House of Representatives, 2007, pp. 43-44). It is thus important to examine the extent to which various curricula can enhance preschool children's academic skills.

Using a classic experimental design with random assignment of students to conditions, this study compares the beginning literacy skills of

Journal of Direct Instruction, Vol. 10, pp. 29-48. Address correspondence to Jean Stockard at jstockard@nifdi.org.

4-year-old Head Start students who used *Funnix Beginning Reading,* a computer-assisted beginning reading program, with the beginning literacy skills of students who had the same amount of extended instruction in their usual language arts curriculum. Each *Funnix* student was paired with a high school tutor who implemented the program 30 minutes a day. The other students worked with their regular classroom teachers. Results indicate that the *Funnix* students had significantly higher beginning reading skills at the end of the school year than students in the control condition.

Related Literature

Our analysis is based on several different strands of literature: (1) studies of early literacy, (2) studies of Direct Instruction (DI), (3) studies specifically related to computer-assisted instruction and tutoring, and (4) previous studies of the implementation of *Funnix* and related programs.

Promoting Early Literacy

The National Reading Panel's report on reading instruction identified five areas of reading instruction that should be part of children's primary grade instruction: phonemic awareness, phonics, fluency, vocabulary, and text comprehension (National Institute of Child Health and Human Development [NICHD], 2000). The panel, accompanying meta-analyses of the research literature, and numerous individual studies have demonstrated that phonemic awareness and phonics-oriented preliteracy and early literacy instruction play a crucial role in enhancing early reading achievement (Ehri, Nunes, Stahl, & Willows, 2001; Ehri et al., 2001; NICHD, 2000).

The areas identified by the National Reading Panel's report parallel theoretical models regarding the development of reading skills (e.g., Chall, 1983; Ehri, 2005; Ehri & McCormick, 1998; Simmons & Kame'enui, 1998). These models describe how the foundational skill of phonological awareness, or being able to hear and manipulate sound structures, precedes the development of alphabetic understanding or the understanding of the relation of print to speech. This, in turn, precedes phonological recoding of letter strings to sounds, which precedes the eventual reading of words and then connected text. The various models see these skills as overlapping but ranging along a continuum, with the end goal of attaining fluency in reading by the end of the primary grades.

All of these analyses emphasize the importance of early achievement and the cumulative effects of learning. The long-term impact of early academic learning has been captured with discussions of the "Matthew effect," using the Biblical quotation that the "rich get richer and the poor get poorer" to describe the long-term and cumulative effects of good or poor reading skills on later academic success (Stanovich, 1986). A large body of empirical evidence demonstrates that early reading ability has lasting impacts on students' academic careers. Those who are able to read fluently in first grade have much more success throughout their school careers (Cunningham & Stanovich, 1997, 1998; Francis et al., 1996; Gough & Juel, 1991; Juel, 1988; Stanovich, 1986). Similar results appear with those who have high reading skills before starting the primary grades (Barnett, 1995; Durkin, 1974-75; Morris, Bloodgood, & Perney, 2003).

Direct Instruction

DI (distinguished from other "direct instruction" approaches, which embody only some of DI's characteristics, by the use of capital letters) has long been recognized as one of the most effective programs in promoting student achievement. The curricula are highly structured and carefully designed to provide systematic and explicit instruction. They are also designed to accelerate students' learning by teaching more than traditional programs in the same amount of time.

All DI programs include five critical features. First, lessons do not focus on a single topic (such as rhyming or vocabulary) but instead work on five or more different skills. Each skill is practiced and applied in more than one lesson, providing repeated and integrated practice. Second, only about 10% of any one lesson involves new skills or concepts, with the remainder involving review and application of material that was introduced previously. This small-step design and continuous review has been found to ensure that all children learn all the skills and concepts presented, even as they become increasingly complex. Third, the DI programs are scripted to ensure that teachers provide explanations that are adequate, quick, and efficient. Fourth, the programs are structured to permit accurate predictions of students' progress, with the expectation that students will progress at the rate of one lesson per day if they are given sufficient time and follow the program carefully. Finally, all DI programs are extensively pretested and then revised based on children's performances during field tests.

Numerous studies have demonstrated that DI promotes significantly higher achievement and more positive self-concepts and school-related attitudes than other curricula. These results have appeared with the general population (e.g., Carlson & Francis, 2002; O'Brien & Ware, 2002; Vitale & Joseph, 2008) and with students with disabilities (e.g., Benner, 2007; Benner, Kinder, Beaudoin, Stein, & Hirschmann, 2005; Cooke, Gibbs, Campbell, & Shalvis, 2004; Flores, Shippen, Alberto, & Crowe, 2004; Malmgren & Leone, 2000; Riepl, Marchand-Martella, & Martella, 2008).

Other studies have shown that the programs are, as would be expected, most effective when delivered with the highest levels of fidelity. The literature increasingly recognizes that teaching is a highly technical and involved process and that training and support are crucial for developing and honing excellent teaching skills. Studies suggest that this assistance should be ongoing and intensive, ideally involving onsite support (Berends, Bodilly, & Kirby, 2002; Blakeley, 2001; Bodilly, 1998; Bodilly, Glennan, Kerr, & Galegher, 2004). Such support may be especially important for programs, such as DI, that require more extensive changes in teacher behavior (Engelmann & Engelmann, 2004). Studies have found that the gap between students in DI programs and those in traditional programs is greater for students of teachers who implemented DI with higher fidelity (Carlson & Francis, 2002; Gersten, Carnine, Zoref, & Cronin, 1986).

Enhancing Achievement Through Computer-Assisted Instruction and Tutoring

Computer-assisted instruction may provide a way to enhance fidelity of implementation. Through very careful design, computer-assisted instructional programs can control the sequencing, pacing, and presentation of material. To the extent that computer-assisted programs are individualized, they also can facilitate the accurate placement of students, ensure that they are studying levels of material that are appropriate for their skills, and allow them to move as quickly as they can. Some evidence suggests that when compared to traditional presentations of material, computer or web-based programs can produce greater gains in reading-related skills (e.g., Cole & Hilliard, 2006; Lonigan et al., 2003).

Another way to individualize the implementation of curricula is through the use of one-on-one tutoring. Studies indicate that tutoring can improve literacy and beginning reading skills, but the effects depend on the quality of the curriculum and on implementing programs with fidelity. Positive effects have been found for tutors from various backgrounds including college students, community members, paraprofessionals, and parents (Allor & McCathren, 2004; Al Otaiba, Schatschneider, & Silverman, 2005; Ritter, Denny, Albin, Barnett, & Blankenship, 2006; Segal-Drori, Korat, Shamir, & Klein, 2009; Vadasy, Jenkins, Antil, Wayne, & O'Conner, 1997; Vadasy, Sanders, & Peyton, 2005)

Funnix: A Computerized DI Program

Funnix Beginning Reading is a computer-assisted reading program designed for one-on-one use with a parent, teacher, or tutor. *Funnix* builds on the DI program called *Horizons* (published by SRA), which was developed in the late 1990s. Research using the final published version of *Horizons* indicates that students using the program have significantly higher achievement than students who use other materials (Tobin, 2003, 2004, 2009).

While *Horizons* is delivered only through printed material, *Funnix* packages the curriculum in a format that is accessed and used on a computer. The program includes 120 carefully sequenced lessons using a computer reading format and narrator, coordinated graphics, and animation. Each lesson requires about 30 minutes to complete. A manual and CD for parents, or other instructors, explains the details of the reading instruction and the procedures that should be followed when using the program. Teachers or tutors must operate a mouse during lessons to preserve the prompting, pacing, and directions built into the *Funnix* program and to facilitate the provision of the feedback that is built in to reinforce appropriate responses and to correct mistakes.

Both *Horizons* and *Funnix* incorporate four phases of instruction, paralleling the development of reading skills described in the theoretical discussions summarized above. The first is a prereading phase in which children learn to identify letters and their sounds. Second is a highly prompted reading phase during which students learn to read words in isolation; then words, phrases, and sentences in stories; then short stories consisting of two sentences; and, eventually, short stories of about 90 words in length. This phase uses prompts to help students recognize the presence of irregular words, letter combinations, and silent letters. A third reading phase includes stories with few prompts. Some of the stories are ones that the students read earlier, but the prompts are removed, while other stories are new and written with few prompts. In the fourth and final phase, all prompts have been removed from the word lists and stories.

As in all the DI materials, the pacing, scope, and sequence of *Horizons* and *Funnix* have been carefully designed and tested. For instance, the programs always preteach sounds and words before they appear in stories, building from teaching of individual sounds, to using the sounds in words, to having the words within a story. In addition, the sequence in which letters are introduced is carefully designed based on research regarding how difficult it is for students to learn each letter sound or letter group. Sounds are introduced from the easiest to the most difficult, and high-utility letter sounds are introduced before low-utility sounds. Blending and segmenting phonemic tasks are taught explicitly and are also carefully sequenced, again building on previous research.

Each *Funnix* lesson includes elements related to decoding and comprehension. *Funnix* promotes fluency throughout the sequence by directing children to reread lists of sounds, lists of words, and stories. *Funnix* also promotes fluency by gradually and systematically adjusting the audio and visual prompts that signal responses. Comprehension activities are presented in connection with story reading throughout the sequence. In the first two phases, before children have learned to read stories, *Funnix Beginning Reading* presents stories orally and asks comprehension questions. After children begin reading stories at the end of the second phase, they answer orally presented comprehension questions during a second reading of the story. At the end of *Funnix Beginning Reading*, children read and write answers to story questions and other comprehension activities using an associated workbook. The workbook activities also include practice in beginning spelling.

Funnix includes strategies for teaching children multisyllabic words and more sophisticated text. After children master sounding out regular words and gain automaticity in decoding whole words, *Funnix* begins teaching children to apply a sounding-out strategy to read more complex words. Children learn and apply strategies that focus on the spelling of words and on familiar word parts. They also learn strategies for expressively reading connected text that contains sophisticated punctuation marks, such as quotes, exclamations, and ellipses. *Funnix* is designed for use in schools as the primary reading program, as an intervention program, as a supplement, or as a summer school or after-school program. It can be used in small groups, as a tutorial with a regular tutor, in a peer-tutoring setting, or as a reinforcer in a paired-practice setting.

Two previous, small-scale studies of student achievement have found a relationship between using *Funnix* and higher achievement. Parlange (2004) examined changes in scores on standardized tests of reading achievement of 10 preschool-aged children who used the *Funnix* program under the guidance of university undergraduates assisting in the preschool. Comparisons of changes in reading achievement over time with normative samples indicated that all of the children experienced substantial improvement in word attack and expressive language after using the program. Similarly, Watson and Hempenstall (2008) compared the achievement growth of 15 kindergarten and first-grade students who used *Funnix* in at-home settings with their parents with a wait-listed comparison group of students in the same grades. The *Funnix* students in both grades had statistically significant improvements over time, but only the gains for the kindergarten students were significantly greater than those in the comparison group.

This study compares the development of early literacy skills for preschool students exposed to *Funnix Beginning Reading* and students exposed to a traditional early literacy program. Students studying *Funnix* were aided by high school-age tutors in a supervised setting, while the other students were taught by their regular classroom teachers. This study adds to the relatively sparse literature on *Funnix* by examining its effectiveness with a different population than used in other studies, by randomly assigning students to treatments, and by employing high school-aged tutors as instructors.

Method

Participants and Procedures

Participants in this study were students in a Head Start program in a suburban area of the southern United States. From approximately 100 four-year-olds in six classrooms, 40 students were randomly selected to participate. The students selected for the study were then randomly assigned to either the control group or the experimental (*Funnix*) group. Because of attrition, the final sample included 37 students (19 in the control group and 18 in the experimental group). Nineteen of the children were non-Hispanic whites, 13 of the children were African American, 4 were Hispanic, and 1 was Asian American. Ten of the children came from homes where English was not the primary language. All of the children were from low-income families, and none had identified disabilities.

All students in the Head Start classrooms received regular in-class instruction with the locally adopted language arts curriculum. The program was developed by a consortium of Head Start programs within the state and was self-published by Head Start. Students in both the *Funnix* and control groups received 30 minutes of supplemental instruction. Those in the control group received 30 minutes of additional instruction in their locally adopted language arts program, while the additional 30 minutes of instruction for those in the experimental group was in *Funnix*. Instruction for the control group occurred in a full class or small group format, as specified by the Head Start curriculum. All instruction for the control group was provided by the regular classroom teachers and teachers' aides.

The *Funnix* instruction was guided by public high school students from the local community. Students chosen to work with the preschoolers were carefully screened and selected based on their past academic performance, school attendance, good conduct, and recommendations from the school counselor. Each high school student was matched with one preschool student and worked with that student throughout the year.

The high school students were trained for a total of 6 hours on how to serve as a *Funnix* instructor. The training focused on details of using the program and on appropriate procedures for reinforcing the preschoolers' responses using materials included with the *Funnix* program. Tutors were told to be positive, upbeat, and encouraging, while allowing the children to have enough "think time" to work through a problem. Tutors were also instructed on how to control the mouse in order to pace the program appropriately. Finally, they were instructed in the proper procedures for correcting different types of student errors and trained on how to model correct answers for children who were hesitant or didn't know an answer.

In addition to technical instruction on the *Funnix* program, the tutors were given detailed guidance on the standards of conduct required for working in the Head Start classrooms, and they signed forms indicating that they understood the standards to which they would be held. The high school students were not allowed to work with the preschoolers until the trainer believed they fully understood all required procedures. At the end of the year the high school tutors were required to write an essay regarding their experience, and they received grades and other recognition for their work.

Funnix students and their individual tutors worked either in the Head Start Media Center or in a classroom that was designated for their exclusive use. Computer stations were installed in these two settings and cardboard study carrels were provided for each computer station to minimize visual distractions. All paired sets of tutors and students had a computer workstation and used earphones to eliminate auditory disturbances from other groups working with the *Funnix* program in the same room.

A certified teacher with a Master's degree who had formerly worked at the tutors' high school supervised the implementation of the daily *Funnix* instruction at the Head Start site. She was present for each day of instruction, actively observed the interactions, and intervened as needed. Tutors were required to keep track of their students' progress. At the end of each tutoring session, tutors completed a log sheet that recorded the lessons and tasks covered, activities completed, their assessment of the child's performance, and any additional comments or concerns. These log sheets were regularly reviewed by the certified teacher.

Instruction began in October 2006 and continued until May 2007. Children were tested three times during the year: in late September 2006, before instruction began; in January 2007, midway through the school year; and in May 2007, at the end of the year. Testers were independent of the Head Start Program and the school district and were supervised and trained by an independent school psychologist. The author was not involved in the implementation or testing.

Measures

Table 1 lists the measures that were administered to all of the children in the experimental and control groups. The Basic Language Concepts Test (BLCT) (Engelmann, Ross, & Bingham 1982), formerly called the Basic Concept Inventory, was administered at all three testing periods. The BLCT is an individually administered instrument designed to screen children, 4 1/2 to 6 years of age, for language skills important for beginning school learning and was especially developed for children in settings such as Head Start. It is also used to diagnose specific skill deficiencies and to provide baseline measures for evaluating progress. The test assesses four general areas: 1) receptive language, the child's ability to understand common words or phrases; 2) imitative function, the child's ability to repeat statements by the tester; 3) representational functions, the child's ability to answer simple questions; and 4) pattern function, the child's ability to repeat a patterned series and to recognize a sequence of actions. Internal reliability of the total scale (Kuder-Richardson 21), based on the original norming population, is 0.90. Predictive validity was assessed through comparing scores on the BLCT with kindergarten and first-grade teachers' ratings of language performance and reading performance (first grade only), with results ranging from 84% to 86% agreement.

In the fall administration, the language skills of many of the children were very low and they were unable to correctly answer many questions on the BLCT. To minimize the students' discomfort the testers ceased administration of the BLCT if students were unable to correctly answer more than 3 of the first 14 receptive language items. As a result, fewer than half (n=13) of the children completed the test in the fall, and we only report data on the BLCT for winter and spring. The total score on the test is the number of errors—that is, a higher score indicates a child has more difficulties with language. The score may be compared to norms from the original test standardization.

The Dynamic Test of Basic Literacy Skills (DIBELS) (5th edition) (DIBELS, 2002) was administered in fall, winter, and spring. The DIBELS measures have high statistical reliability and can be compared against established benchmarks that indicate the level at which students should achieve to reach generally accepted literacy goals. All of the measures

Table 1

Measures Administered to Students by Testing Time

	Fall	Winter	Spring
Basic Language Concepts Test	X	X	X
Initial Sound Fluency (ISF)	X	X	
Letter Naming Fluency (LNF)	X	X	X
Phoneme Segmentation Fluency (PSF)		X	X
Nonsense Word Fluency (NWF)		X	X
Controlled Text Passages (2 passages)			X
100-Word Test			X

are timed assessments, and scores reflect the number of correct answers given within a set duration. The Kindergarten Benchmark Assessment form, the earliest form available, was used. Initial sound fluency (ISF) was assessed in the fall and winter, letter naming fluency (LNF) was assessed at all three time points, and phoneme segmentation fluency (PSF) and nonsense word fluency (NWF) were assessed in both winter and spring. Both the raw scores (the number of correct responses in a minute) and whether or not the child reached kindergarten benchmarks were examined.

Two measures of oral reading fluency were assessed in the spring. The first was the 100-Word Test, which presents children with a set of 100 words that are typically learned early in a reading program. Words on the list are no more than two syllables but contain a variety of vowel and consonant combinations. Students are asked to read the words, and responses are marked as incorrect if the child reads the word incorrectly or does not respond within four seconds. Testing is terminated after the child misses four words in a row or indicates that he or she doesn't know how to read any more words (after being asked about "a" and "I"). The score on the test is simply the number of words read correctly.

The other measure of oral reading was derived from two short, controlled text passages, each involving a very short story. The number of words that students read correctly in the two passages was highly correlated (r = .97), so these numbers were summed for analysis.

Analysis

The scores of students in the control and experimental groups were compared in several different ways. First, simple descriptive statistics (means and standard deviations) were obtained for the raw scores for each testing period. To obtain an estimate of the magnitude of the difference between the two groups, t-tests and effect sizes (Cohen's d) were calculated. If *Funnix* were more effective than the regular classroom curriculum in promoting beginning literacy, we would expect greater differences between the two groups at the later administrations. Based on this directional hypothesis, we used one-tail tests of significance. We also used the published norms for the BLCT and benchmarks for the DIBELS measures to calculate the proportion of children in each group who would be considered at risk for future academic problems. If *Funnix* were more effective we would expect fewer *Funnix* students to be at risk at the end of preschool.

Second, we used multivariate analyses. Students' growth over time on the BLCT and the DIBELS measures was examined using repeated measures analyses of variance (ANOVA), with the tests at each time point as repeated measures and experimental condition as a factor. If *Funnix* were more effective, stronger gains would be expected for the experimental group. This would result in a significant interaction between the repeated measures and condition (*Funnix* vs. control). An analysis of covariance (ANCOVA) was used to analyze the two measures of reading fluency (i.e., the 100-Word Test and the controlled text passages) in which fall LNF scores were used as covariates in the assessment of group differences. Fall LNF was chosen as the covariate after preliminary analysis indicated that it had the highest correlation of the fall scores with the spring measures of reading fluency (r = 0.57 for the 100-Word Test and 0.53 for the controlled text passages).

Third, we divided the children into two groups based on their race-ethnicity and home language: 1) students who were non-Hispanic whites and whose families spoke English at home, and 2) students who came from families whose home language was not English and/or who belonged to a racial-ethnic minority. Although all the children came from low-income homes, it could be expected that race-ethnicity and home language might provide further educational barriers and that children from the nonminority English-speaking homes could have had an educational advantage that might have affected their progress. We examined the average scores on tests completed in the spring for minority and nonminority students in the *Funnix* and control groups and then used two-way ANOVA, with group and minority status as factors, as an inferential test.

Finally, we used post-hoc matching to create a reduced sample of students that was as closely matched as possible on beginning literacy scores as well as gender and race ethnicity. This technique was used to provide yet another way of introducing controls into our analyses. To obtain the cases, the students were rank ordered on their pretest scores on LNF. Next, pairs of children with similar LNF scores and equivalent race-ethnicity, home language, and gender were selected. One member of each pair had been randomly assigned to the control group and one had been randomly assigned to the *Funnix* group. We calculated descriptive statistics, t-tests, and Cohen's d values for this reduced sample and compared the results to those obtained with the total group. Then, in a descriptive analysis, we focused on two pairs of children, one with very low scores at pretest and one with high scores, and compared changes in their scores over the academic year.

It should be remembered that the sample size for this study is very small, with fewer than 20 students within each group. However, the research design includes several important elements that enhance its internal validity. First, the students were randomly assigned to the experimental and control groups, the classic method of ensuring comparability of groups. Second, all children received the same amount of additional instruction in language arts. The only difference was the curriculum used for this additional instruction. Third, while studies with small samples are often limited by having only one teacher for each condition, multiple instructors were involved with both the experimental and control groups. This effectively eliminates the possibility of an instructor-treatment interaction. Fourth, multiple measures of achievement were used, and these measures were gathered at three different time points by testers not involved with the instruction. Finally, while the effect sizes are not influenced by the sample size, the inferential statistical tests adjust for sample size, making it more difficult to attain statistical significance than if a larger sample had been used.

Results

Descriptive Statistics and Bivariate Analyses

Table 2 gives descriptive statistics (means and standard deviations) on all measures for all administrations, the results of t-tests comparing mean scores for the two groups at each time point, and the corresponding effect sizes. As would be expected, given the random assignment design, differences between the two groups on scores obtained in the fall were not statistically significant, although the *Funnix* students had slightly higher scores on both measures (LNF and ISF).

In winter, as expected, the *Funnix* students had scores that were significantly higher than those of students in the control conditions on all of the measures. The *Funnix* students had significantly fewer errors on the BLCT and significantly higher scores on the four DIBELS measures: LNF, ISF, PSF, and NWF. The Cohen's d values comparing scores of students in the two groups range from 0.72 to 1.55, well beyond the level of 0.25 typically characterized as large or educationally significant.

The results in spring continue to show a strong advantage for the *Funnix* students. Although *Funnix* students had higher scores on all measures, *t*-tests reveal nonsignificant dif-

Table 2

Descriptive Statistics, t-tests, and Effect Sizes

BLCT—Total Score (# Errors)

	M	*SD*	*n*	*t*	*p*	Cohen's *d*
Control Group—Winter	30.7	9.5	15	-3.04	0.003	-1.14
Funnix Group—Winter	19.2	10.7	14			
Control Group—Spring	22.4	7.9	19	-1.54	0.07	-0.51
Funnix Group—Spring	18.0	9.6	18			

Letter Naming Fluency (LNF)

	M	*SD*	*n*	*t*	*p*	Cohen's *d*
Control Group—Fall	8.8	9.1	19	1.17	0.12	0.39
Funnix Group—Fall	12.8	11.3	18			
Control Group—Winter	17.6	11.7	19	2.19	0.02	0.72
Funnix Group—Winter	27.7	16.1	18			
Control Group—Spring	24.6	15.5	19	2.47	0.01	0.81
Funnix Group—Spring	37.8	17.1	18			

Initial Sound Fluency (ISF)

	M	*SD*	*n*	*t*	*p*	Cohen's *d*
Control Group—Fall	4.4	5.6	19	0.49	0.31	0.17
Funnix Group—Fall	5.4	6.3	18			
Control Group—Winter	10.7	5.3	18	2.82	0.004	0.98
Funnix Group—Winter	17.7	9.0	17			

Phoneme Segmentation Fluency (PSF)

	M	*SD*	*n*	*t*	*p*	Cohen's *d*
Control Group—Winter	4.1	4.7	19	4.19	<0.001	1.55
Funnix Group—Winter	18.0	13.4	18			
Control Group—Spring	8.2	9.4	19	6.79	<0.001	2.24
Funnix Group—Spring	32.1	12.0	18			

ferences ($p > 0.05$) among groups on the BLCT, 100-Word Test, and controlled text passages. Differences were significant for all of the DIBELS measures of beginning literacy ($p < 0.05$). All of the effect sizes comparing the scores for the *Funnix* and control groups, including those that were not statistically significant, surpassed the usual criterion of educationally significant, ranging from 0.51 to 2.24.

We compared the preschool students' DIBELS scores at the spring administration (winter scores for ISF) to the established benchmarks for the DIBELS measures for kindergarten students (Good, Simmons, Kame'enui, Kaminski, & Wallin, 2002). Scores were compared with the fall kindergarten norms for LNF and ISF and with the winter norms for NWF and PSF, the first point at which this DIBELS measure is typically administered. Thus, results indicate the extent to which the Head Start preschool children would be considered at risk for later academic problems compared to other children at the beginning or middle of kindergarten.

Results indicate that the control students achieved the greatest success in learning their letters and initial sounds. None of the students in either the control group or the *Funnix*

Table 2, continued

Descriptive Statistics, t-tests, and Effect Sizes

Nonsense Word Fluency (NWF)

	M	*SD*	*n*	*t*	*p*	Cohen's *d*
Control Group—Winter	3.6	5.8	19	2.07	0.03	0.80
Funnix Group—Winter	13.9	19.6	17			
Control Group—Spring	4.7	6.7	19	2.13	0.02	0.85
Funnix Group—Spring	20.8	31.5	18			

100-Word Test

	M	*SD*	*n*	*t*	*p*	Cohen's *d*
Cotrol Group—Spring	5.5	12.3	19	1.60	0.06	0.56
Funnix Group—Spring	16.2	25.7	18			

Controlled Text Passages

	M	*SD*	*n*	*t*	*p*	Cohen's *d*
Control Group—Spring	5.4	17.6	19	1.42	0.08	0.51
Funnix Group—Spring	20.3	41.2	18			

Note. The t-tests and effect sizes were computed for each testing period and each measure. For example, the t-test and effect size for fall LNF compares the average LNF raw score of the control group and the *Funnix* group for that time period. All probabilities are one-tailed.

group would be considered at risk based on their LNF scores. Similarly, none of the *Funnix* students and only one of the control students would be considered at risk given their ISF scores. The results are strikingly different with the measures of NWF and PSF, measures which are much more closely related to actual reading. Using the mid-year kindergarten norms (the earliest that are available), over half of the control students but only a handful of the *Funnix* students (1 to 3 students) would be considered at risk of not meeting literacy goals given these measures.

The scores on the BLCT also may be compared to established norms (Engelmann, et al., 1982, p. 49). Recall that scores were not available for most of the children for the fall administration because very few children could complete the test. A majority of the children completed the test by winter, and all of the children completed the test by spring. At both the winter and spring administrations, a majority of the children in the *Funnix* group had scores above the median of the normative distribution for their age group, and half of the *Funnix* students scored in the top quartile (twice the proportion that would be expected). One-third of the control students scored above the median at the winter testing and almost two-thirds scored above the median at the spring testing. None of the control students scored in the top quartile at the winter testing, while four students (slightly more than one-fifth of the group) did so in the spring.

Multivariate Analyses

Table 3 reports the results of the multivariate analyses. Even though students were randomly

Table 3

Multivariate Analyses

Repeated Measures Analyses of Variance

	Time		Condition		Interaction	
	F	*p*	*F*	*p*	*F*	*p*
BLCT—Total	26.268	<0.001	7.49	0.01	5.237	0.03
LNF	39.766	<0.001	5.36	0.03	2.199	0.13
ISF	50.595	<0.001	4.13	0.05	6.018	0.02
PSF	24.334	<0.001	44.13	<0.001	7.305	0.01
NWF	6.176	0.02	5.33	0.03	3.68	0.06

Analyses of Covariance

	LNF-Fall		Condition		Interaction	
	F	*p*	*F*	*p*	*F*	*p*
100-Word Test	12.21	0.001	0.44	0.51	4.2	0.05
Controlled Text Passages	9.63	0.004	0.55	0.46	3.9	0.06

assigned to treatment condition, these analyses are arguably more accurate than those presented in Table 2 because they control for children's initial levels of achievement before exposure to the curriculum. The first panel of Table 3 reports the results of the repeated measures ANOVA, with scores at each time period as repeated measures and condition (control group and *Funnix*) as the factor. The F values associated with time (the repeated measures) and condition are significant in all five analyses, and the F values associated with the interaction effect are significant in all but the analysis of LNF (using the more conservative level of 0.10 to determine significance for an interaction effect).

These results indicate changes in the scores over time were significant but that, for most variables, the pace of change varied significantly between the two groups. As can be seen from Table 2, the changes in the DIBELS measures were substantially greater for the *Funnix* group than for the control group. From the fall to the spring administration, the average LNF score increased by 25 points for the *Funnix* students but only by 16 points for the control students. From fall to winter the average ISF score increased by 12 points for the *Funnix* students, but only 6 points for the control students. Differences in changes in PSF and NWF from winter to spring were even stronger: a 14-point increase in average PSF scores for the *Funnix* students compared to only a 4-point increase for the control students; and a 7-point increase in average NWF for *Funnix* students, but only a 1-point average increase for the control students.

Although the ANOVA results with the BLCT are similar to the other results, the changes that underlie the F scores were slightly different. At both the winter and spring testing periods the *Funnix* students had fewer errors but the significant interaction effect resulted from greater change in the control group over time. The greater change within the control group may reflect both a regression toward the mean for the control group and a ceiling effect for the *Funnix* students. As noted above, by the winter testing the *Funnix* group was performing well above the expected level for 4 year olds, while the control group was not.

The results of the ANCOVA are shown in the second panel of Table 3. A significant interaction reveals a difference between groups' fall LNF scores and spring reading scores. Inspection of the data indicated that this result reflects a much smaller association of fall LNF with oral reading for the control students than for the *Funnix* students. Within the control group, the majority of children had very low spring scores, whether or not they would be considered at low risk given their fall LNF scores, and the correlation of fall LNF scores and spring oral reading scores was nonsignificant ($r = 0.29$, $p = 0.23$ for the 100-Word Test and $r = 0.26$, $p = 0.28$ for the controlled text passages). In contrast, the majority of students in the *Funnix* group, no matter what their fall LNF scores, had higher scores on the spring measures of oral reading than those in the control group, and there was a significant correlation of fall LNF scores and spring oral reading scores ($r = 0.67$, $p < 0.001$, for the 100-Word Test and $r = 0.62$, $p < 0.001$ for the controlled text passages).

Controlling for Race-Ethnicity and Home Language

Table 4 summarizes the results when minority status is controlled. The top panel of Table 4 gives the average scores on each measure at the spring testing for students in the control and *Funnix* group for minority and nonminority students, and the bottom panel gives the results of the associated two-way ANOVA for each measure. The results generally confirm those obtained through other analyses. Within the two groups based on minority status, the students in the *Funnix* condition outperformed students in the control group in all but one comparison. They had fewer errors on the

BLCT, higher scores on the three DIBELS measures, and higher scores on the 100-Word Test. With the controlled text passage scores, the differences between *Funnix* and control students were quite large among the minority students but were near zero for the nonminority students. Inspection of the raw data indicated that this reflected a very high score of one student within the control group, who had a score of 76. No other student in the control group (minority or nonminority) scored over 14, while 6 students in the *Funnix* group had scores above that level.

The ANOVA results in the bottom panel of Table 4 confirm these results. Even with the very small sample size and the reduced degrees of freedom with the addition of the control variable, 4 of the 8 differences between the groups were significant at the

Table 4

Spring Scores by Group and Minority Status

Mean Scores by Group and Minority Status

	Minority Students		Nonminority Students	
	Control	Funnix	Control	Funnix
Spring BLCT (errors)	23.4	17.8	21.1	18.3
Spring LNF	19.6	38.6	31.4	36.6
Spring PSF	8.0	32.4	8.4	31.6
Spring NWF	4.1	24.4	5.5	15.3
100-Word Test	3.0	19.1	9.0	11.7
Controlled Text Passage	1.2	26.8	11.0	10.1
n	11	11	8	7

Two-Way Analysis of Variance Results

	Group		Minority Status		Interaction	
	F	*p*	*F*	*p*	*F*	*p*
BLCT (errors)	1.95	0.17	0.09	0.77	0.20	0.66
LNF	4.96	0.03	0.79	0.38	1.61	0.21
PSF	41.44	<0.001	0.00	0.96	0.03	0.88
NWF	3.84	0.06	0.25	0.62	0.47	0.50
100-Word Test	1.92	0.18	0.01	0.92	0.97	0.33
Controlled Text Passage	1.36	0.25	0.11	0.75	1.56	0.22

Note: Nonminority students are non-Hispanic whites whose families speak English at home. All scores are from the spring administration.

0.05 level, and an additional comparison was significant at the 0.06 level.

Matched Samples

Even though the differences between the two groups on the measures used as pretests in the fall were not statistically significant, the *Funnix* students had slightly higher scores than the control students, and the Cohen's d value associated with the fall LNF score would be considered educationally significant (0.39). The multivariate tests reported above adjust for these differences statistically, but to provide additional controls we examined a reduced sample of students matched on fall scores, gender, and race-ethnicity.

This reduced sample included 24 of the 37 students. Two of the pairs (4 students) were boys, while the remaining pairs were girls. In seven of the pairs both children were white, in four pairs both children were African American, and in one pair the *Funnix* child was African American while the child in the control group was Asian American. Children in the control group of this reduced sample had slightly higher LNF and ISF scores in the fall, although the t-tests were not significant and the d values fell shy of the level generally considered educationally significant. This result reverses that obtained for the total sample, in which students in the *Funnix* group had slightly higher fall scores.

As hypothesized, the winter and spring scores indicate a consistent advantage for students in the *Funnix* group. *Funnix* students made fewer errors on the BLCT, had higher scores on the DIBELS measures of beginning literacy, and read more words correctly in the tests of oral reading. Even with the very small sample size the differences in the spring scores were statistically significant (one-tail test) on 3 of the 6 comparisons. All but one of the effect sizes met or surpassed the usual criterion of educationally important, ranging from 0.14 (for the 100-Word Test) and 0.25 (for the BLCT) to 1.67 (for phoneme segmentation fluency). The average effect size for the differences in spring scores was 0.64 for the reduced sample compared to 0.91 for the total sample.

We also examined results for students in two of the matched pairs, one with high fall scores and one with low fall scores. The students in the high-scoring pair were both white females. At the fall testing, their LNF scores were quite close (24 for the control student and 26 for the *Funnix* student), but the student in the control group had a markedly higher ISF score (19 compared to 7 for the *Funnix* student). By the winter testing period the *Funnix* child had markedly higher scores on all of the measures. She scored 32 on ISF, 43 on LNF, 33 on PSF, and 25 on NWF, compared to scores, respectively, of 16, 22, 14, and 0 for the child in the control group. At the spring testing the girl in the *Funnix* group had markedly fewer errors on the BLCT (9 versus 14). She also continued to have markedly higher scores on the beginning literacy measures: 42 on LNF, 46 on PSF, and 9 on NWF, compared to 37, 11, and 0 for the girl in the control group. The differences in the two measures of oral reading are the most striking. The child in the *Funnix* group easily read both isolated words and connected text, compared to the child in the control group who read only 3 words, including "a" and "I" (100-Word Test score of 37 and controlled text passages score of 31 for the *Funnix* student compared to scores of 3 and 0 for the control student).

The two children matched for their low pretest scores were African American males. Both children had scores of zero on the fall administration of the LNF and ISF measures. However, during winter testing, the boy in the *Funnix* group had markedly higher scores than the boy in the control group on three of the DIBELS measures: 15 on ISF, 21 on LNF, and 13 on PSF, compared to 9, 5, and 0; yet the *Funnix* child scored 0 on NWF, while the boy in the control group scored 3. During spring testing the child in the *Funnix* group had markedly higher scores on LNF, PSF, and NWF: 24, 25, and 10 compared to 0, 0, and 1 for the boy in the control group. Both boys, however, had very low scores on the two oral reading measures. Thus, even though the boy in the *Funnix* group had not reached the point of independent reading by spring, he was much better prepared to succeed in kindergarten than the boy in the control group.

Summary and Discussion

This paper examined the relationship of instruction in the computer-based *Funnix Beginning Reading* program to the development of beginning reading skills. The study employed a pretest-posttest control group design with Head Start students from a suburban community in the southern United States. Students from six different classrooms were randomly assigned to the control or experimental group. Those in the control group received 30 minutes of additional instruction each day in their usual language arts curriculum. All instruction for students in the control group was provided by their classroom teachers and teaching assistants. Students in the experimental group received 30 minutes of instruction with *Funnix Beginning Reading*. Instruction for the *Funnix* group was provided by high school-aged tutors, who were trained and supervised by an experienced teacher.

Pretesting indicated no significant differences between the two groups in beginning literacy skills. However, by winter and spring the students in the *Funnix* group had significantly higher scores on numerous measures of beginning literacy. These results occurred with simple comparisons of means, comparisons of scores to established benchmarks, and multivariate analyses that controlled for minority status and for initial levels of skill. The results also appeared with a reduced sample of individually matched children on their pretest scores. Two case-wise comparisons of children with similar initial skill levels illustrated the magnitude of these changes.

In general, the results indicate that 4-year-old children in a Head Start program can develop strong beginning literacy skills with instruction in *Funnix Beginning Reading*. By the end of the academic year the vast majority of all the students in the study—both those who received enhanced instruction in their regular Head Start curriculum and those in the *Funnix* group—had expertise in letter naming and knowledge of initial sounds that would bode well for their future success. In addition, a large proportion of the children in the *Funnix* program acquired skills much closer to true beginning reading, with significantly higher scores on the DIBELS measures of NWF and PSF as well as higher scores on the two text-reading measures. Only one of the *Funnix* students would be considered at risk of later literacy problems based on the spring PSF score, and only three of these students would be considered at risk based on their NWF scores. Note that both of these measures of "at risk" are based on norms developed for kindergarten students at mid-year, fully 9 months after the testing period for the Head Start students.

The high scores of the *Funnix* students on the test of basic language concepts (BLCT) also illustrate this superior achievement. By definition, one would expect students' scores to be equally distributed across the percentiles developed through the testing norms (e.g., 25% below the 25th percentile, etc.). By spring testing, students in the control group had a distribution similar to this expectation, and slightly more students than expected had scores above the median (12 versus 9 to 10). A slightly larger proportion of the *Funnix* students were above the median, but, even more striking, half of the *Funnix* students scored in the top quartile—twice the proportion that would be expected by chance. A close inspection of the data suggests that these increases in language skills occurred soon after beginning *Funnix* instruction. At winter testing, half of the *Funnix* students, but none of the control students, scored in the top quartile. This suggests that instruction in *Funnix* contributed to both beginning literacy and general language development.

Even though the sample size was relatively small, most results were statistically significant. In addition, almost all effect sizes were quite large, well beyond the level (0.25) traditionally cited as educationally important. The fact that students were randomly assigned to treatment and came from several different classrooms enhances the internal validity of the findings. The use of multivariate statistics and replication of results with a smaller, closely matched sample also help to validate the findings.

The results obtained in this study largely replicate findings obtained in other studies of *Funnix Beginning Reading*. Like Parlange (2004) and Watson and Hempenstall (2008), these results indicate significantly stronger gains in pre- and beginning reading skills among *Funnix* students than among the control students. The results are especially notable given the random assignment of students to conditions and the use of high school-age tutors, rather than college students or parents, as employed in other studies.

It is important to continue examination of the *Funnix* program. Future studies should include larger samples, samples of different ages from other areas of the country, and compare adult volunteers, teen volunteers, teacher aides, and parents as tutors. Finally, it would be important to examine factors related to students' pace of completing the program. Some students in the *Funnix* group progressed very rapidly through the lessons, while others had slower progress. Factors that could explain these variations might include those related to children's initial skills, their English proficiency, characteristics of the tutor, and the relationship between the tutor and child.

Other preschool programs, funded by Head Start or with other sources of funding, could potentially learn from these results. They illustrate the ways in which low-income students can develop strong beginning literacy skills that provide a solid foundation for early reading. The *Funnix* program was implemented in a low-cost manner, using high school volunteers and involving only 30 minutes a day of additional instruction. Introducing such a program could involve a relatively minor alteration in a preschool schedule and potentially utilize volunteers who are already active and committed. In addition, our results suggest the *Funnix* program would fulfill the mandates of the National Reading Panel and of the 2007 Improving Head Start for School Readiness Act calling for research-based curricula.

Previous work has indicated the importance of high fidelity of implementation in producing strong achievement gains for students in DI programs (Carlson & Francis, 2002; Gersten et al., 1986). Key elements of implementation involve the sequencing, pacing, and presentation of material and the way feedback is provided. As noted above, the *Funnix* program controls these factors and helps ensure these critical elements are presented as designed. In addition, there was strong support for the tutors with onsite guidance, regular reporting procedures, and consultations for difficulties. While not extraordinarily expensive, these supportive measures helped promote the smooth operation of the tutorial program. Including such support would be important to help promote success.

Finally, the potentially positive impact of the program on the high school tutors should not be ignored. Interviews with supervisory personnel indicate tutors found the program rewarding and satisfying. Providing recognition to the students, both at the preschool site and at their high school, helped make the experience prestigious among the high school peers and promoted commitment. Many high schools are located close to preschools and encourage their students to provide community service. In addition, high school tutors may benefit from implementation of the program described in this article (B. Primm, personal communication, March 11, 2009).

References

Al Otaiba, S., Schatschneider, C., & Silverman, E. (2005). Tutor-assisted intensive learning strategies in kindergarten: How much is enough? *Exceptionality, 13*(4), 195–208.

Allor, J., & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research and Practice, 19*(2), 116–129.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children, 5*(3), 25–50.

Benner, G. J. (2007). The relative impact of remedial reading instruction on the basic reading skills of students with emotional disturbance and learning disabilities. *Journal of Direct Instruction*, 7, 1–15.

Benner, G. J., Kinder, D., Beaudoin, K. M., Stein, M., & Hirschmann, K. (2005).The effects of the "Corrective Reading Decoding" program on the basic reading skills and social adjustment of students with high-incidence disabilities. *Journal of Direct Instruction, 5*, 67–80.

Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American schools after a decade.* Santa Monica, CA: Rand.

Blakeley, M. R. (2001). A survey of levels of supervisory support and maintenance of effects reported by educators involved in Direct Instruction implementations. *Journal of Direct Instruction, 1*(2), 73–83.

Bodilly, S. J. (1998). *Lessons from New American Schools' Scale-up Phase.* Santa Monica, CA: Rand.

Bodilly, S. J., Glennan, T. K., Jr., Kerr, K. A., & Galegher, J. R. (2004). Introduction: Framing the problem. In T.K. Glennan, Jr., S. J. Bodilly, J. R. Galegher, & K. A. Kerr (Eds.), *Expanding the reach of education reforms: Perspectives from leaders in the scale-up of educational interventions* (pp. 1–35). Santa Monica, CA: Rand.

Carlson, C. D., & Francis, D. J. (2002). Increasing the reading achievement of at-risk children through Direct Instruction: Evaluation of the Rodeo Institute for Teacher Excellence (RITE). *Journal of Education for Students Placed at Risk*, 7, 141–166.

Chall, J. S. (1983). *Stages of reading development.* New York: McGraw-Hill.

Cole, J. M. & Hilliard, V. R. (2006). The effects of web-based reading curriculum on children's reading performance and motivation. *Journal of Educational Computing Research, 34*, 353–380.

Cooke, N. L., Gibbs, S. L., Campbell, M. L., & Shalvis, S. L. (2004). A comparison of Reading Mastery Fast Cycle and Horizons Fast Track A-B on the reading achievement of students with mild disabilities. *Journal of Direct Instruction, 4*, 139–151.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934–945.

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator, Spring/Summer*, 1–8.

Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *The American Economic Review, 85*, 341–364.

Currie, J., & Thomas, D. (2000). School quality and the longer-term effects of Head Start. *The Journal of Human Resources, 35*, 755–774.

DIBELS (2002). Assessment Committee Analysis of Reading Assessment Measures: Coding Form. Retrieved December 2, 2009, from https://dibels.uoregon.edu/techreports/dibels_5th_ed.pdf

Durkin, D. (1974-75). A six year study of children who learned to read in school at the age of four. *Reading Research Quarterly, 10*(1), 9–61.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167–188.

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly, 14*, 135–163.

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

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's Meta-Analysis. *Reading Research Quarterly, 36*(3), 250–287.

Engelmann, S., Ross, D., & Bingham, V. (1982). *Basic Language Concepts Test.* Chicago: Science Research Associates.

Engelmann, S. E., & Engelmann, K. E. (2004). Impediments to scaling up effective comprehensive school reform models. In T. K. Glennan, Jr., S. J. Bodilly, J. R. Galegher, & K. A. Kerr (Eds.), *Expanding the reach of education reforms: Perspectives* *from leaders in the scale-up of educational interventions* (pp. 107–133). Santa Monica, CA: Rand.

Flores, M. M., Shippen, M. E., Alberto, P., & Crowe, L. (2004). Teaching letter-sound correspondence to students with moderate intellectual disabilities. *Journal of Direct Instruction, 4,* 173–188.

Francis, D. J., Stuebing, K. K., Shaywitz, S. E., Shaywitz, B.A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17.

Gersten, R., Carnine, D., Zoref, L., & Cronin, D. (1986). A multifaceted study of change in seven inner-city schools. *The Elementary School Journal, 86,* 257–276.

Good, R. H., Simmons, D. S., Kame'enui, E. J., Kaminski, R. A., & Wallin, J. (2002). *Summary of decision rules for intensive, strategic, and benchmark instructional recommendations in kindergarten through third grade* (Technical Report No. 11). Eugene, OR: University of Oregon.

Gough, P. B., & Juel, C. (1991). The first stages of word recognition. In L. Rieben & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 47–56). Hillsdale, NJ: Erlbaum.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437–447.

Lee, V. E., & Loeb, S. (1995). Where do Head Start attendees end up? One reason why preschool effects fade out. *Educational Evaluation and Policy Analysis, 17,* 62–82.

Lipson, M. Y., & Wixson, K. K. (1997). *Assessment and instruction of reading and writing disability: An interactive approach* (2nd ed.). New York: Plenum.

Lonigan, C. J., Driscoll, K., Phillips, B. M., Cantor, B. G., Anthony, J. L., & Goldstein, H. (2003). A computer-assisted instruction phonological sensitivity program for preschool children at-risk for reading problems. *Journal of Early Intervention, 25,* 248–262.

Malmgren, K. W., & Leone, P. E. (2000). Effects of a short-term auxiliary reading program on the reading skills of incarcerated youth. *Education and Treatment of Children, 23,* 239–247.

Morris, D., Bloodgood, J., & Perney, J. (2003). Kindergarten predictors of first- and second-grade reading achievement. *The Elementary School Journal, 104,* 93-109.

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

O'Brien, D. M., & Ware, A. M. (2002). Implementing research-based reading programs in the Fort Worth Independent school district. *Journal of Education for Students Placed at Risk, 7,* 167–195.

Parlange, L. A. (2004). *The effects of* Funnix Beginning Reading *program on the beginning reading skills of preschoolers.* Unpublished manuscript, Eastern Washington University.

Riepl, J. H., Marchand-Martella, N. E., & Martella, R. C. (2008). The effects of Reading Mastery Plus on the beginning reading skills of students with intellectual and developmental disabilities. *Journal of Direct Instruction, 8,* 29–39.

Ritter, C., Denny, G., Albin, G., Barnett, J., & Blankenship, V. (2006). The effectiveness of volunteer tutoring programs: A systematic review. *Campbell Systematic Reviews* 2006:7.

Segal-Drori, O., Korat, O., Shamir, A., & Klein, P. S. (2009). Reading electronic and printed books with and without adult instruction: Effects on emergent reading. *Reading and Writing, May,* 1–18.

Simmons, D. C., & Kame'enui, E. J. (Eds.) (1998). *What reading research tells us about children with diverse learning needs: Bases and basics.* Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–407.

Tobin, K. G. (2003). The effects of the Horizons reading program and prior phonological awareness training on the reading skills of first graders. *Journal of Direct Instruction, 3,* 1–16.

Tobin, K. G. (2004). The effects of beginning reading instruction in the Horizons reading program on reading skills of third and fourth graders. *Journal of Direct Instruction, 4,* 129–137.

Tobin, K. G. (2009). A comparison of two reading programs on the reading outcomes of first-grade students. *Journal of Direct Instruction, 9,* 35–46.

U.S. House of Representatives (2007). *Improving Head Start for School Readiness Act of 2007, H.R. 1429.* Washington, DC: U.S. House of Representatives.

Vadasy, P. F., Jenkins, J. R., Antil, L. R., Wayne, S. K., & O'Connor, R. E. (1997). The effectiveness of one-to-one tutoring by community tutors for at-risk beginning readers. *Learning Disability Quarterly, 20,* 126–139.

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring. *Journal of Learning Disabilities, 38*(4), 364-380.

Vitale, M. R., & Joseph, B. L. (2008). Broadening the institutional value of Direct Instruction implemented in a low-SES elementary school: Implications for scale-up and reform. *Journal of Direct Instruction, 8*, 1–18.

Watson, T., & Hempenstall, K. (2008). Effects of a computer based beginning reading program on young children. *Australasian Journal of Educational Technology, 24*, 258–274.

Wharton-McDonald, R., Pressley, M., & Hampston, J. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *The Elementary School Journal, 99*, 101–128.

Author's Notes

This project could not have been completed without the dedication and hard work of Betsy Primm and Tracy Britt in their onsite supervision of the project and their concern for the welfare of both the preschoolers and high school students involved in the project. We also thank the officials of the Buford, GA, Head Start organization for allowing the study to proceed and Michael Rebar for compiling the data that are analyzed in this report. Special thanks is owed to Ms. Linda Murray at Buford High School. Finally, we thank Betsy Primm, Bonnie Grossen, Siegfried Engelmann, and Craig Darch for their comments on earlier drafts of this paper; Owen Engelmann and Kurt Engelmann for providing material regarding Direct Instruction curricula; and Cristy Couglin for summaries of the research literature on Direct Instruction. Any errors that remain are the responsibility of the author.