Effects of Tutoring in Phonological and Early Reading Skills on Students at Risk for Reading Disabilities

Patricia F. Vadasy, Joseph R. Jenkins, and Kathleen Pool

Abstract

This study examined the effectiveness of nonprofessional tutors in a phonologically based reading treatment similar to those in which successful reading outcomes have been demonstrated. Participants were 23 first graders at risk for learning disability who received intensive one-to-one tutoring from noncertified tutors for 30 minutes, 4 days a week, for one school year. Tutoring included instruction in phonological skills, letter–sound correspondence, explicit decoding, rime analysis, writing, spelling, and reading phonetically controlled text. At year end, tutored students significantly outperformed untutored control students on measures of reading, spelling, and decoding. Effect sizes ranged from .42 to 1.24. Treatment effects diminished at follow-up at the end of second grade, although tutored students continued to significantly outperform untutored students in decoding and spelling. Findings suggest that phonologically based reading instruction for first graders at risk for learning disability can be delivered by nonteacher tutors. Our discussion addresses the character of reading outcomes associated with tutoring, individual differences in response to treatment, and the infrastructure required for nonprofessional tutoring programs.

Research on students at risk for learning disability suggests that early, explicit instruction in phonological and decoding skills can help these students stay on track to successful reading acquisition (Adams, 1990; Blachman, 1994; Fielding-Barnsley, 1997; Juel, 1988; Liberman, Shankweiler, Blachman, Camp, & Werfelman, 1980; Spear-Swerling & Sternberg, 1996; Torgesen, Wagner, & Rashotte, 1994; Williams, 1985). Indeed, such instruction may be critical for first graders who exhibit very low literacy skills because they are at serious risk for developing long-term problems in learning to read (Juel, 1988; Vellutino et al., 1996).

Early in first grade, most teachers can identify children who need more intense, individual instruction to acquire word-level reading skills. However, competing demands on classroom teachers' time usually preclude giving the students the level of individualized instruction they require. This situation has led to the development of supplemental tutoring programs designed to forestall reading problems. Some of these early intervention programs use highly trained teachers as tutors (Clay, 1985), whereas others rely on paraprofessionals and volunteers (Invernizzi, Juel, & Rosemary, 1997).

The research on tutoring provides limited information on the tutor qualifications associated with successful reading interventions. Wasik and Slavin's (1993) review of five first-grade tutoring programs concluded that "programs using certified teachers as tutors appeared to obtain substantially larger impacts than those using paraprofessionals" (p. 196), with effect sizes (ESs) ranging from +.55 to +2.37 for the former versus ESs from +.20 to +.75 for the latter. However, because the tutors' educational background in these studies was confounded with several potentially important factors—that is, the programs' theoretical orientation; components of reading emphasized (e.g., prior knowledge, reading strategies, and error correction strategies); coordination with the child's classroom program; latitude given tutors to make instructional decisions; and duration and frequency of tutoring—firm conclusions about the relative effects of teachers and volunteers are not possible. Nevertheless, some program differences were likely a function of tutor qualification (e.g., the latitude given tutors to employ a variety of strategies in addressing individual student needs vs. utilizing explicitly scripted or programmed instruction materials).

If achievement effects were at all comparable for teacher-tutors and noncertified tutors, cost factors would clearly favor using noncertified tutors. Many schools use volunteer tutors to help struggling readers, and recent federal programs (e.g., America Reads) have encouraged schools to pursue this strategy. Despite the face value of one-to-one help, daily tutoring by noncertified teachers, even for an entire

school year, does not guarantee improved reading achievement (Ellson, Harris, & Barber, 1968). One reason for uncertainty about the effectiveness of noncertified tutors can be traced to individual differences in their teaching skills. Juel (1996) analyzed video and audio recordings of tutoring sessions in which college athletes tutored first-grade children in reading. In examining the performance of more and less successful tutors (as defined by the achievement of their tutees), Juel noted a relationship between children's reading growth and the character of tutor–student verbal interactions. Tutors whose children showed larger reading gains provided significantly more scaffolded reading and writing experiences and explicit cognitive modeling of reading and writing. Juel also found that the amount of time tutors gave to specific activities was significantly related to children's reading growth: Spending more time on letter–sound and word-reading activities was associated with larger reading gains.

Juel's (1996) findings suggest ways that schools might structure the work of volunteer tutors so that their effects better approximate those of teacher-tutors (i.e., designing tutoring lessons that target critical early reading skills and incorporate explicit modeling and response-contingent scaffolding). Beyond these ideas, Wasik (1998) suggested additional guidelines for strengthening volunteer-tutor programs: tutor training and supervision by reading specialists, consistent and intensive teaching, use of quality reading materials, ongoing assessment of tutees, consistent attendance by tutors, and coordination of the program with classroom instruction. Nevertheless, the research base on factors that contribute to the efficacy of volunteer-tutor programs remains thin.

Over the last several years we have been developing and testing a tutoring system called Sound Partners (Jenkins, Vadasy, Firebaugh, & Profilet, in press; Vadasy, Wayne, O'Connor, Jenkins, & Pool, 1998), which uses nonprofessional tutors and a structured program to provide intensive early literacy instruction for children who need more focused and individualized help than classrooms typically provide. The program consists of 100 lessons on phonological awareness, letter–sound activities, word identification, text reading, and writing. Tutors are recruited from the school community and trained to work with individual students for one school year. Instruction is systematic and explicit, combining phonemic awareness, phonics, and gradually increasing amounts of reading time—features that are associated with successful early reading interventions (Foorman, Francis, Beeler, Winikates, & Fletcher, 1997; Torgesen, Wagner, Rashotte, Alexander, & Conway, 1997).

In the 2 years prior to the study reported here, we tested versions of the treatment, randomly assigning at-risk first graders to tutoring and nontutoring control groups. Tutoring occurred 4 days a week, one half hour per day, for the school year. Results from the first year (Vadasy, Jenkins, Antil, Wayne, & O'Connor, 1997a) showed that tutored students outperformed students in the control group on segmentation and spelling posttests, with nonsignificant but positive effect sizes in word recognition (.31) and nonword reading (.34).

In a second experiment, children who received a revised version of the treatment (additional instruction in decoding, long-vowel words, and word endings, along with different text selections) performed significantly higher than controls in nonword reading and spelling (Vadasy, Jenkins, Antil, Wayne, & O'Connor, 1997b). A post hoc analysis of these data revealed a relation between children's achievement and the quality of tutoring they received: Students whose tutors were "high implementors" (i.e., consistently followed the lesson formats) scored significantly higher than students of low implementors and students in the control group in word reading, nonword reading, and spelling, with ESs that averaged .83.

The present study reports the results of the third iteration of tutoring refinements, following up on two problems that surfaced in the previous field tests: one involving variability in tutors' implementation and one involving instructional content. We addressed the first concern, variability in tutors' implementation of program protocols, through increased training and supervision. Regarding the second problem, having identified various stumbling blocks in teaching and learning in previous renditions of the program (e.g., the lack of a soundingout routine), we sought to test the efficacy of our revised lessons.

In addition, we examined two other issues. The first was the permanence of tutoring effects. If schools are to mount tutoring programs, it is important to determine not only whether such efforts result in reading improvements by year's end, but whether effects are sustained beyond the year of intervention. Only a few first-grade literacy interventions have followed students beyond the year in which they were treated (e.g., Madden, Slavin, Karweit, Dolan, & Wasik, 1993; Shanahan & Barr, 1995). In the present study, we followed children through completion of second grade.

The second issue involved the character of the reading outcomes produced by our treatment. Torgesen, Wagner, and Rashotte (1997) recently spelled out a set of implicit assumptions underlying early intervention programs that focus on developing phonological reading skill (i.e., using knowledge of the alphabetic principle to decode unfamiliar words). According to Torgesen et al., phonological reading skill stands at the base of a reading skills hierarchy. For children to reach a point in their reading development where they can *independently* learn words, they must first develop phonological reading skills (Share, 1995; Share & Stanovich, 1995). Phonological reading skill allows the developing reader to engage in independent learning trials, which in turn help to forge representations of words in memory, as amalgamations of word-specific orthographic and

phonological information (Ehri, 1980; Ehri & Wilce, 1985). Well-developed orthographic–phonological representations of words in memory are required to reach the next level in the reading hierarchy (i.e., automatic, fluent word recognition). Finally, the ability to process print automatically and effortlessly frees attentional resources for comprehension, which sits at the top of the reading hierarchy (Samuels & Flor, 1997). Torgesen et al. reminded us that the expected relationship between phonological reading skill and word knowledge and fluency has only modest empirical support. In light of questions about the permanence and character of treatment effects, we retested students on phonological, word identification, fluency, and spelling tasks at the end of second grade, 1 year after the termination of treatment.

Method

Participants

First-grade children were selected from four elementary schools in a large urban school district. In September, teachers were asked to review their class lists and select up to six students they feared would not learn to read by the end of the year. Teachers from 11 classrooms identified a total of 64 students, who were administered pretests. Via pretest scores, those students were rank ordered on the following four key measures: letter names, WRAT-R Reading raw score, WRAT-R Spelling raw score, and PPVT-R raw score (see Pretest section). Next, we randomly assigned the 46 students who scored lowest on most of these variables to treatment or control groups, then assigned the remaining students to a replacement group. Students in the control group received the schools' regular (i.e., classroom instruction and Title I services) reading instruction. Because some children moved within the first 2 months of tutoring, we replaced two students in the treatment and two students in the control group.

All of the participants except two were from three schools that served a large proportion of students from minority and low-income backgrounds. The student populations of these three schools were 67% minority, and 47% were eligible for free or reduced-price lunch. The sample included 23 first graders in the treatment group and an equal number in the control group. There were 9 girls and 14 boys in each group. Twenty students in the tutored group and 21 students in the control group were members of minority groups. None of the students were identified for special education at entry to the program, as this does not typically occur in this district until third or fourth grade. During the year, Title I services were provided to 14 students in the tutored group and 18 students in the untutored group. Title I service delivery varied across schools. In one school, funds were used to reduce class sizes across the school; in the other schools, Title I provided pull-out or in-class small-group instruction.

Pretests

As Vellutino, Scanlon, and Tanzman (1994) noted, various tasks are used to evaluate phonological skills. Most of these tasks, however, do not have documented psychometric properties or norms. Like most researchers, we used tasks that were widely known and a variety of formats to assess phoneme analysis. Measures of naming rate were administered at pretest only, as we did not expect the intervention to influence these underlying processing capacities (Blachman, 1994).

Peabody Picture Vocabulary Test–Revised (PPVT-R; Dunn & Dunn, 1981). The PPVT-R is an individually administered, norm-referenced test of receptive vocabulary. From four simple black-and-white illustrations, children select the picture that best illustrates the meaning of a stimulus word presented orally by the tester. One point is awarded for each correct response. Age-based standard scores were used.

Rapid Letter Naming. Presented with a card displaying the uppercase letters in random order, students name as many letters as they can during 1 minute. Letter sounds are also accepted as correct responses. The score is the number of letters named per minute.

Wide Range Achievement Test–Revised (WRAT-R) Reading Subtest (Jastak & Wilkinson, 1984). The WRAT-R is an individually administered, norm-referenced achievement test of basic skills. The Reading subtest consists of letters and words that the child is asked to name. The number of words and letters correctly identified is transformed to an age-based standard score.

Rapid Automatized Naming (RAN; Catts, 1993). Presented with a chart containing pictures of animals (e.g., a pig, a cow, a horse) in three different colors (red, blue, and black), students name the animals and their colors as rapidly as they can. The time required to name all 24 items is the score.

Sound Repetition (O'Connor, Jenkins, & Slocum, 1995). Students listen to 12 items consisting of two to four phonemes each. Items are presented with a 1-second delay between phonemes, and, after a 2-second delay, students repeat the sounds. The score is the number of phonemes correctly repeated.

Modified Rosner. Students are given a version of the Rosner Test of Auditory Analysis (Rosner, 1979), modified by Berninger, Thalberg, DeBruyn, and Smith (1987), in which they segment multisyllable words by deleting one syllable. Five items require deletion of the initial syllable, and five items require deletion of the last syllable. The score is the total items correctly segmented.

Segmenting Sounds (O'Connor et al., 1995). Students listen to 10 words consisting of two to three phonemes each. The examiner models onset-rime seg-

mentation and asks the student to repeat each word in an onset-rime format. Students receive 1 point for each correctly segmented portion of the word (1 to 2 points for onset-rime segmentation and up to 3 points for segmentation into three phonemes). The score is the total number of onset-rimes and phonemes segmented.

WRAT-R Spelling Subtest. The Spelling subtest requires the examinee to copy marks, print his or her name, and print a list of dictated words. The number of items correct is transformed to an age-based standard score.

Alphabet Writing (Berninger, 1990). Students write the alphabet in lowercase letters. Capital letters, omissions, additions, transpositions, and reversals count as errors. This task is scored as the number of correct letters written in the first 15 seconds, as well as total time and total correct.

First-Grade Posttests

Students in the treatment and control groups were posttested on a variety of norm- and criterion-referenced measures assessing phonological, word reading, passage reading, and spelling skills. The Reading and Spelling subtests of the WRAT-R were readministered along with the following tests.

Woodcock-Johnson Psycho-Educational Battery–Revised (WJ-R), Word Attack Subtest (Woodcock & Johnson, 1989). For this test, the examinee pronounces pseudowords that increase in difficulty. One point is awarded for each correct response, and the number of correct items is transformed into age-based standard scores.

Bryant Pseudoword (Bryant, 1975). A list of 50 pseudowords is read until five consecutive items are missed. One point is assigned to each correct response.

Dolch Word Recognition Test (Dolch, 1939). The student reads from a list of 220 short, frequently used words arranged in groups according to basal reading levels, until 10 consecutive items are missed. The score is the total number of words correctly identified.

Analytical Reading Inventory (ARI; Woods & Moe, 1977). This is a criterion-referenced, individually administered test of oral passage reading. Both primer and first-grade passages were administered at the end of first grade. Testers record oral reading fluency (time and accuracy). The score is a rate measure of number of words correctly read per minute.

Yopp-Singer Segmentation Task (Yopp, 1988). Students segment sounds of 22 orally given words with corrective feedback. Testing continues until students miss 10 consecutive items, and the score is the total number of words segmented correctly.

Curriculum-Based Spelling List. Ten words taken from the storybooks used in the lessons compose a written spelling test. One point is awarded for each word spelled correctly.

Writing Sample–Spelling (Deno, 1985). Students write for 5 minutes in response to a prompt ("It was a dark and stormy night"). The writing score is the number of words correctly spelled.

Second Grade Follow-Up

Students were retested on the WJ-R Word Attack subtest and the ARI (using a Grade 2 passage). In addition, we administered the WJ-R Word Identification subtest and the Test of Written Spelling, described below.

WJ-R Word Identification Subtest. The examiner directs the child to read from a list of increasingly difficult words. The total number of words correctly identified is transformed to an age-based standard score.

Test of Written Spelling (Larsen & Hammill, 1994). Students write words dictated by the examiner. A subtest of 50 predictable words conforming to phonics rules and a subtest of 50 words with irregular spelling were analyzed separately. Each student was scored on the total items spelled correctly on each subtest.

All measures were individually administered except the writing sample, which was administered to students in their classrooms.

Procedure

The tutoring treatment was administered for 27 weeks. Students were retested in the spring of first and second grades. Following is a description of the tutoring treatment.

Tutoring Content

There are 100 scripted lessons, each designed to last approximately 30 minutes and consisting of 5 to 10 short activities that required between 1 and 15 minutes to implement. Lessons focused on segmenting, blending, letter–sound correspondences, word families, writing with invented spelling, and reading text with controlled vocabulary—skills that have been found to be helpful in assisting children who are slow to develop reading skill (Bradley & Bryant, 1985; Juel, 1996; Pinnell, Lyons, DeFord, Byrk, & Seltzer, 1994; Slavin, Madden, Karweit, Livermon, & Dolan, 1990; Torgesen, Wagner, Rashotte, et al., 1997; Vellutino et al., 1996). The two previous versions of the program emphasized similar skills. For the current field test, several new components were added, and several components used in previous field tests were revised or expanded.

Explicit Decoding Instruction. Early lesson versions did not include specific instruction in sounding out, and tutors did not model a consistent strategy for students to apply to sound out words. Iversen and Tunmer (1993) reported accelerated progress for students receiving a modified Reading Recovery program that included explicit code instruction, and Fielding-Barnsley (1997) reported the benefits of

explicit instruction in decoding, encoding, and letter–sound correspondence. We gave weight to these findings in our program redesign, incorporating into lessons an explicit teaching routine in decoding as well as encoding skills using letter tiles.

Rime Analysis. Instruction in developing orthographic coding skills for word families may help some children develop word recognition skills (Adams, 1990; Berninger, 1990; Goswami & Bryant, 1990), in particular once they have already developed letter–sound analysis skills (Ehri & Robbins, 1992). For this field test, lessons were revised to increase opportunities to identify newly learned rime units in word lists and story-reading components.

Story Reading. Lessons provided daily practice in reading and rereading phonetically regular text selections to maximize opportunities for children to apply their developing phonological and decoding skills while constructing meaning from text (Juel & Roper/Schneider, 1985). In this field test we increased the match between the discrete reading skills taught in each lesson and those needed to read that day's story. This approach to intervention is supported by the phonological linkage hypothesis (Hatcher, Hulme, & Ellis, 1994), and by Reading Recovery's success with rereading familiar books (Clay, 1985).

Other Revisions. We revised letter–sound instruction to include many letter-pair combinations, building upon the systematic English phonologies at the letter-cluster level (Venezky, 1970) that, when taught, seem to help students with learning disabilities (Berninger et al., 1998).

To standardize tutor instruction in word endings and silent-*e* words, we added practice in these skills prior to the time that students encountered these word types in their reading. Finally, nonwords were occasionally introduced in the lessons to increase practice opportunities to identify letter pairs and words by analogy.

Tutor Recruitment, Training, and Supervision

Tutors were recruited through announcements in school newsletters, then hired as employees of the schools and paid $5 an hour for tutoring and time spent in training. At the beginning of the year, seven tutors were mothers of children in the schools, and one tutor was a father. Two tutors were replaced in the middle of the year by an unemployed actor. A certified special education teacher was also hired to tutor, and to provide us with expert feedback on instructional content.

Finally, for this field test we increased the intensity of tutor training to accommodate the added lesson activities, and to address the problem of weak implementation observed in previous years. Tutors received 8 hours of training before commencing tutoring and 6 hours of training during the school year. Initial training included explanations, modeling, and role playing of each lesson component. Tutors also received guidelines for behavior management, record keeping, and error correction strategies. Follow-up training was scheduled during the year, when project staff noted a need to review strategies, when tutors requested a review or help in a new lesson component, or when tutors reported problems in using lessons or teaching a particular skill.

Record Keeping

Tutors maintained daily logs of attendance and lesson progress as a measure of treatment intensity. According to these logs, children attended from 54 to 89 sessions, with a mean of 72 days. Because tutor and student absences reduced treatment intensity, tutors were encouraged to make up missed lessons when possible.

Fidelity of Intervention

To address our concerns about variable implementation, research staff observed each tutor at least once a week. During these 15- to 30-minute observations, project staff (Vadasy or Pool) looked for the following actions: starting lessons on time, making error corrections, following lesson formats, managing student behavior, using positive encouragement strategies, and providing a full 30 minutes of instruction. A total percentage of these six behaviors was obtained for each tutor, averaging across behaviors (reported under Results). Both observers at times observed each tutor, and they frequently compared their notes. In conjunction with the observations, tutors were often given brief written or oral feedback (e.g., suggestions for another way to teach a child having difficulty, or praise for a tutor's instructional skills). At other times, project staff modeled a strategy or adjusted a student's placement in the program (e.g., directing the tutor to go back to review previous lessons or lesson components until skills were solidly mastered, or to skip lessons when students had clearly mastered a skill and needed more challenging material).

Finally, students were tested every 10 lessons on mastery of lesson content. Project staff administered these curriculum-based tests with items drawn directly from a recently completed lesson. The mastery tests were a check on the tutor's lesson pacing and the student's acquisition of skills.

Results

On the basis of weekly observations of tutors by project staff, tutors implemented the program with a high degree of fidelity. The average of tutor scores across all observations reveals that tutors demonstrated an implementation rate of 89% on the six established criteria (e.g., conducted all lesson components according to specification). Student mastery of instructional content as measured by the curriculum-based tests administered to students every 10 lessons was high, with a mean score of 94% across students and tests. The lowest average score for an individual student was 84%.

Group means and standard deviations for the 10 pretests are shown in Table 1, along with results of one-way analyses of variance for each measure. One difference (WRAT-Spelling) was significant, favoring the treatment group.

Table 2 provides the posttest means and standard deviations, adjusted for pretests. We used a composite z of pretest scores to adjust posttests because some posttests of interest did not have that same measure administered as a pretest. A multivariate analysis of covariance indicated significant group differences, Wilks's lambda = .61, $F(10, 33) = 2.14$, $p < .05$. All univariate tests significantly favored the treatment group except the ARI first-grade level, which did not differ between treatment and control. Effect sizes (i.e., treatment and control adjusted mean differences divided by the pooled unadjusted standard deviations for treatment and control groups) ranged from .42 to 1.24. The largest effect size was observed for nonword reading (1.24), the smallest for reading in context (.42 and .60). Three measures provided norm-referenced standard scores; the mean standard score for the treatment group at posttest exceeded the 50th percentile for WRAT-R Reading and Woodcock-Johnson Word Attack, and was within 3 points of this criterion on WRAT-R Spelling.

Figure 1 shows the distribution of treatment and control students on word identification (WRAT-R Reading) and nonword reading (Woodcock-Johnson Word Attack). Twelve of 23 tutored students scored above the 50th percentile on word identification (vs. 2 of 23 students in the control group), and 15 of 23 tutored students surpassed this criterion on nonword reading (vs. 8 of 23 students in the control group). Al-

TABLE 1
One-Way ANOVA of Means of Treatment and Control Groups for All Pretreatment Means

Variable	Treatment		Control		
	M	*SD*	*M*	*SD*	*F*
Age	6.56	0.39	6.66	0.40	0.81
PPVT-R standard	83.57	16.57	84.74	13.12	0.07
Letter-naming rate	30.68	19.11	33.68	17.07	0.90
WRAT-R Reading standard	83.26	9.88	81.22	9.51	0.51
Rapid automatized naming	18.62	6.33	20.02	6.84	1.99
Sound repetition	23.00	5.70	24.62	5.52	0.95
Modified Rosner segmentation	5.70	3.01	5.30	2.74	0.21
Segmenting sounds	5.22	5.60	6.00	6.20	0.20
WRAT-R Spelling standard	81.13	8.92	75.27	10.06	4.28*
Alphabet writing	2.43	3.81	1.70	1.29	0.39

Note. For each group, $n = 23$. PPVT-R = Peabody Picture Vocabulary Test–Revised; WRAT-R = Wide Range Achievement Test–Revised.
*$p < .05$ ($df = 1, 44$).

TABLE 2
Adjusted Means and Standard Deviations of Treatment and Control Groups at Posttest

Measures	Treatment		Control				
	M	*SD*	*M*	*SD*	*F*	*p*	*ES*
Reading							
WRAT-R Reading Subtest (standard)	102.45	18.81	88.77	11.38	8.44	.006	.91
Dolch Word List	144.74	54.95	102.67	47.37	8.21	.006	.82
Analytical Reading Inventory (words per minute)							
Primary level	45.36	34.77	29.42	18.19	3.92	.054	.60
First-grade level	36.57	33.38	25.43	19.69	1.96	.169	.42
Decoding							
Woodcock-Johnson Word Attack (standard)	109.27	13.66	94.12	10.71	16.93	0.000	1.24
Bryant	19.45	11.65	8.94	7.79	12.78	0.001	1.08
Segmenting and spelling							
WRAT-R Spelling Subtest (standard)	97.33	16.60	85.30	12.67	7.24	.010	.82
Curriculum-based spelling measure	8.00	1.98	5.95	2.42	10.41	0.002	.93
Words correct (%) on writing measure	0.71	0.22	0.55	0.19	6.69	0.013	.76
Yopp-Singer Segmentation	15.51	3.79	11.15	5.53	9.89	0.003	.85

Note. For each group, $n = 23$. WRAT-R = Wide Range Achievement Test–Revised.

though the effect of treatment on both measures is readily apparent, there was nevertheless considerable variation in the children's response to treatment, with standard scores ranging from 72 to 141 in word reading and from 90 to 141 in nonword reading. Seven tutored students scored below the 25th percentile (standard score of 90) on word identification; all scored at or above the 25th percentile on nonword reading.

Follow-Up at Second Grade

We were able to locate and retest 20 treatment students and 17 students in the control group near the end of second grade. A multivariate analysis of covariance, controlling for the pretest composite *z* score, was significant, Wilks's lambda = .54, $F(9, 26) = 2.45$, $p < .05$. Univariate tests were significant for word attack and spelling ($p < .05$), as shown in Table 3.

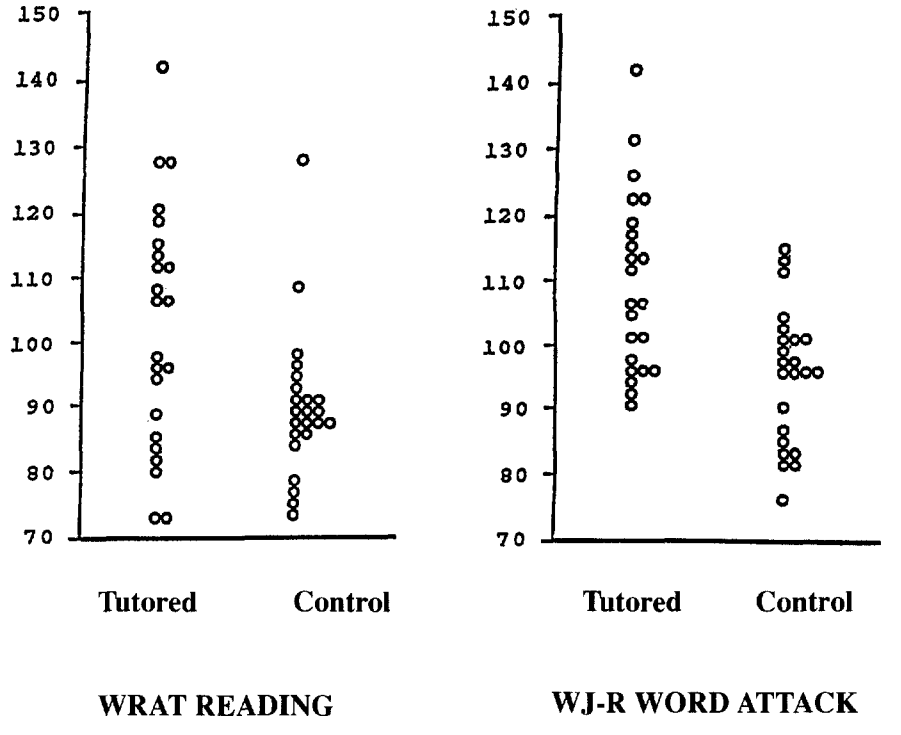

FIGURE 1. Individual performance on two word-level measures of reading.

Discussion

Lesson Revisions and Program Implementation

We undertook the present experiment, in part, to study treatment effects after addressing problems observed in instructional content and implementation of tutoring for students at risk for learning disability. Regarding revisions in lesson content, the children's performance (mean of 94%) on the periodic mastery tests given every 10 lessons suggests that they acquired the skills targeted by the program. Regarding fidelity of implementation, we found that providing more training in lesson components *before* tutors began working with children, along with increased supervision, resulted in more accurate implementation, relative to levels observed in prior field tests. Whereas in the previous field test only 30% of tutors were observed to implement the majority of the lesson activities consistent with program protocols

TABLE 3
Adjusted Means and Standard Deviations of Treatment and Control Groups at Grade 2 Follow-Up

Measures	Treatment		Control				
	M	*SD*	*M*	*SD*	*F*	*p*	*ES*
Reading							
Woodcock-Johnson Word Identification (standard)	97.05	15.63	94.49	11.57	.31	.580	.19
Analytical Reading Inventory, Second-grade level	57.42	38.87	61.10	40.39	.09	.764	−.09
Decoding							
Woodcock-Johnson Word Attack (standard)	102.32	14.93	91.64	9.20	6.46	.016	.87
Spelling							
TWS–Predictable (standard)	91.15	11.12	82.35	8.51	6.88	.013	.89
TWS–Unpredictable (standard)	88.55	11.52	81.36	7.25	4.88	.034	.75

Note. Treatment group $n = 20$; control group $n = 17$. TWS = Test of Written Spelling.

(Vadasy et al., 1997b), in this field test 71% of tutors were observed to be high implementors. Moreover, anecdotal evidence (e.g., tutors who increasingly followed program elements and implemented them with greater skill) suggests that the frequent supervision and technical assistance contributed to improved implementation. Obtaining more accurate program implementation was important because a previous finding had indicated a relation between fidelity of implementation and reading outcomes (Vadasy et al., 1997b). Although these two modifications (revised lesson content and increased supervision) are confounded with respect to their effects on the children's reading and spelling achievement, the results of the periodic performance tests and the tutors' higher fidelity of implementation suggest that, together, the changes were successful.

Treatment Effects

Comparisons of pretests revealed no hint of differences between the treatment and control groups on 9 of 10 measures (all Fs < 1.0), but on 1 measure (WRAT-Spelling), the treatment group seemed to have an advantage ($p < .05$). This advantage translated into a 2-point difference between groups on raw WRAT-R Spelling scores. The treatment mean was 17 and the control mean was 15. A student who copied all of the marks correctly and was able to print two letters in his or her name earned 20 points on the Spelling subtest. This spelling pretest advantage for the treatment group may qualify the results, but we have attempted to adjust for any pretreatment differences with an analysis of covariance.

At the end of first grade, tutored children significantly surpassed controls on a broad range of reading and spelling measures. Relative to national norms, mean posttest performance of the treatment group surpassed the 50th percentile on WRAT-R Reading and WJ-R Word Attack. It is instructive to compare our results with those of other intensive first-grade reading interventions that have reported positive effect sizes. Table 4 shows results for other first-grade studies that have tested tutoring against a comparable untutored control group. Requiring a control group excludes several recent tutoring reports (Invernizzi, Juel, & Rosemary, 1997; Juel, 1996; Vellutino et al., 1996). Most of the interventions in Table 4 were delivered by certified teachers. On the dimension of treatment intensity (minutes per week and number of weeks), our treatment is near the median, relative to previous studies. On the dimension of effectiveness, our effect sizes are just above the median on each measure (word, nonword, and composite reading).

Two tutoring studies included measures that we also employed (Iversen & Tunmer, 1993; Juel, 1996). Posttest results from those studies and ours are shown in Table 5. Dolch Word Recognition results from Iversen and Tunmer's Standard Reading Recovery group were virtually identical to ours, but their Modified Reading Recovery group, which included explicit training in phonological recoding, earned higher scores. We also rescored our participants' WRAT-R Reading subtest using the scoring system reported by Juel (1996). Her word recognition posttests from the WRAT-R were similar to ours, but our spelling scores seemed to be somewhat higher. Pretest word identification levels were compa-

TABLE 4
Intensity and Effects of One-to-One Intervention in First-Grade Tutoring Studies

Study	Tutors	Sessions/ week	Minutes/ week	Number of weeks	Effect size: Real word	Effect size: Nonword	Effect size: Composite reading
Ellson, Harris, & Barber (1968)	Paraprofessionals	5	75	35	0.10[a]	NA	0.01
					0.26[b]	NA	0.36
Hatcher, Hulme, & Ellis (1994)	Teachers	2	60	20	0.30	0.30	0.35
Iversen & Tumner (1993)	Teachers	4	120	12–20	3.41[c]	1.32	2.39
					3.40[d]	1.25	2.68
Pinnell, Lyons, DeFord, Byrk, & Seltzer (1994)	Teachers	5	150	20	NA	NA	0.73
Silver, Hagin, & Beecher (1981)	Teachers	3–5	120	35	0.94	1.39	1.06
Slavin et al. (1990)	Teachers	5	100	35	0.58	1.39	0.48
Vadasy et al. (1998)	Paraprofessionals	4	120	27	0.89	1.16	0.85
Wallach & Wallach (1976)	Paraprofessionals	5	150	35	0.64	NA	NA

Note. NA indicates the information was not available.
[a]Programmed tutoring group (one session daily). [b]Directed tutoring group (one session daily). [c]Standard Reading Recovery. [d]Modified Reading Recovery.

rable in Juel's and our study (2.85 vs. 1.87, respectively), as were spelling pretest levels (1.33 vs. .78, respectively). Although there is risk in comparing performance from samples that may differ on unknown characteristics, it is nevertheless interesting to note the comparability in posttest performance across first-grade tutoring studies that target children at risk in reading.

Follow-Up Testing and the Character of Treatment Effects

Tests at the end of the treatment period showed the strongest reading effects on phonological reading skill (i.e., naming nonwords), followed by word recognition, then reading fluency. For an approach like ours, which emphasized phonological decoding skills, these results were consistent with expectations. But as Torgesen, Wagner, Rashotte, et al. (1997) indicated, focusing instruction on phonological reading skill derives from the idea that these skills facilitate development of accurate and fluent reading, which in turn frees attention for comprehension. Each of these stages (phonological reading, accurate word recognition, fluent reading in context) is considered necessary, if not sufficient, for the development of subsequent stages. Like other recent studies (Lovett et al., 1994; Olson, Wise, Ring, & Johnson, 1997; Torgesen, Wagner, Rashotte, et al., 1997), our second-grade follow-up results indicated that phonological reading skills did not automatically translate into superior word recognition and fluent reading. One year after treatment, the tutored group still enjoyed a large advantage in phonological reading skills and spelling of regular words, but it performed similarly to the control group in both word recognition and fluency. Not only did the treated group fail to use its superior phonological reading skill to increase its advantage in word learning, but it also appears to have lost most of its original advantage. We must, however, acknowledge that our regrettable decision to change word recognition tests at second-grade follow-up testing (i.e., using the Woodcock-Johnson rather than the WRAT-R) resulted in a confounding of the word recognition measure with the treatment and control differences measured at two points in time. Nevertheless, these results, combined with those of other investigators, suggest that researchers need to revisit the assumed role of phonological reading skill in advancing related reading skills. Additional instruction may be required to encourage children to use these skills to improve word learning during independent learning trials, or to help them create more complete representations of words in memory during assisted learning trials. Or, if volunteer tutors can successfully help at-risk students develop phonological reading skill, then the expertise of reading specialists might be needed to help these children extend this advantage to word learning, fluency, and skilled comprehension.

Response to Treatment

There is widespread belief that if children at risk for learning disabilities could receive early intervention (e.g., daily explicit, intensive, one-to-one tutoring in phonological skills over the course of several months), many reading difficulties could be overcome, especially if intervention began in first grade, before the cascading effects of instructional failures are felt (Stanovich, 1986). As Figure 1 illustrates, many children in this study responded positively to tutoring, achieving reading skills that were at or above grade level. Indeed, a number of our students emerged as remarkably fearless decoders. Nevertheless, we also observed a small but significant group of children who were unable to master first-grade literacy skills, despite 27 weeks of one-to-one instruction. Five students (22%) scored in the lowest 25th quartile in reading, as did 5 students in spelling, suggesting that some children require more intensive or longer assistance than this program provided, or different assistance altogether.

Variable response to early intervention appears to be the norm. In their study of first graders with reading impairments who were provided one semester of daily 30-minute one-to-one tutoring, Vellutino et al. (1996) found that 33% still scored below the 30th

TABLE 5

Posttest Means (for Raw Scores) and Standard Deviations on Common Measures in First-Grade One-to-One Interventions

	WRAT-R Word Recognition		WRAT-R Spelling		Dolch		Yopp	
Iversen & Tumner (1993)								
Modified Reading Recovery					[a]153.88	(44.61)	[b]16.88	(4.53)
Standard Reading Recovery					[a]143.41	(40.41)	[b]17.63	(4.46)
Juel (1996)								
Overall	20.55	(9.35)	12.44	(5.64)				
Vadasy et al. (1998)	22.04	(10.05)	28.83	(3.83)	144.74	(54.95)	15.51	(3.79)

Note. Parentheses denote *SD.*

[a]End of year measures. [b]Discontinuation measures.

percentile, and 15% scored below the 15th percentile, on standardized tests. O'Connor (1997) observed that about 10% of students who received an intensive one-to-one intervention in first grade failed to benefit significantly, as measured on tests of rapid letter naming, segmenting, and standardized reading. Vandervelden and Siegel (1997) noted that 2 of the 10 lowest scoring kindergartners in their 12-week phonemic awareness intervention failed to improve their phoneme recognition skills. Even the strong positive effects found for Reading Recovery apparently exclude a subset of children who do not respond to treatment (Shanahan & Barr, 1995).

Juel (1996) suggested that interventions longer than 1 year may be needed by children attending schools with a large population of children from low-SES homes. Such children often enter school with fewer literacy experiences, along with more significant health and social welfare needs (Bowey, 1995). Multiyear treatments may succeed in returning some children to a typical developmental trajectory (Blachman, 1994), but even extended, state-of-the-art treatments seem to fall short for a small percentage of children (Torgesen, Wagner, & Rashotte, et al., 1997; Vellutino et al., 1996). Early tutoring can function as a screening mechanism to identify children who require more expert and intensive instruction (Vellutino et al., 1996). In fact, several Title I and special education teachers told us that they regarded a student's failure to improve in our program to be an indicator for special education assessment.

Infrastructure for Volunteer Nonprofessional Tutoring Programs

When nonprofessional tutors in this study received regular supervisory support, they demonstrated better teaching skills and more accurately implemented elements of this tutoring program, relative to the levels observed in previous field tests, when supervision was more intermittent. Relatedly, they were able to significantly affect first-grade literacy outcomes for children at risk for reading failure. This constitutes a good news/bad news situation: Tutors can provide the kind of critical instruction that spells the difference between success and failure for some children, but the tutors must receive considerable training, support, and supervision. With less intensive supports, as we reported previously (Vadasy et al., 1997b), tutors may be far less effective. Large individual differences exist among those who turn out for tutoring; some nonprofessional tutors require significant guidance to be effective. At the same time, we also observed a group of individuals with impressive talents for teaching students with low reading skills, including the ability to pace instruction briskly, adjust to students' needs for modeling and scaffolding, and manage behavior—observations consistent with Juel (1996). Nonprofessional tutors can develop strong teaching repertoires, especially if they stay with the program beyond 1 year and continue to receive support from reading teachers (Invernizzi, Rosemary, Juel, & Richards, 1997). Nonprofessionals who tutored for more than 1 year tended to be more successful, in part because they seemed to gain a deeper understanding of reading acquisition.

Besides providing information on treatment efficacy, field tests disclose challenges that schools face in bringing research-based practices to scale. To establish and maintain a systematic tutoring approach like ours requires considerable infrastructure that extends beyond the lesson materials. There must be individuals who can recruit reliable and conscientious tutors, provide training and supervision, integrate tutoring into the schools' schedules, give technical assistance on instructional and management problems, and help in assessing student progress. In several schools that are now implementing this treatment, Title I and special education teachers have assumed these tasks. At other schools, however, we have observed that even very enthusiastic teachers are unable to spare the time to provide sufficient supervision.

Another infrastructure element is funding to hire tutors. Schools participating in our research devised two strategies to continue the program after grant funds were no longer available. One group of schools reallocated Title I funds to pay parent tutors and instructional assistants to tutor for 1 or 2 hours each day. Another group of schools raised supplementary funds through grants from local foundations or obtained assistance from their Parent–Teacher–Student Associations to hire parent tutors from the community.

Our approach to early intervention using nonprofessional tutors includes most of the features that Wasik (1998) identified in her review of volunteer tutoring programs. Like all of the programs she reviewed, our program was not coordinated with classroom reading instruction. While we agree that such a match is desirable, the tremendous diversity across classroom teachers' literacy instruction (ranging from literature-based to basals to explicit phonics approaches) makes achieving this match difficult, unless reading teachers can adjust tutoring lessons according to each child's classroom circumstances and progress (e.g., Invernizzi, Juel, & Rosemary, 1997). Because we lacked sufficient resources to adjust each child's program in this manner, we opted to provide tutors with one set of structured materials and explicit teaching strategies, along with help in implementing them. Our results suggest that positive outcomes can be achieved using a standard tutoring approach across schools and classrooms.

Conclusions

This study adds to the findings that explicit training in phonological skills improves word-level reading and spelling skills. Results suggest that some Grade 1 effects are not sustained at the

end of second grade. Our change in word recognition measures in Year 2, however, makes it more difficult to interpret the Grade 2 findings. Together with findings from our previous research (Vadasy et al., 1997a, 1997b), our results indicate that programs using nonteacher tutors can produce broad and meaningful reading improvements for first-grade students at risk for reading disability, but only if the programs provide carefully designed, structured lessons, along with regular training and supervision for the tutors. Findings from this series of studies have implications for instructionally sound public policy regarding tutors.

ABOUT THE AUTHORS

Patricia F. Vadasy *is a senior research associate at the Washington Research Institute. Her specializations are early reading instruction and implementation of research-based practices.* ***Joseph R. Jenkins*** *is a professor in special education in the College of Education at the University of Washington, Seattle. His specializations are learning disabilities and instruction and assessment.* ***Kathleen Pool*** *is a first-grade teacher at Captain Johnston Blakely Elementary, Bainbridge Island, WA. Her specialization is early reading instruction. Address: Patricia F. Vadasy, Washington Research Institute, 150 Nickerson St., Ste. 305, Seattle, Washington 98109 (e-mail: pvadasy@wri-edu.org).*

AUTHORS' NOTE

This research was supported in part by Grant No. H023R20019 from the U.S. Department of Education to the Washington Research Institute, and by the Paul G. Allen Charitable Foundation.

REFERENCES

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

Berninger, V. W. (1990). Multiple orthographic codes: Key to alternative instructional methodologies for developing the orthographic–phonological connections underlying word recognition. *School Psychology Review, 19,* 518–533.

Berninger, V. W., Thalberg, S., DeBruyn, I., & Smith, R. (1987). Preventing reading disabilities by assessing and remediating phonemic skills. *School Psychology Review, 16,* 554–565.

Berninger, V. W., Vaughan, K., Abbott, R. D., Brooks, A., Abbott, S. P., Rogan, L., Reed, E., & Graham, S. (1998). Early intervention for spelling problems: Teaching spelling units of varying size with a multiple connections framework. *Journal of Educational Psychology, 90,* 1–19.

Blachman, B. A. (1994). What we have learned from longitudinal studies of phonological processing and reading, and some unanswered questions: A response to Torgesen, Wagner, and Rashotte. *Journal of Learning Disabilities, 27,* 287–291.

Bowey, J.A. (1995). Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. *Journal of Educational Psychology, 87,* 476–487.

Bradley, L., & Bryant, P. E. (1985). *Rhyme and reason in reading and spelling.* Ann Arbor: University of Michigan Press.

Bryant, N. D. (1975). *Diagnostic test of basic decoding skills.* New York: Columbia University, Teachers College.

Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *Journal of Speech and Hearing Research, 36,* 948–958.

Clay, M. (1985). *The early detection of reading difficulties* (3rd ed.). Tadworth, Surrey: Heinemann.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Dolch, E. W. (1939). *A manual for remedial reading.* Champaign, IL: Garrard Press.

Dunn, L. M., & Dunn, L. (1981). *Peabody picture vocabulary test–Revised.* Circle Pines, MN: American Guidance Service.

Ehri, L. C. (1980). The development of orthographic images. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 311–338). London, England: Academic Press.

Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly, 27,* 13–26.

Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20,* 163–179.

Ellson, D. G., Harris, P., & Barber, L. (1968). A fieldtest of programmed and directed tutoring. *Reading Research Quarterly, 3,* 307–367.

Fielding-Barnsley, R. (1997). Explicit instruction in decoding benefits children high in phonemic awareness and alphabet knowledge. *Scientific Studies of Reading, 1,* 85–98.

Foorman, B. R., Francis, D. J., Beeler, T., Winikates, D., & Fletcher, J. M. (1997). Early interventions for children with reading problems: Study designs and preliminary findings. *Learning Disabilities, 8,* 63–71.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hove, England: Erlbaum.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65,* 41–57.

Invernizzi, M., Juel, C., & Rosemary, C. A. (1997). A community volunteer tutorial that works. *The Reading Teacher, 50,* 304–311.

Invernizzi, M., Rosemary, C., Juel, C., & Richards, H. C. (1997). At-risk readers and community volunteers: A 3-year perspective. *Scientific Studies of Reading, 1,* 277–300.

Iversen, S., & Tunmer, W. E. (1993). Phonological processing skills and reading recovery program. *Journal of Educational Psychology, 85,* 112–126.

Jastak, S., & Wilkinson, G. S. (1984). *The wide range achievement test–Revised.* Wilmington, DE: Jastak Associates.

Jenkins, J. R., Vadasy, P. F., Firebaugh, M., & Profilet, K. (in press). Tutoring first-grade struggling readers in phonological reading skills. *Learning Disabilities Research and Practice.*

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.

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 30,* 268–289.

Juel, C., & Roper/Schneider, D. (1985). The influence of basal readers on first-grade reading. *Reading Research Quarterly, 20,* 134–152.

Larsen, S. C., & Hammill, D. D. (1994). *Test of written spelling* (3rd ed.). Austin, TX: PRO-ED.

Liberman, I. Y., Shankweiler, D., Blachman, B. A., Camp, L., & Werfelman, M. (1980). Steps towards literacy. In P. Levinson & C. Sloan (Eds.), *Auditory processing and language: Clinical and research perspectives* (pp. 189–215). New York: Grune & Stratton.

Lovett, M. W., Borden, S. L., Deluca, T., Lacerenza, L., Benson, N. J., & Brackstone, D. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically and strategy-based reading training programs. *Developmental Psychology, 30,* 805–822.

Madden, N. A., Slavin, R. E., Karweit, N. L., Dolan, L. J., & Wasik, B. A. (1993). Success for all: Longitudinal effects of a restructuring program for inner-city elementary schools. *American Educational Research Journal, 30,* 123–148.

O'Connor, R. E. (1997, February). *Layers of intervention.* Paper presented at the annual meeting of the Pacific Coast Research Conference, La Jolla, CA.

O'Connor, R. E., Jenkins, J. R., & Slocum, T. A. (1995). Transfer among phonological tasks in kindergarten: Essential instructional content. *Journal of Educational Psychology, 87,* 202–217.

Olson, R. K., Wise, B., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the post-training development of word recognition. *Scientific Studies of Reading, 1,* 235–254.

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Byrk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly, 29,* 9–39.

Rosner, J. (1979). *Helping children overcome learning disabilities* (2nd ed.). New York: Walker.

Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading and Writing Quarterly, 13,* 107–122.

Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk children. *Reading Research Quarterly, 30,* 958–996.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151–218.

Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education, 1,* 1–58.

Silver, A. A., Hagin, R. A., & Beecher, R. (1981). A program for secondary prevention of learning disabilities: Research in academic achievement and in emotional adjustment. *Journal of Preventive Psychiatry, 1,* 77–87.

Slavin, R. E., Madden, N. A., Karweit, N. L., Livermon, B. J., & Dolan, L. (1990). Success for All: First-year outcomes of a comprehensive plan for reforming urban education. *American Educational Research Journal, 27,* 255–278.

Spear-Swerling, L., & Sternberg, R. J. (1996). *Off track: When poor readers become 'learning disabled'.* Boulder, CO: Westview Press.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–406.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities, 27,* 276–286.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1997). Prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies in Reading, 1,* 217–234.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Alexander, A. W., & Conway, T. (1997). Preventive and remedial interventions for children with severe reading disabilities. *Learning Disabilities, 8,* 51–61.

Vadasy, P. F., Jenkins, J. R., Antil, L., Wayne, S., & O'Connor, R. E. (1997a). Community-based early reading intervention for at-risk first graders. *Learning Disabilities Research and Practice, 12,* 29–39.

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

Vadasy, P. F., Wayne, S. R., O'Connor, R. E., Jenkins, J. R., & Pool, K. (1998). *Sound Partners: One-to-one tutoring in early reading skills.* Manuscript submitted for publication.

Vandervelden, M. C., & Siegel, L. S. (1997). Teaching phonological processing skills in early literacy: A developmental approach. *Learning Disability Quarterly, 20,* 63–82.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88,* 601–638.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1994). In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 279–332). Baltimore: Brookes.

Venezky, R. (1970). *The structure of English orthography.* The Hague: Mouton.

Wallach, M. A., & Wallach, L. (1976). *Teaching all children to read.* Chicago: University of Chicago Press.

Wasik, B. A. (1998). Volunteer tutoring programs in reading: A review. *Reading Research Quarterly, 33,* 266–292.

Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28,* 179–200.

Williams, J. P. (1985). The case for explicit decoding instruction. In J. Osborn, P. Wilson, & R. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 205–213). Lexington, MA: D.C. Health.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson psycho-educational battery–Revised.* Allen, TX: DLM Teaching Resources.

Woods, M. L., & Moe, A. J. (1977). *Analytical reading inventory.* Sydney, Australia: Charles E. Merrill.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23,* 159–177.