Reading and Writing: An Interdisciplinary Journal **11:** 239–273, 1999.

Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children

BENITA A. BLACHMAN[1], DARLENE M. TANGEL[2], EILEEN WYNNE BALL[3], ROCHELLA BLACK[4] & COLLEEN K. MCGRAW[5]
[1]*Syracuse University;* [2]*Oriskany School District and Syracuse University;* [3]*University of Illinois at Chicago;* [4]*Northport – East Northport School District;* [5]*University of Pittsburgh at Bradford, USA*

Abstract. Low-income, inner-city children were involved in a two-year intervention delivered in the regular classroom by regular classroom teachers to develop phonological awareness and word recognition skills. For the treatment children, an 11-week phoneme awareness program in kindergarten was followed by a first grade reading program (extended to grade 2 for some children) that emphasized explicit, systematic instruction in the alphabetic code. Control children participated in the school district's regular basal reading program. Both groups participated in a phonetically-based spelling program mandated by the district. At the end of grade 1, treatment children (n = 66) significantly outperformed control children (n = 62) on measures of phonological awareness, letter name and letter sound knowledge, and three measures of word recognition, and reached marginal significance (0.056) on a fourth. They also significantly outperformed the control children on two measures of spelling. One year later, at the end of grade 2, the treatment children (n = 58) significantly outperformed the control children (n = 48) on all four measures of word recognition. For the groups as a whole, there were no differences on the one measure of spelling readministered at the end of grade 2. However, there were significant differences in spelling between the treatment (n = 16) and control children (n = 13) who remained in the bottom quartile of spellers at the end of grade 2 when partial credit was given for phonetically correct spelling, and significant differences in reading favoring these treatment children on all four measures of word recognition.

Keywords: Early reading, Literacy, Phonological Awareness, Phoneme awareness, Reading intervention, Word recognition

Introduction

During the last 25 years much has been learned about how young children acquire literacy skills, especially about the role of phonological processes in learning to read an alphabetic writing system and about how to facilitate the development of early reading and spelling skills (e.g., Adams 1990; Beck & Juel 1995; Blachman 1997; Brady & Shankweiler 1991; Goswami & Bryant 1990; Gough, Ehri & Treiman 1992; Rieben & Perfetti 1991). Researchers have concluded that during the early stages of reading acquisition 'instruction that facilitates both phoneme awareness and alphabetic coding is vitally

important to success' (Vellutino 1991: 442). Although no one is suggesting that these elements of instruction are sufficient, there is now consensus among researchers that these elements are indeed necessary (e.g., Adams & Bruck 1993; Liberman & Liberman 1990; Stanovich 1986, 1994; Torgesen, Wagner & Rashotte 1997a).

Data from numerous intervention studies in preschool, kindergarten, and first grade classrooms (e.g., Ball & Blachman 1991; Blachman, Ball, Black & Tangel 1994; Bradley & Bryant 1983; Byrne & Fielding-Barnsley 1991, 1993, 1995; Castle, Riach & Nicholson 1994; Lie 1991; Lundberg, Frost & Petersen 1988; O'Connor, Notari-Syverson & Vadasy 1996) indicate that early intervention to heighten phonological awareness – an awareness of the phonological segments in spoken words – has a facilitating effect on early reading and spelling acquisition. These results provide support for the insight of Isabelle Liberman, made more than 25 years ago, that one of the fundamental tasks facing the beginning reader is understanding that speech can be segmented and that these segmented units can be represented by printed forms (Liberman 1971, 1973). Becoming aware of the phonemic segments in speech, however, is complicated by the fact that the phonemes are 'merged' or coarticulated during speech production (Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967), with the result that we hear only a single acoustic unit – the syllable. Consequently, although an awareness of the phonemic segments in spoken words will make it easier for a child to understand how an alphabetic orthography represents speech (e.g., why the written representation of /sun/ has three letters), this awareness does not come easily for many beginning readers. It has been found repeatedly that children who lack this linguistic insight are likely to be among our poorest readers (Blachman 1984; Bradley & Bryant 1983; Byrne & Fielding-Barnsley 1993; Fletcher et al. 1994; Høien, Lundberg, Stanovich & Bjaalid 1995; Iversen & Tunmer 1993; Juel 1988; Lundberg, Olofsson & Wall 1980; Mann & Liberman 1984; Share, Jorm, Maclean & Matthews 1984; Stanovich, Cunningham & Cramer 1984; Torgesen & Burgess 1998; Vellutino & Scanlon 1987; Vellutino et al. 1996).

In addition to the knowledge we have gained in the area of phonological awareness, we also know that accuracy and fluency in decoding are critical to reading success: without accuracy and fluency at the level of the individual word, there will always be constraints on comprehension (Adams 1990; Adams & Bruck 1995; Beck & Juel 1995; Chall 1989; Ehri 1991, 1997; Rieben & Perfetti 1991; Vellutino 1991; Williams 1979, 1985, 1994). Researchers have pointed out, however, that 'sometimes children have trouble learning to decode because they are completely unaware of the fact that spoken language is segmented – into sentences, into syllables, and into phonemes' (Williams 1987: 25–26). Support for this idea comes from Juel's

(1988, 1994) longitudinal study of the reading development of fifty-four children from first to fourth grade. Juel found that the fourth grade poor readers entered first grade with limited phonological awareness. This lack of understanding of the internal structure of words contributed to their slowness in learning letter sound correspondences and decoding. Indeed, at the end of grade 4, the decoding of the poor readers was not yet equivalent to that of the average and good readers at the beginning of grade 2. Gough and Juel conclude that 'what seems essential is to insure that children learn to decode in first grade. If decoding skill does not arrive then, it may be very hard to change the direction that reading achievement takes' (1991: 55).

Developing phonological awareness and accurate and fluent decoding skills/word recognition early in the child's progress of literacy acquisition seem to be essential. When we began this project, most early intervention studies, however, had investigated *either* the effect of phonological awareness (e.g., Ball & Blachman 1991) *or* the benefits of a decoding approach to reading (e.g., Felton 1993). These studies had not investigated the effects of an intervention that (1) included phonological awareness instruction in kindergarten followed by a first grade reading program that built on this awareness and emphasized the alphabetic code, and (2) was provided in the regular classroom by regular classroom teachers.

Several studies with somewhat older severely learning disabled children (ranging in age from 7 to 13) have found that a treatment that emphasizes *both* phonological awareness and word recognition/decoding facilitates reading acquisition (Alexander, Anderson, Heilman, Voeller & Torgesen 1991; Lovett et al. 1994; Williams 1980). With poor readers in grades 1 and 2 and at-risk children selected in kindergarten and treated through grade 2, there is also evidence of a positive effect on reading when intervention combines phonological awareness instruction and direct instruction in reading (Hatcher, Hulme & Ellis 1994; Iversen & Tunmer 1993; Torgesen, Wagner & Rashotte 1997b; Vellutino et al. 1996).

With even younger children – those in preschool and kindergarten classrooms – the phonological awareness studies have often included instruction in connecting the phonological segments to letters (e.g., Ball & Blachman 1991; Byrne & Fielding-Barnsley 1991), but have not included more formal instruction in decoding, both because it would be developmentally inappropriate and also because the goal of these studies has been to isolate the effect on early reading of phonological awareness instruction provided *before* children were engaged in formal reading instruction. Our clinical experience, however, and feedback from the teachers of children who had participated in our earlier kindergarten study that did not include a first grade treatment (Ball & Blachman 1991), suggested that teachers wanted guidance in creating

a link between the kindergarten activities designed to enhance the phonological awareness of their children and the development of decoding skills in the early grades. We set out to create and evaluate a program that would incorporate what is known about developing both phonological awareness and good decoding skills.

Another factor motivating this research was the fact that most studies, including some of our own previous work (Ball & Blachman 1991), have utilized specially trained teachers who are brought to the schools to implement whatever phonological awareness intervention or special reading program is being evaluated. Although such studies have demonstrated the value to early reading acquisition of an emphasis on phonological awareness and alphabetic coding, studies conducted by specially trained teachers brought to schools by the experimenter may not be as convincing to the practitioners being asked to implement these models of instruction as studies in which the experimental treatment is provided during the regular school day in the general education classroom by regular classroom teachers. Thus, our goal in the present study was to develop and evaluate a two-stage intervention delivered in mainstream classrooms by kindergarten, first and second grade teachers. The intervention included an 11-week phonological awareness program delivered during the second half of the kindergarten year, followed by a reading program in grade 1 (and continuing to grade 2 for some children) that extended the earlier phonological awareness instruction and emphasized the alphabetic code. The results of the kindergarten phonological awareness intervention were presented in Blachman et al. (1994) and will be summarized below. The remainder of this paper will focus on the results of the first grade year and follow-up results from the end of grade 2.

Kindergarten intervention

As previously reported (Blachman, Ball, Black & Tangel 1994), children in this study were selected during the kindergarten year from the total population of 18, all-day kindergartens in four, demographically comparable low-income, inner-city schools, in a large, urban district in upstate New York. Of the 21 elementary schools in the district, we chose 4 of the 5 schools in the district with the lowest achievement scores in reading on standardized tests. Treatment and control children attended different schools to ensure that the control children would not be exposed to the treatment activities (for a more detailed description of subject selection criteria, see Blachman et al. 1994). During the kindergarten year, 159 children participated in the study. The 84 treatment children included 47 boys and 37 girls with a mean age of 5.62; the 75 control children included 38 boys and 37 girls with a mean age of 5.64.

Prior to the intervention, there were no significant differences between the treatment and control children on age [$t(157) = 0.46$; $p = 0.64$], sex [$\chi^2(1, n = 159) = 0.44$; $p = 0.50$], race [$\chi^2(2, n = 159) = 0.56$; $p = 0.76$], SES, measured by number of children in each group receiving free lunch [$\chi^2(1, n = 159) = 0.28$; $p = 0.60$], developmental level, based on a screening instrument designed by the school district and administered to all kindergarten children [$t(157) = 1.50$; $p = 0.13$], sound counting [$t(157) = 0.56$; $p = 0.58$], or any of the other pretest variables (i.e., PPVT-R, phoneme segmentation, letter name and letter sound knowledge or word identification), as shown in Table 1. These children, as a group, were in the low average range in receptive vocabulary as measured by a mean PPVT-R score of 91 (mean = 100; sd = 15), and approximately 85% of our sample received free or reduced lunch. The children had extremely limited knowledge of the alphabet as indicated by the fact that the children knew, on average, only 2 letter sounds in January of kindergarten prior to the beginning of the intervention, despite the fact that the district's kindergarten program included letter name and letter sound instruction. As pointed out to us by the classroom teachers, most of the children could not yet write their names in January of their kindergarten year. (As a point of reference, these children were lower skilled than the children in our earlier study, Ball and Blachman (1991): the mean PPVT-R of the children was 101, and the children, who were from a mix of high-, middle-, and low-income families, knew an average of 9 letter sounds when the intervention began.)

From February to May of kindergarten, treatment children participated in 41, 15 to 20 minute phonological awareness lessons (adapted from the shorter, 28 lesson program used in Ball and Blachman, 1991), delivered over an 11 week period to small, heterogeneous groups of four or five children. Classroom teachers and their teaching assistants were responsible for teaching all of the lessons. Each lesson followed the same daily format: (1) a phoneme segmentation activity (called say-it-and-move-it) in which children learned to move disks to represent the sounds in one-, two-, and three-phoneme words spoken by the teacher, (2) a segmentation-related activity, such as grouping words on the basis of shared sounds (e.g., hat and hot go together because they share the same initial sound), and (3) one of a variety of activities to teach the letter names and sounds of eight letters (a, m, t, i, s, r, f, b) (for a more detailed description of these lessons, see Blachman et al. 1994).

After completing these 41 lessons, the treatment children were superior to the control children in terms of phonological awareness, letter name and letter sound knowledge, reading phonetically regular real words and pseudowords, and developmental spelling (see Table 1).

Table 1. Pretest and posttest means for treatment and control groups

	Treatment			Control			
Variable	Mean	sd	n	Mean	sd	n	*p*
Pretests							
January of kindergarten							
PPVT-R	91.4	11.3	84	90.7	9.6	75	0.655
Phoneme segmentation	11.9	4.2	84	11.8	4.6	75	0.897
Letter names	11.2	6.9	84	10.7	7.4	75	0.653
Letter sounds	2.4	3.9	84	2.8	4.0	75	0.493
Woodcock word ID	0.1	0.4	84	0.1	0.4	75	0.432
Posttests							
End of kindergarten							
Phoneme segmentation	23.6	6.9	84	13.2	4.4	75	0.0001
Letter names	19.0	6.2	84	17.1	7.1	75	0.0201
Letter sounds	13.3	5.8	84	9.4	6.8	75	0.0001
Woodcock Word ID	0.6	1.6	84	1.0	3.0	75	0.0807
Regular Words I	4.2	5.3	84	0.4	1.7	75	0.0001
Nonwords I	2.3	3.2	84	0.2	0.8	75	0.0001
Developmental spelling	11.6	6.8	77	6.0	5.0	72	0.0001
End of first grade							
Phoneme segmentation	27.9	3.2	66	23.1	6.5	62	0.0001
Letter names	25.5	1.0	66	24.2	3.9	62	0.0190
Letter sounds	24.6	1.6	66	21.3	5.0	62	0.0001
Woodcock Word ID	33.9	14.1	66	28.6	15.8	62	0.0560
Decoding Skills Test							
Real words	26.7	13.7	66	16.2	15.1	62	0.0001
Nonwords	15.2	6.8	66	8.7	8.2	62	0.0000
Regular Words II	36.9	11.0	66	25.4	14.7	62	0.0001
Developmental spelling	47.1	5.8	66	41.4	10.8	62	0.0004
WRAT-R spelling	10.2	3.7	66	8.6	4.7	62	0.0264
Modified WRAT-R	47.0	15.1	66	36.4	19.8	62	0.0010
End of second grade							
Woodcock Word ID	54.9	10.4	58	49.6	14.6	48	0.0287
Decoding Skills Test							
Real words	45.0	12.9	58	38.1	17.8	48	0.0278
Nonwords	36.4	14.2	58	29.0	18.1	48	0.0196
Regular Words II	47.6	6.9	58	41.2	12.6	48	0.0024
WRAT-R spelling	15.6	4.2	58	15.6	5.5	48	0.9378
Modified WRAT-R	65.6	14.8	58	61.8	20.5	48	0.2967

Table 2. Pretest means for treatment and control groups for students remaining at the end of first grade

	Treatment (n = 66)		Control (n = 62)			
Variable	Mean	sd	Mean	sd	*t*	*p*
Pretests						
PPVT-R	91.5	11.6	90.3	9.7	0.616	0.539
Phoneme segmentation	12.2	4.1	12.0	4.6	0.277	0.782
Letter names	11.5	6.6	10.3	7.3	0.962	0.333
Letter sounds	2.3	3.6	2.8	4.1	0.780	0.437
Woodcock Word ID	0.1	0.5	0.1	0.4	0.538	0.591

Grade one intervention

Method

Participants. Between June of kindergarten and May of first grade, 31 children were eliminated from the study. The treatment group lost 18 children; 15 moved, 1 was dropped due to behavior problems and 2 were dropped because of excessive absence due to illness. The control group lost 13 children; 11 moved and 2 children were placed in self-contained special education classes. This left a sample of 128 children (66 treatment and 62 control). Prior to analyzing the first grade posttest data, pretest data from kindergarten of the 128 remaining children were reanalyzed. No significant differences were found on age [treatment mean = 5.61; control mean = 5.63, $t(126) = 0.35$, $p = 0.72$], sex [Yates $\chi^2(1, n = 128) = 0.245$, $p = 0.62$], race [Yates $\chi^2(2, n = 128) = 2.15$, $p = 0.34$], or any of the other pretraining variables (see Table 2).

Procedure

Grouping for instruction. At the end of kindergarten, all treatment children within a school were rank ordered based on their letter name and letter sound knowledge, phoneme segmentation scores, and word reading, and then separated into groups based on these scores. Each group (ranging in size from 6 to 9 children)[1] was assigned to a different first grade classroom in one of the two treatment schools. Each of the eleven first grade teachers in these two schools became responsible for one of these groups of children.

We separated the children into homogeneous groups for first grade reading instruction (consistent with the model used in this district) because,

as expected, all children did not move through the kindergarten program at the same pace. Many children caught on quickly in kindergarten and, during the say-it-and-move-it activity, went from using blank tiles to represent the phonemes in words to using tiles with printed letters on them. Other children at the end of kindergarten were not yet successful segmenting three-phoneme words using blank tiles. There was also wide variation in the number of letter sounds children knew at the end of kindergarten. It was clear that different children would need different amounts of time in first grade to review and to extend the phoneme awareness activities that we introduced in kindergarten. Separating the children into ability groups made it easier for first grade teachers to provide appropriately paced instruction – instruction tailored to meet the needs of a particular group and, within each group, instruction tailored to meet the needs of individual children.

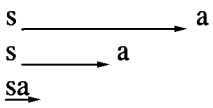
First grade reading program. It is important to note that the first grade reading program for the treatment children (described below) was provided *in place of* and *not in addition to* the basal reading program used with other children in this district. The length of each lesson (30 minutes) was consistent with the length of time both treatment and control classroom teachers reported spending with each of their reading groups. Children receiving the treatment approach, when compared to the control group, did not receive any extra time devoted to reading instruction in first grade.

Treatment children began the first grade year with a review of the phoneme awareness and letter sound activities presented in kindergarten. To help teachers accommodate individual differences, guidelines for the first grade reading program were prepared at three different levels: one level for groups needing minimal review of the kindergarten activities (1 to 2 weeks at the beginning of the school year), one level for groups needing a longer review period (4 to 6 weeks), and one level for children needing a more extensive review and gradual transition to the first grade reading program (about 12 weeks). In addition to practicing phoneme segmentation activities and games used during kindergarten, part of the first grade review and transition phase included introducing all letter sounds (not just the 8 sounds introduced during the kindergarten program), using selected workbook pages for the first time for additional practice connecting letters to sounds, and learning some high frequency words that would appear in the early readers (e.g., *I, to, said*). Following the review and transition phase, the first grade reading program for the treatment children consisted of a daily, 30-minute, 5-step reading program (adapted from Blachman 1987) that continued to reinforce phoneme awareness skills and emphasized the alphabetic code.

Although groups began the first grade 5-step reading program at different times in first grade (depending on the number of weeks spent in phonological awareness and sound-symbol review activities), once a group began the 5-step program, the group followed the steps described below.

1) Each lesson began with a brief and quick-paced (1 to 2 minutes) review of sound-symbol associations learned in previous lessons and the introduction of new sound-symbol correspondences. For this part of the lesson, teachers utilized a sound pack (set of cards) containing each of the graphemes (i.e., letters and letter clusters) being reviewed. In order to highlight the vowels, vowel letters were printed in red.

2) The second step in the program was instruction in phoneme analysis and blending skills. To avoid the pitfalls of the letter-by-letter blending strategy that teachers often use, a blending technique adapted from Engelmann (1969) was utilized. In the typical approach used to teach blending, children are taught to attack an unknown word by sounding it out letter-by-letter (e.g., buh-a-tuh) and then blending it to produce the word (e.g., bat). It is impossible with this approach, however, to recover the original word 'bat', regardless of how quickly the child tries to blend the sounds together (A. Liberman et al. 1967; Liberman 1971; Liberman & Shankweiler 1979). The Engelmann procedure avoids much of the distortion that comes with trying to produce sounds (e.g., stop consonants) in isolation. Children were taught to pronounce as a single unit a consonant (continuant) followed by a vowel. To begin, the teacher represented this strategy on the board as follows:



The teacher pointed to the first letter, and the child was taught to produce that letter's sound and hold the sound until the teacher's finger reached the second letter. When her finger touched the second letter, the second sound was produced and held. With each successive practice opportunity, the length of time between sounds was decreased until the two sounds were pronounced as a single unit. By adding final consonants (initially, stop consonants) and pronouncing the whole word, a set of real words was built (e.g., sat, sam). Words containing new short vowels were also introduced in this manner. Depending on the needs of the group, this activity was used for a few days or a few weeks and then eliminated from the lessons.

A second phoneme analysis and blending technique was used throughout the first grade year to help children learn to synthesize sounds without resorting to letter-by-letter blending. Adapted from a technique suggested by Slingerland (1971), each child used a small pocket-chart (that we called a

'sound board') to manipulate letters to form words. Consonants and vowels previously mastered by the children were written on individual letter cards and placed in the top pocket. First, the teacher pronounced a word, such as fat, emphasizing the medial (vowel) sound. Then children repeated the word, listened for the vowel sound, and selected the appropriate vowel grapheme card (vowels were color coded) from the top pocket and placed it in the lower pocket. The teacher then repeated the word and asked the child to select the letter that represented the first sound in the word and place it in the appropriate position (i.e., in front of the vowel) in the bottom pocket. The teacher pronounced part of the word saying, 'Now we have fa. Our word is fat. What is the last sound we fear in fat?' The child then selected the t and placed it at the end of the word. The whole word was then read either by an individual child or by the group.

Once the child was successful representing words in this fashion, phoneme manipulation was introduced. For this task, the child might be asked to change fat to fan and, when new vowels were mastered, change fan to fin. A later lesson might require changing fin to shin and, eventually, as new syllable types were introduced on the sound board, changing shin to shine.

3) The third activity in this 5-step plan gave children the opportunity to develop more automatic recognition of words that they had practiced previously on the sound board. Once they could construct and accurately read on the sound board a pool of phonetically regular words, these words were put on flash cards and the children practiced reading them quickly. High frequency words that have to be memorized, such as said, were selected from stories the children would be reading and were also introduced at this time. These words were written in a different color. For approximately 2–3 minutes daily, children practiced reading both phonetically regular words and irregular, high frequency words. The goal of this quick-paced activity was to build automaticity.

4) Next, children engaged in 10 to 15 minutes of reading connected text. Children read phonetically controlled readers from the *Primary Phonics* series (published by Educator's Publishing Service) and selected stories from the Scott Foresman basal reading series used throughout the school district. (None of the other materials, such as workbooks, from the Scott Foresman series was used). Each classroom also had trade books for independent reading at other times during the day, and children went to the school library for additional reading materials. (It should be noted that although the materials used during the 30-minute reading lesson were consistent across treatment groups, treatment teachers were free to use whatever books they thought appropriate for the rest of the day. There was considerable variation across classrooms.)

5) The last step of each lesson included a short writing to dictation activity. Generally, 4 to 6 words and a sentence were dictated. Teachers dictated words drawn from word lists that were practiced on the sound boards or words encountered in the phonetically controlled readers. Children were directed to print vowel headings at the top of each dictation page (e.g., a and i, or later, ai, oa, ea). These headings represented the particular vowel sounds that were target sounds for that day's lesson. The dictation activity gave teachers an opportunity to evaluate student progress on the target sounds for the day. The dictation notebooks became a record of student growth over the first grade year, as both students and their teachers could review the progress that was made as students progressed from writing and reading simple closed syllable words (e.g., ham) to more complex patterns (e.g., hike, rain).

By the time children completed the program, they had been introduced to words representing all six syllable types, including closed syllables, such as fat and flat, final 'e' syllables, such as cake and shine, open syllables, such as me and cry, vowel team syllables, such as pain, teach, and crawl, vowel + r syllables, such as burn and start, and consonant le syllables, as in bottle and table. Although the focus of many of the early activities was on developing accurate and automatic word recognition skills, vocabulary development and comprehension were not neglected. Teachers were encouraged to make sure that children knew the meaning of all words that they were asked to read or spell and comprehension of stories was developed using a variety of strategies (e.g., retellings, making predictions). As the children progressed through the program and were able to recognize more words, more time in each lesson was devoted to reading new stories and rereading old ones. To continue to stay within the time allocated for reading groups, it was suggested to teachers that they begin to alternate the use of the sound board and dictation, using the sound board two days per week and dictation three days per week. This allowed for more time to be spent reading connected text.

The control children followed the traditional Scott Forseman basal reading program used in the district. During the approximately 30-minute reading lessons, teachers used the standard basal reader and workbooks, and word cards were used to introduce new vocabulary for each lesson. At another time during the day, the basal reader was supplemented by a single phonics workbook that the children used independently. Control teachers reported that 17 of the 62 control children also used a structured phonetic approach, *Reading Mastery*, for part of their reading instruction. All control children also engaged in reading trade books from their classroom libraries and their school libraries.

Both treatment and control children participated in whole class, phonetically-based spelling instruction using *Spelling: Words and Skills* by

Scott Foresman (1986) as mandated by the school district. Specifically, this first grade spelling program introduces lists of phonetically regular words containing short vowels (e.g., hen) and long vowels (e.g., name), as well as initial blends (e.g., bl) and digraphs (e.g., ch).

Teacher training. Prior to and during the first grade intervention (from August to June), treatment teachers and their teaching assistants participated in 13, 2-hour inservice workshops (two in August and September and then one every month through June). During these sessions, teachers learned how young children acquire literacy skills and about the role of phonological processes in learning to read. Teachers learned how to provide explicit instruction in the alphabetic code using the 5-step plan, developed sample lessons, and discussed matters such as pacing. We stressed that adequate time to develop phonological awareness and letter sound recognition and to develop accuracy and fluency in word recognition as the first grade year progressed was more important than 'covering the material' (for an extensive discussion of this point, see Juel 1994). Putting a high premium on appropriate pacing meant that some children would complete this intervention program, while others would not get through as many lessons or as many readers. The expectation, however, was that all children would experience high degrees of accuracy (and consequently success) during reading instruction. In addition to the inservice program, the project coordinator and a doctoral student in education visited the treatment schools weekly to answer questions, monitor treatment fidelity, assess the progress of the various groups, and monitor the progress of individual children.

Posttesting. At the end of May of the first grade year children were assessed by examiners blind to whether children were in the treatment or control group. Children were retested on phoneme segmentation, letter name and letter sound knowledge, and the Woodcock Reading Mastery Tests-Revised, Word Identification subtest (Woodcock 1987). Posttest only measures at the end of first grade included an experimenter-devised phonetically regular word list (Regular Words II), the Phonic Patterns subtest from The Decoding Skills Test (Richardson & DiBenedetto 1985), the 10-word, experimenter-devised Developmental Spelling Test (DST) (Tangel & Blachman 1995), and the spelling subtest from the Wide Range Achievement Test-Revised (WRAT-R) (Jastak & Wilkinson 1984).

Measures

Phoneme segmentation test. The phoneme segmentation test (Ball & Blachman 1988), originally adapted from a segmentation test designed by Liber-

man and her colleagues (1974), consists of 34 randomly arranged one-, two-, and three-phoneme items. It was preceded by four practice training trails. The training sequence provided modeling and corrective feedback in segmenting one-, two-, and three-phoneme words. The method and order of items during the actual test presentation were identical for all subjects. Internal reliability for this measure has been reported to be 0.91 (Ball & Blachman 1988).

Alphabet letter names and sounds. Each of the 26 letters of the alphabet was written in lower-case on a 3×5 index card. The cards were mixed and presented in fixed random order. The child was required to name the letter and give the sound each letter makes. Interrater reliability for this measure was established by Ball ($r = 0.997$) (Ball & Blachman 1991).

Woodcock Reading Mastery Tests-Revised, Word Identification subtest. The Word Identification subtest of the Woodcock Reading Mastery Tests measures word recognition by asking children to read individual words from a graded word list. The internal reliability of the Word Identification subtest for first grade is reported to be 0.98 (Woodcock 1987).

Phonetically Regular Word List (Regular Words II). On this experimenter-devised measure, each child was asked to read 54 phonetically regular words, which were grouped by phonetically regular patterns (e.g., closed syllable, final 'e' syllable, vowel team syllable, and vowel + r syllable).[2]

The Decoding Skills Test, The Phonic Patterns subtest (Richardson & DiBenedetto 1985). The phonic patterns subtest on the Decoding Skills Test assesses a child's ability to use phonic patterns to decode words. Six 5-word lists of monosyllabic real words, six 5-word lists of monosyllabic nonwords, and six 5-word lists of polysyllabic real words were presented to the children. (The lists of polysyllabic nonwords that are part of this subtest were not administered to the first grade children.) The nonwords, or nonsense words as they are called in this test, were created by changing one or two letters of a real word (e.g., grain was changed to thrain). Split-half reliability coefficients are reported to range from 0.95 to 0.99 across the three subtests.

The Developmental Spelling Test (DST). The experimenter-devised Developmental Spelling Test was used to assess developmental spelling in first grade (Tangel & Blachman 1995). This test was expanded to 10 words from the 5-word version previously used by Tangel and Blachman (1992) to assess developmental spelling at the end of kindergarten. The additional 5 words (hunt, kissed, street, order, snowing) that were added to the original kindergarten list reflected later developing spelling patterns (e.g., the ed in kissed,

and the spelling of vowel + r words). The 10-words were embedded in sentences so that the format of the test would follow traditional spelling tests given in the classroom.[3] A 7-point scale (0–6) was designed to measure the sophistication of the spelling productions, taking into consideration both the number of phonemes represented and the level of orthographic representation. Interrater reliability for this measure is reported to be $r = 0.999$ (Tangel & Blachman 1995).

The Wide Range Achievement Test-Revised (WRAT-R), Level 1. The WRAT-R (Jastak & Wilkinson 1984), Level 1 spelling subtest was individually administered to assess the student's ability to write single words from dictation. Both the standard method of scoring (number right) and a modified 4-point (0–3) scoring system were used. The modified scoring system was similar to the scoring system used for the DST in that it awarded points on the basis of developmental sophistication of the spellings that were incorrect but captured the phonetic structure of the word. Interrater reliability for the modified scoring system was reported to be $r = 0.999$ (Tangel & Blachman 1995).

Results

Effects of training on end of year performance

Phoneme segmentation. One of the questions explored in this study was whether the group that received phoneme awareness instruction in kindergarten and a first grade reading program that emphasized the alphabetic code would be better able than the control group to segment words into their constituent phonemes. Posttest results from the end of first grade were evaluated with analysis of covariance using kindergarten pretest segmentation score as the covariate. Results indicate that the treatment group performed significantly better that the control group on the phoneme segmentation posttest [$F(1,125) = 31.84$; $p < 0.0001$, see Table 1].

Letter names and sounds. At the end of first grade, differences between the two groups in letter name and letter sound knowledge were also evaluated using analysis of covariance with the appropriate kindergarten pretest as covariate. Results indicate that the treatment group performed significantly better than the control group on the test of letter name knowledge [$F(1,125) = 5.64$; $p = 0.019$], and on the test of letter sound knowledge [$F(1,125) = 26.49$; $p < 0.0001$, see Table 1].

Reading. The effects of the training on reading were evaluated using the posttest score on the Woodcock Reading Mastery Test-Revised, Word Iden-

tification subtest. To evaluate group differences on the Woodcock posttest, analysis of covariance with the pretest score as covariate was used. The posttest difference between the groups on this measure was marginally significant [$F(1,125) = 3.72$; $p = 0.056$, see Table 1].

Independent *t*-tests were used to evaluate differences between the treatment and control group on the posttest-only measures. The groups were given the Phonic Patterns subtest of the Decoding Skills Test requiring the children to read both monosyllabic and polysyllabic phonetically regular words and monosyllabic nonwords. Results indicate that the treatment children read significantly more real words than the control group [$t(126) = 4.106$; $p < 0.0001$], and the treatment group read significantly more nonwords than the control group [$t(126) = 4.877$; $p < 0.00001$]. In addition, the children were evaluated on an experimenter-devised measure of phonetically regular real words (referred to on Table 1 as Regular Words II). Results indicate that the children in the treatment group read significantly more words on this measure than the control group [$t(112) = 4.985$; $p < 0.0001$, the degrees of freedom were adjusted for unequal variances, see Table 1].

Spelling. The experimenter-devised Developmental Spelling Test (DST) (for a detailed analysis of the first grade spelling results, see Tangel & Blachman 1995), administered as a posttest-only measure, was used to evaluate the developmental sophistication of the children's spelling. Scores on the DST were analyzed using independent *t*-tests. Results indicate that the total number of points earned on the DST was significantly higher for the treatment group than the control group [$t(126) = 3.66$; $p = 0.0004$, see Table 1].

The children were also given a standardized spelling test (the WRAT-R, Level 1 Spelling subtest). Two methods of scoring the children's productions were used. First, the raw score (number right) was computed. Using this method, treatment children significantly outperformed control children [$t(126) = 2.24$; $p = 0.0264$, see Table 1]. A modified scoring system of the WRAT-R (for a detailed description of this scoring system, see Tangel & Blachman 1995) was developed to determine whether treatment children differed from control children in phonetically correct spelling (e.g., spelling light as lite). Partial credit was given for phonetically accurate spelling. An independent *t*-test was used to determine whether treatment and control children differed on the mean number of points earned as measured by their modified WRAT-R score. The treatment group showed a significantly higher level of spelling sophistication than the control group [$t(126) = 3.38$; $p = 0.001$, see Table 1].

Performance over time

Split-plot factorial 2×3 ANOVAs were performed on the average scores earned by treatment and control groups on phoneme segmentation, letter name knowledge, letter sound knowledge, and the Woodcock Word Identification subtest at each of the three administrations of these tests (January of kindergarten, end of kindergarten, and end of first grade). The sphericity assumption of repeated measures ANOVA was not met for any of the four sets of hypotheses. As a result, the critical and *p*-values were adjusted using the Greenhouse-Geisser epsilon method. In addition, the experiment-wise error rate was controlled for in all cases where multiple comparisons were made.

Phoneme segmentation. The analysis of phoneme segmentation scores indicated (see Table 3) a significant treatment effect [$F(1,126) = 73.98$; $p < 0.0001$], a significant time effect [$F(2,252) = 289.50$; $p < 0.0001$], and a significant Time × Treatment interaction [$F(2,252) = 41.22$; $p < 0.0001$]. Tests of the simple main effects were examined to further explain the source of the interaction. Significant treatment effects were found at the end of kindergarten and the end of first grade. The treatment group scored significantly higher than the control group on phoneme segmentation at both testing times, with the largest difference between the two groups occurring at the end of kindergarten ($p < 0.0001$). Significant time effects were found for the control and treatment groups. Scheffe's multiple comparison test was used to further examine the significant time effect. The treatment group significantly improved their segmentation scores from the beginning of kindergarten to the end of kindergarten ($p < 0.001$) and from the end of kindergarten to the end of first grade ($p < 0.0001$). The control group significantly improved ($p < 0.0001$) their segmentation scores only from the end of kindergarten to the end of first grade (see Figure 1).

Letter name and letter sound knowledge. The analysis of letter name knowledge indicated (see Table 4) a significant treatment effect [$F(1,126) = 5.11$; $p = 0.0254$], and a significant time effect [$F(2,252) = 384.00$; $p < 0.0001$]. The lack of a significant interaction indicates that the treatment and control group behaved similarly across the three levels of the time factor, with both groups improving on letter name knowledge at similar rates. The significant treatment effect, however, indicates that the treatment group scored significantly higher than the control group on letter name knowledge, with most of the differences due to significant differences at the end of kindergarten (see Figure 2).

The analysis of letter sound knowledge indicated (see Table 5) a significant treatment effect [$F(1,126) = 16.62$; $p < 0.0001$], a significant time

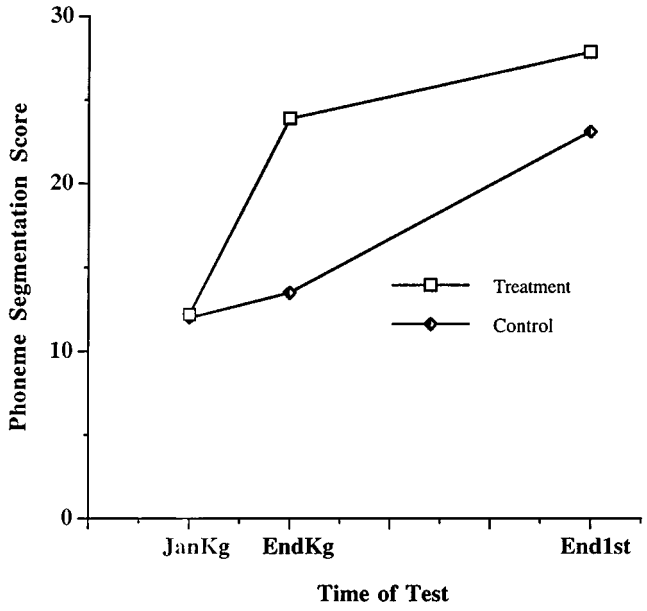


Figure 1. Phoneme segmentation mean raw scores of treatment and control groups across three test administrations.

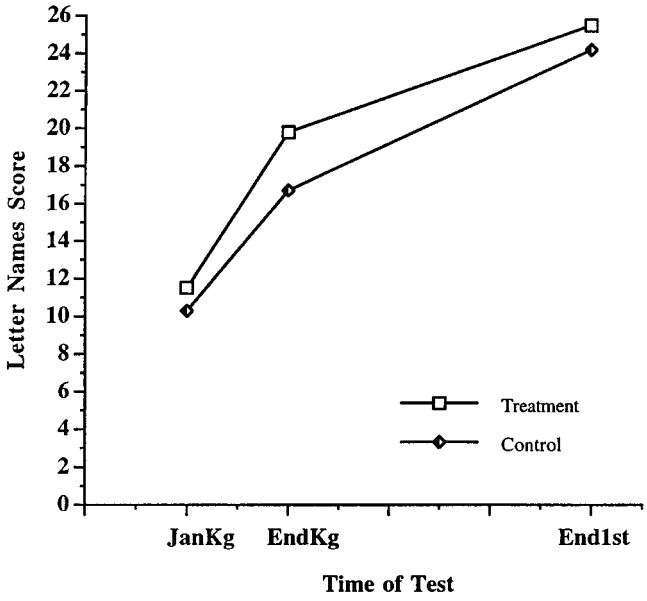


Figure 2. Letter name mean raw scores of treatment and control groups across three test administrations.

Table 3. Repeated measures analysis of variance on phoneme segmentation

Source	df	MS	F	*p*
Between subjects	127			
Treatment	1	2528.62	73.98	<0.0001
Error between	126	34.18		
Within subjects	256			
Time of test	2	5777.95	289.50	<0.0001
Time × Treatment	2	822.74	41.22	<0.0001
Error within	252	19.96		

Table 4. Repeated measures analysis of variance on letter name knowledge

Source	df	MS	F	*p*
Between subjects	127			
Treatment	1	331.47	5.11	0.0254
Error between	126	64.81		
Within subjects	256			
Time of test	2	6240.61	384.00	<0.0001
Time × Treatment	2	36.77	2.26	0.1167
Error within	252	16.25		

effect [F(2,252) = 918.32, $p < 0.0001$], and a significant Time × Treatment interaction [F(2,252) = 16.32; $p < 0.0001$]. Tests of the simple main effects indicated a significant treatment effect at the end of kindergarten ($p < 0.0001$) and the end of the first grade ($p < 0.001$). The treatment group scored significantly higher than the control group on letter sound knowledge at both testing times, with the largest difference occurring at the end of kindergarten. Significant time effects were found for the control and treatment groups. Both groups significantly improved ($p < 0.0001$) their letter sound scores from the beginning of kindergarten to the end of kindergarten and from the end of kindergarten to the end of first grade (see Figure 3).

Reading. The analysis of word identification scores on the Woodcock indicated (see Table 6) a trend level treatment effect [F(1,126) = 3.15; $p = 0.0782$], a significant time effect [F(2,252) = 568.69; $p < 0.0001$], and a significant interaction [F(2,252) = 4.42; $p = 0.0364$]. Tests of the simple main effects indicated a significant treatment effect at the end of first grade. The treatment

Table 5. Repeated measures analysis of variance on letter sound knowledge

Source	df	MS	F	*p*
Between subjects	127			
Treatment	1	597.54	16.62	<0.0001
Error between	126	35.96		
Within subjecs	256			
Time of test	2	13394.88	918.32	<0.0001
Time × Treatment	2	239.03	16.32	<0.0001
Error within	252	14.59		

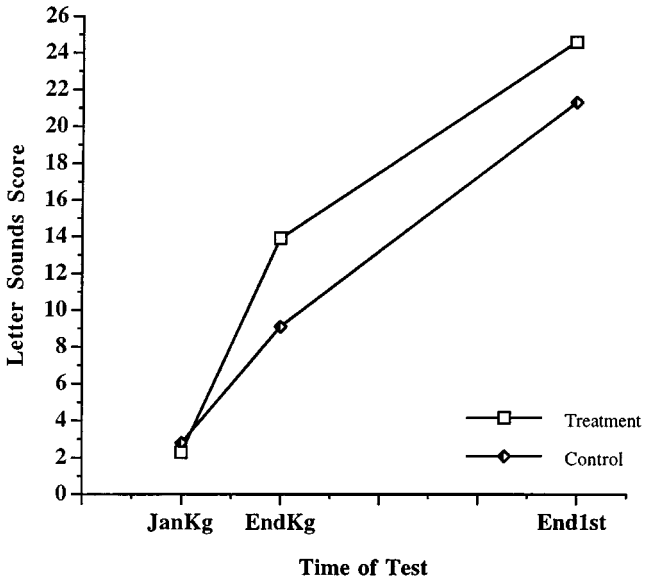


Figure 3. Letter sound mean raw scores of treatment and control groups across three test administrations.

group scored significantly higher than the control group ($p < 0.01$) on the Woodcock Word Identification subtest. Significant time effects were found for the control and treatment groups from the end of kindergarten to the end of first grade ($p < 0.0001$) (see Figure 4).

Table 6. Repeated measures analysis of variance on Woodcock Word Identification (over 3 test administrations)

Source	df	MS	F	*p*
Between subjecs	127			
Treatment	1	271.06	3.15	0.0782
Error between	126	85.99		
Within subjecs	256			
Time of test	2	40788.78	568.69	<0.0001
Time × Treatment	2	317.30	4.42	0.0364
Error within	252	71.72		

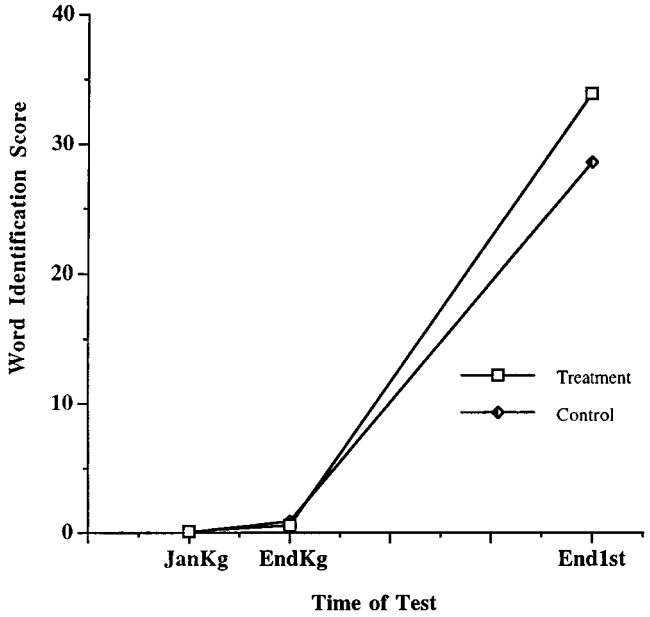


Figure 4. Woodcock Word Identification mean raw scores of treatment and control groups across three test administrations.

Grade two follow-up

It is important to point out that this was not a follow-up year in the strictest sense of the term. That is, it was not a year during which we had no contact with any of the children. As will be described below, we provided a modest amount of training for second grade teachers during the first half of the second grade year and materials for second grade children who needed them.

Table 7. Pretest means for treatment and control groups for students remaining at the end of second grade

	Treatment (n = 58)		Control (n = 48)			
Variable	Mean	sd	Mean	sd	*t*	*p*
Pretests						
PPVT-R	91.6	10.2	90.9	10.2	0.159	0.874
Phoneme segmentation	12.2	4.1	11.7	5.0	0.523	0.602
Letter names	11.3	6.6	10.3	7.3	0.740	0.461
Letter sounds	1.9	3.0	2.6	3.9	1.026	0.307
Woodcock Word ID	0.1	0.4	0.1	0.5	0.171	0.865

Method

Participants. Between June of first grade and May of second grade, we lost an additional 22 children from the study because they moved out of this school district. The treatment group lost 8 children and the control group lost 14 children. This left a sample of 106 children (58 treatment and 48 control[4]). Thus, between January of kindergarten and May of second grade, approximately one-third of the original sample of 159 children left the school district. Prior to analyzing the second grade posttest data, pretest data from kindergarten of the 106 remaining children were reanalyzed. No significant differences were found on age [treatment mean = 5.62; control mean = 5.64, $t(104) = 0.35$; $p = 0.71$], sex [Yates $\chi^2(1, n = 106) = 0.022$; $p = 0.88$], race [Yates $\chi^2(2, n = 106) = 0.662$; $p = 0.72$], or any of the other pretraining variables (see Table 7).

Procedure

Grouping for instruction and second grade reading program. At the end of the first grade year, treatment children were rank ordered on the basis of our end of year reading measures (described earlier). Recommendations were made for second grade reading instruction based on these scores and on input from first grade teachers regarding individual progress in reading and how much of the program had been completed by the child. Children who had completed our program (which meant they had been introduced to words representing all six syllable types, including closed syllables, such as fat and flat, final 'e' syllables, such as cake and shine, open syllables, such as me and cry, vowel team syllables, such as pain, teach, and crawl, vowel + r syllables, such as burn and start, and consonant le syllables, as in bottle and table)

and/or children who were reading above grade level received no further intervention. These children were assigned to second grade classrooms and placed in reading groups by their second grade teachers. We had no further contact with them until the end of second grade assessment. The remaining treatment children were placed in homogeneous reading groups and continued to receive reading instruction based on the 5-step plan (outlined previously) until they completed the program midyear in grade 2. Once the program was completed, the children were placed in the same basal reading program being used by the control children and our research team had no further contact with them until the end of year assessment. Two of the six second grade treatment teachers, one at each school, chose to continue the intervention beyond the first half of second grade with a few children (3 children at one school and 5 at the other) who needed more time to complete the program.

For children continuing in the treatment program in grade 2, the program followed the 5 steps outlined previously, and time devoted to the program was again 30 minutes. Because children were now spending more time reading in context, the sound board was used less frequently (once or twice a week) to introduce new vowel teams (e.g., oa, oi) and dictation was used approximately three times per week. Between the first grade year and second grade year, the school district changed from the Scott Foresman basal series to the Silver, Burdett and Ginn basal series (1989). Thus, the second grade children who continued in the treatment program read selected stories from the Silver, Burdett and Ginn series in addition to phonetically controlled readers from the *Primary Phonics* series. Trade books from classroom and school libraries were read at other times during the day.

Treatment children who had completed our program at the end of grade 1 and the control children used the newly adopted Silver, Burdett and Ginn basal series for their reading instruction, as mandated by the school district. Instruction for all of these children followed roughly the traditional basal approach outlined previously, supplemented at other times during the school day by reading trade books from classroom and school libraries. Six of the 48 control children also used an alternate program selected by the district – *Reading Mastery*, a structured phonetic approach – for part of their primary reading program.

Both treatment and control children participated in whole class, phonetically-based spelling instruction using *Spelling: Words and Skills* by Scott Foresman (1986) as mandated by the school district. Specifically, in the second grade speller, the first grade concepts (e.g., short and long vowels) were reviewed and expanded and lists of words containing vowel teams (e.g., play, plain, keep, dream) and vowel + r were introduced (e.g., car, start).

Teacher training. Second grade teachers of the treatment children participated in 3 inservice training sessions (one session in September, one in October, and one in December, totalling approximately 6 hours) to familiarize them with the 5-step reading program. Compliance with the program was monitored much less frequently than during the first grade year and participation on the part of the teachers during the second grade year was voluntary.

Posttesting. At the end of May of the second grade year, children who remained within this school district were again assessed by examiners blind to whether children had been in the treatment or control group. Whereas when we began the study, the 159 children in the original sample were located in 4 different schools, at the end of second grade, the 106 who remained in the district were now located in 14 of the 21 elementary schools in this urban district. Children were retested on several measures of word recognition, including the Woodcock Word Identification subtest (Woodcock 1987), the experimenter-devised phonetically regular word list (Regular Words II), the Phonic Patterns subtest from The Decoding Skills Test (Richardson & DiBenedetto 1985) (including administration of six 5-word lists of polysyllabic nonwords that had not been administered the previous year), and the spelling subtest from the Wide Range Achievement Test-Revised (WRAT-R) (Jastak & Wilkinson 1984).

Results

Effects of training on end of year performance

Reading. The effects of the training on reading were also evaluated at the end of second grade using posttest scores on the Woodcock Word Identification subtest. Group differences on the Woodcock posttest were evaluated using analysis of covariance with the kindergarten pretest score as the covariate. The posttest difference between the groups on this measure was significant [$F(1,103) = 4.92$; $p = 0.0287$, see Table 1].

Independent *t*-tests were used to evaluate differences between the treatment and control group on the posttest-only measures. The groups were given the Phonic Patterns subtest of the Decoding Skills Test. For the end of second grade administration of this test, children were required to read monosyllabic and polysyllabic phonetically regular words and monosyllabic and polysyllabic nonwords. (The polysyllabic nonwords were not administered at the end of first grade, but were added to the end of second grade assessment.) Results indicate that the treatment children read significantly more real words than the control group [$t(84.0) = 2.240$; $p = 0.0278$, the degrees of freedom were adjusted for unequal variances], and the treatment

children read significantly more nonwords than the control group [$t(104) = 2.370$; $p = 0.0196$, see Table 1].

The children were also evaluated on the experimenter-devised measure of phonetically regular real words (referred to as Regular Words II on Table 1). Results indicate that the treatment children read significantly more phonetically regular words than the control group [$t(70.1) = 3.154$; $p = 0.0024$, the degrees of freedom were adjusted for unequal variances].

Spelling. The only measure of spelling given at the end of second grade was the standardized spelling test, the WRAT-R, Level 1 Spelling subtest, that was also given at the end of grade 1. As with the first grade data, two methods of scoring were used for this spelling test, the standard method and a modified method giving partial credit for phonetically accurate responses (e.g., spelling light as lite) (for detailed scoring procedures, see Tangel & Blachman 1995). There were no significant differences between the treatment and control children using either the standard scoring method [$t(86.3) = 0.078$; $p = 0.9378$, the degrees of freedom were adjusted for unequal variances], or the modified method [$t(83.7) = 1.050$; $p = 0.2967$, the degrees of freedom were adjusted for unequal variances, see Table 1].

To investigate spelling further, we rank ordered children in the combined sample of treatment and control children on the basis of raw scores on the WRAT-R spelling subtest and identified those in the lowest quartile. Twenty-nine children (16 treatment children and 13 control children), or 27.4 percent of the sample, were identified when we included children who had scores that tied with those in the lowest quartile. As shown in Table 8, there were no significant differences between these low scoring spellers in the treatment and control groups on pretest measures administered at the beginning of kindergarten. Independent *t*-tests were used to evaluate differences between these treatment and control children at the end of second grade. When the standard method of scoring was used on the WRAT-R, there was not a statistically significant difference between the means of these groups [$t(27) = 1.5503$; $p = 0.1327$], although the scores favor the treatment children. (It should be noted that the means for the treatment (mean = 10.4, sd = 2.4) and control children (mean = 8.8, sd = 3.4) in the lowest quartile of spellers at the end of grade 2 are quite similar to the means generated one year earlier by our entire sample of 66 treatment children (mean = 10.2, sd = 3.7) and 62 control children (mean = 8.6, sd = 4.7) at the end of grade 1. With these larger sample sizes one year earlier, the difference between the means was statistically significant, as shown in Table 1.) When the modified scoring system for the WRAT-R (giving partial credit for phonetically accurate responses) was used, there was a significant difference between the means of the treat-

Table 8. Pretest and posttest means for treatment and control children in the bottom quartile of spellers at the end of grade 2

	Treatment (n = 16)		Control (n = 13)		
Variable	Mean	sd	Mean	sd	*p*
Pretests					
January of kindergarten					
PPVT-R	88.8	7.3	86.0	7.0	0.3146
Phoneme segmentation	11.4	4.5	9.2	4.1	0.1830
Letter names	7.5	5.7	6.2	3.8	0.5009
Letter sounds	0.9	2.5	0.6	0.8	0.7014
Woodcock Word ID	0.0	0.0	0.0	0.0	–
Posttests					
End of second grade					
Woodcock Word ID	42.6	6.9	30.9	12.0	0.0059
Decoding Skills Test					
Real words	28.5	12.5	15.2	10.2	0.0047
Nonwords	21.1	12.7	7.9	6.2	0.0014
Regular Words II	39.1	7.4	25.6	11.6	0.0008
WRAT-R spelling	10.4	2.4	8.8	3.4	0.1327
Modified WRAT-R	49.4	11.3	36.8	14.7	0.0140

ment (mean = 49.4, sd = 11.3) and control children (mean = 36.8, sd = 14.7) in our lowest quartile of spellers [$t(27) = 2.629$; $p = 0.0140$, see Table 8]. Perhaps even more important in terms of future literacy success are the significant differences between these groups of low spellers, again favoring the treatment children, on the four reading measures administered at the end of grade 2, the Woodcock Word Identification subtest [$t(18.3) = 3.116$; $p = 0.0059$, the degrees of freedom were adjusted for unequal variances], Decoding Skills Test – Real Words [$t(27) = 3.079$; $p = 0.0047$], Decoding Skills Test – Nonwords [$t(22.7) = 3.643$; $p = 0.0014$, the degrees of freedom were adjusted for unequal variances], and Regular Words II [$t(27) = 3.791$; $p = 0.0008$, see Table 8].

Performance over time

Reading. A split-plot factorial 2×4 ANOVA was performed on the average scores earned during each of the four administrations of the Woodcock Word Identification subtest (January of kindergarten, end of kindergarten, end of first grade, and end of second grade) by the 106 treatment and control children

Table 9. Repeated measures analysis of variance on Woodcock Word Identification (over 4 test administrations)

Source	df	MS	F	*p*
Between subjects	105			
Treatment	1	593.66	3.38	0.0689
Error between	104	175.74		
Within subjects	318			
Time of test	3	68395.97	1062.95	<0.0001
Time × Treatment	3	241.39	3.75	0.0386
Error within	312	64.35		

who remained at the end of second grade. The sphericity assumption of repeated measures ANOVA was not met for the set of hypotheses associated with this analysis. As a result, the critical and *p*-values were adjusted using the Greenhouse-Geisser epsilon method. In addition, the experiment-wise error rate was controlled for in all cases where multiple comparisons were made.

The analysis of the word identification scores for all four testing times (n = 106) indicated (see Table 9) a trend level treatment effect [$F(1,104) = 3.38$; $p = 0.0689$], a significant time effect [$F(3,312) = 1062.95$; $p < 0.0001$], and a significant Time × Treatment interaction [$F(3,312) = 3.75$; $p = 0.0386$]. Tests of the simple main effects indicated a significant treatment effect at the end of second grade, indicating that the treatment group scored significantly higher on word identification ($p < 0.05$). Significant time effects were found for the control and the treatment groups. Scheffe's multiple comparison test was used to further examine the significant time effect. Both the treatment and control groups significantly improved ($p < 0.0001$) their word identification scores from the end of kindergarten to the end of first grade and from the end of first grade to the end of second grade (see Figure 5).

Discussion

The major finding from this study is that children who participated in a phonological awareness program in kindergarten followed by a reading program in grade 1 (and extended to grade 2 for some children) that built on this awareness and emphasized explicit instruction in the alphabetic code demonstrated a significant advantage in reading at the end of grades 1 and 2. Specifically, at the end of grade 1 the treatment children significantly outperformed the control children on standardized measures of phonetically regular

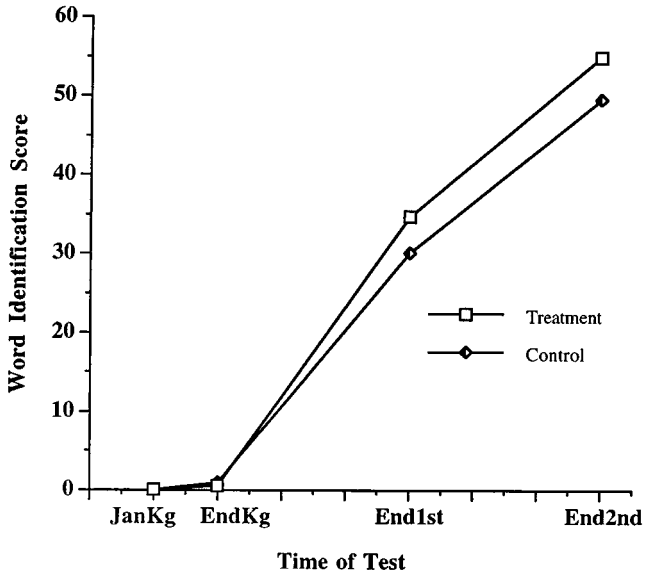


Figure 5. Woodcock Word Identification mean raw scores of treatment and control groups across four test administrations.

words and nonwords and on an experimenter-devised measure of phonetically regular words, with performance on a standardized measure of word recognition reaching marginal significance ($p = 0.056$). At the end of grade 2, the treatment children read significantly more words on all four measures of word recognition. It is important to note that nonwords were neither taught nor practiced during the intervention and, as such, serve as a measure of the child's ability to transfer decoding skills to words never seen before. At the end of grades 1 and 2, treatment children were better able than control children to read nonwords.

With regard to spelling, at the end of first grade, treatment children demonstrated significantly superior performance on our Developmental Spelling Test and on a standardized measure of spelling, the Wide Range Achievement Test-R (WRAT-R) (for details see Tangel & Blachman 1995). One year later, there were no significant differences on the WRAT. We did not readminister the Developmental Spelling Test at the end of grade 2, so we don't know if that measure might have detected developmental differences between the groups. When we looked more closely just at the children in the lowest quartile of our sample in spelling, we found striking differences favoring the treatment children on the adjusted scores on the WRAT-R (i.e., giving partial credit for phonetically correct spelling, such as writing lite for light), and on our four measures of word recognition. From comparing the second

grade data in Tables 1 and 8, it appears that the lowest scoring children in the treatment group (Table 8) differentially benefitted from the intervention when compared to the treatment group as a whole (Table 1).

It is important to remember that both treatment and control children participated in the same phonetically-based spelling program during grades 1 and 2, as required by the school district. (Only the general classroom reading programs of the treatment and control children differed – and then, the programs differed only during the thirty minutes of group instruction provided to both groups of children.) One interpretation of the spelling data is that the lack of differences between groups is a testimonial to the benefits of a phonetically-based spelling program. After two years of this program, the control children were no longer behind the treatment children on a standardized measure of spelling. The fact that 11 of the first 17 words dictated on the WRAT-R were included in the district mandated spelling program used by both groups also may have contributed to the lack of differences on the standardized spelling measure. Although by the end of second grade the control children had reached a level of spelling that was equivalent to that of the treatment children, their improved second grade spelling performance did not appear to transfer to reading words. That is, the treatment children who had the benefit of phonological awareness training in kindergarten and a phonetic approach for both reading and spelling remained significantly ahead of the control children in reading words and nonwords at the end of second grade.

Two additional features of this research that we find particularly important are that (1) all instruction was provided to groups of children by regular classroom teachers in general education classrooms, and (2) the children involved in these studies were from low-income, inner-city schools. Most of the children in our study received free lunch (used as a measure of SES), were in the low average range on a measure of receptive vocabulary (i.e., mean = 91) and had limited knowledge of the alphabet in the middle of kindergarten as indicated by knowledge of only two letter sounds when our intervention began. The children attended 4 of the 5 lowest achieving elementary schools (out of 21 elementary schools) in this large, urban district. These are the schools that often have the fewest resources and yet serve children who come to school with the greatest needs. As Hiebert has suggested, we should be searching for ways to ‘maintain the features of effective instruction in low-income schools’ (1994: 24). Our results indicate that it is possible to enhance the reading success of low-income children who begin school with limited literacy skills and who go to low-achieving schools by providing reading instruction in the early grades that emphasizes phonological awareness and provides explicit instruction in the alphabetic code. A variety of modifica-

tions could have been added to the treatment, such as the involvement of parents, making sure additional opportunities existed throughout the day to practice new skills in the context of additional reading, extending the 30-minute block of instructional time, and making sure that spelling and writing activities were integrated with what was being taught in reading. Although it is possible that these curricular changes would have added to the benefit provided to the treatment children, it is encouraging to us that it is possible to provide demonstrable benefits to children by changing very little and by working within the time, personnel, and regular classroom constraints of a large, urban district.

In an attempt to evaluate our instructional model in as normal a classroom setting as possible, we also worked within the group structure already in place (this resulted in instructional groups between 6 and 9 children – larger than we would have liked) and involved no additional teaching personnel beyond those employed in these classrooms by the school district. After providing our experimental children within 10 to 13 hours of phoneme awareness training over an 11 week period in kindergarten, the treatment children received no additional instructional time in reading in grades 1 and 2, beyond what was routinely provided for all children in this school district. The type of reading instruction provided to the treatment children was systematic and focused on explicit teaching of the alphabetic code. Teachers were encouraged to pace the lessons so that high rates of word reading accuracy were maintained, as opposed to just 'covering' the material in order to reach a certain lesson or complete a set number of stories (for a detailed discussion of the value of maintaining high rates of word reading accuracy in first grade, see Juel 1994). Such teaching resulted in significantly higher levels of word recognition for the treatment children at the end of grades 1 and 2 and demonstrated that strategies to promote phonological awareness and alphabetic coding can be incorporated into general classroom instruction **before** children have had a chance to fail.

Another integral part of the instructional model that we used was the in-service training received by the teachers and their teaching assistants. During these sessions, teachers learned about the role of phonological processes in learning to read and were provided with a theoretical framework to support the teaching of phonological awareness and explicit teaching of the alphabetic code. There is evidence that teachers typically lack this information (Liberman 1987; Moats 1995) and that they are 'undereducated for the very demanding task of teaching reading and spelling explicitly' (Moats 1995: 43). Given that teacher training programs often fail to address the beginning reading skills that research has shown to be critical to reading success, the need to provide training appears essential.

As discussed in the report of the kindergarten year (Blachman et al. 1994), there are limitations to studies conducted in naturalistic settings. We did not, for example, utilize strict random assignment. Although treatment and control children did not differ on important pretreatment linguistic and cognitive variables, the children attended different schools. This helped to maintain the integrity of the experimental treatment, but leaves open the possibility that individual school factors influenced the outcome of the study.

The results of our study provide clear evidence of transfer from training in phonological awareness and alphabetic coding to measures of word recognition. We do not, however, have direct evidence of the effects of our treatment on fluency and reading comprehension. Despite substantial evidence that reading comprehension in later grades is highly correlated with the ability to decode in grades 1 and 2 (see, for example, Foorman, Francis, S.E. Shaywitz, B.A. Shaywitz & Fletcher 1997; Juel 1994), there is a need to investigate the influence on reading comprehension of interventions, like the one reported here, that strive to ensure that children learn to decode early in their school careers. Some indirect support for greater fluency of word recognition among the treatment children is provided by their superior scores on the Woodcock at the end of grade 2, a test that requires 'a natural reading of the word in about 5 seconds' (Woodcock 1987). We also learned from interviews with teachers at the end of the second grade year that while 31% (n = 15) of the control children were still reading in a first grade reader, only 5% (n = 3) of the treatment children were still reading in a first grade reader. Teachers also reported recommending 10 of the control children for retention; whereas none of the children who had participated in our intervention were recommended for retention at the end of second grade. Retention was determined primarily by end of year performance on a reading test created by the school district. Although end of year book placement and retention data are suggestive of group differences in reading ability, we cannot rule out the influence that individual school factors and differential teacher training may have had on these data. These school and teacher training factors are mitigated somewhat by the fact that our treatment and control children were dispersed across 14 of the 21 elementary schools in this urban district when we last tested the children at the end of second grade.

It is also important to point out that design limitations, due to budgetary constraints, leave important questions unanswered. We don't know, for example, how our treatment and control children would have compared to a group of children who received **only** the phonological awareness program in kindergarten or how our groups would have compared to a group of children who received **only** the phonetically-based program in grade 1 (and grade 2 if necessary) without a phoneme awareness program in kindergarten. A study

like the one reported here but that also included these two additional groups would add considerably to our knowledge about the value of different models of early reading intervention.

Acknowledgments

This article was supported by a research grant to the first author from the National Center for Learning Disabilities. We wish to thank the children, teachers, and administrators of the Syracuse City School District, Syracuse, New York, for their participation in this project. We also want to thank Frank Vellutino for his insights regarding our spelling data and Richard Olson for his feedback on an earlier draft of this manuscript.

Notes

1. It should be noted that most of our treatment children were in reading groups that were larger than we would recommend. One of our goals, however, was to integrate our reading program into the existing organizational structure of the classroom. Teachers generally had 3 to 4 reading groups in their classrooms, only one of which was involved in our study. In order to make sure that our treatment could be conducted without additional classroom personnel and within the time constraints for group reading instruction, we created groups that approximated the size of the other reading groups in these classrooms.
2. No attempt was made to teach the treatment children to read the words on the experimenter-devised, phonetically regular word list. It is possible that children encountered some of these words during reading instruction. We did not control for exposure to these words.
3. No attempt was made to teach the treatment children to spell the words on the Developmental Spelling Test (DST). As noted in Tangel and Blachman (1995), a review of the district mandated spelling program by Scott Foresman (1986) revealed that one of the words on the DST, the word *snow* from the test word *snowing*, was taught to both treatment and control children. We have no reason to believe that treatment and control children had differential exposure to the 10 words on the DST.
4. Based on the results of an end of grade 1 reading test created by the school district, 3 of the 48 control children and none of the treatment children were required to repeat first grade. These 3 control children were retained in our control sample and included with the grade 2 children for all analyses.

References

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

Adams, M.J. & Bruck, M. (1993). Word recognition: The interface of educational policies and scientific research, *Reading and Writing: An Interdisciplinary Journal* 5: 113–139.

Adams, M.J. & Bruck, M. (1995). Resolving the 'Great Debate', *American Educator* 19: 7, 10–20.

Alexander, A.W., Anderson, H.G., Heilman, P.C., Voeller, K.K.S. & Torgesen, J.K. (1991). Phonological awareness training and remediation of analytic decoding deficits in a group of severe dyslexics, *Annals of Dyslexia* 41: 193–206.

Ball, E.W. & Blachman, B.A. (1988). Phoneme segmentation training: Effect on reading readiness, *Annals of Dyslexia* 38: 208–225.

Ball, E.W. & Blachman, B.A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling?, *Reading Research Quarterly* 26: 49–66.

Beck, I.L. & Juel, C. (1995). The role of decoding in learning to read, *American Educator* 19: 8, 21–25, 39–42.

Blachman, B.A. (1984). Relationship of rapid naming ability and language analysis skill to kindergarten and first grade reading achievement, *Journal of Educational Psychology* 76: 610–622.

Blachman, B.A. (1987). An alternative classroom reading program for learning disabled and other low achieving children. In: R. Bowler (ed.), *Intimacy with language: A forgotten basic in teacher education* (pp. 49–55). Baltimore, MD: Orton Dyslexia Society.

Blachman, B.A. (ed.) (1997). *Foundations of reading acquisition and dyslexia: Implications for early intervention*. Mahwah, NJ: Erlbaum.

Blachman, B.A., Ball, E., Black, R. & Tangel, D. (1994). Kindergarten teachers develop phoneme awareness in low-income, inner-city classrooms: Does it make a difference?, *Reading and Writing: An Interdisciplinary Journal* 6: 1–17.

Bradley, L. & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection, *Nature* 30: 419–421.

Brady, S. & Shankweiler, D. (eds.) (1991). *Phonological processes in literacy: A tribute to Isabelle Y. Liberman*. Hillsdale, NJ: Erlbaum.

Byrne, B. & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children, *Journal of Educational Psychology* 83: 451–455.

Byrne, B. & Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A 1-year follow-up, *Journal of Educational Psychology* 85: 104–111.

Byrne, B. & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial, *Journal of Educational Psychology* 87: 488–503.

Castle, J.M., Riach, J. & Nicholson, T. (1994). Getting off to a better start in reading and spelling: The effects of phonemic awareness instruction within a whole language program, *Journal of Educational Psychology* 86: 350–359.

Chall, J.S. (1989). Learning to read: The great debate 20 years later: A response to 'Debunking the great phonics myth', *Phi Delta Kappan* 70: 521–538.

Ehri, L.C. (1991). Learning to read and spell words. In: L. Rieben & C.A. Perfetti (eds.), *Learning to read: Basic research and its implications* (pp. 57–73). Hillsdale, NJ: Erlbaum.

Ehri, L.C. (1997). Sight word learning in normal readers and dyslexics. In: B. Blachman (ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 163–189). Mahwah, NJ: Erlbaum.

Engelmann, S. (1969). *Preventing failure in the primary grades*. Chicago, IL: Science Research Associates.

Felton, R.H. (1993). Effects of instruction on the decoding skills of children with phonological-processing problems, *Journal of Learning Disabilities* 26: 583–589.

Fletcher, J.M., Shaywitz, S.E., Shankweiler, D.P., Katz, L., Liberman, I.Y., Stuebing, K.K., Francis, D.J., Fowler, A.E. & Shaywitz, B.A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions, *Journal of Educational Psychology* 86: 6–23.

Foorman, B.R., Francis, D.J., Shaywitz, S.E., Shaywitz, B.A. & Fletcher, J.M. (1997). The case for early reading intervention. In: B.A. Blachman (ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 243–264). Mahwah, NJ: Erlbaum.

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

Gough, P.B., Ehri, L.C. & Treiman, R. (eds.) (1992). *Reading acquisition*. Hillsdale, NJ: Erlbaum.

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.

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.

Hiebert, E.H. (1994). Reading recovery in the United States: What difference does it make to an age cohort?, *Educational Researcher* 23(9): 15–25.

Høien, T., Lundberg, I., Stanovich, K.E. & Bjaalid, I. (1995). Components of phonological awareness, *Reading and Writing: An Interdisciplinary Journal* 7: 171–188.

Iversen, S. & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery Program, *Journal of Educational Psychology* 85: 112–126.

Jastak, S. & Wilkinson, G.S. (1984). *Wide Range Achievement Test-Revised*. Wilmington, DE: Jastak Associates.

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. (1994). *Learning to read and write in one elementary school*. New York/Berlin: Springer Verlag.

Liberman, A.M., Cooper, F.S., Shankweiler, D. & Studdert-Kennedy, M. (1967). Perception of the speech code, *Psychological Review* 74: 731–761.

Liberman, I.Y. (1971). Basic research in speech and lateralization of language: Some implications for reading disability, *Bulletin of the Orton Society* 21: 72–87.

Liberman, I.Y. (1973). Segmentation of the spoken word and reading acquisition, *Bulletin of the Orton Society* 23: 65–67.

Liberman, I.Y. (1987). Language and literacy: The obligation of the schools of education. In: W. Ellis (ed.), *Intimacy with language: A forgotten basic in teacher education* (pp. 1–9). Baltimore: The Orton Dyslexia Society.

Liberman, I.Y. & Liberman, A.M. (1990). Whole language vs. code emphasis: Underlying assumptions and their implications for reading instruction, *Annals of Dyslexia* 40: 51–76.

Liberman, I.Y. & Shankweiler, D. (1979). Speech, the alphabet, and teaching to read. In: L.B. Resnick & P.A. Weaver (eds.), *Theory and practice of early reading*, Vol. 2 (pp. 109–134). Hillsdale, NJ: Erlbaum.

Liberman, I.Y., Shankweiler, D., Fischer, F.W. & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child, *Journal of Experimental Child Psychology* 18: 201–212.

Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade children, *Reading Research Quarterly* 26: 234–250.

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

Lundberg, I., Frost, J. & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children, *Reading Research Quarterly* 23: 263–284.

Lundberg, I., Olofsson, A. & Wall, S. (1980). Reading and spelling skill in the first school years predicted from phonemic awareness skills in kindergarten, *Scandinavian Journal of Psychology* 21: 159–173.

Mann, V.A. & Liberman, I.Y. (1984). Phonological awareness and verbal short-term memory: Can they presage early reading problems?, *Journal of Learning Disabilities* 17: 592–599.

Moats, L.C. (1995). The missing foundation in teacher education, *American Educator* 19(2): 9, 43–51.

O'Connor, R.E., Notari-Syverson, A. & Vadasy, P.F. (1996). Ladders to literacy: The effects of teacher-led phonological activities for kindergarten children with and without disabilities, *Exceptional Children* 63: 117–130.

Richardson, E. & DiBenedetto, B. (1985). *Decoding Skills Test*. Parkton, MD: York Press.

Rieben, L. & Perfetti, C.A. (eds.) (1991). *Learning to read: Basic research and its implications*. Hillsdale, NJ: Erlbaum.

Scott Foresman Spelling: Words and Skills (1986). Glenview, IL: Scott, Foresman and Company.

Share, D.J., Jorm, A.F., Maclean, R. & Matthews, R. (1984). Sources of individual differences in reading achievement, *Journal of Educational Psychology* 76: 1309–1324.

Silver Burdett & Ginn Reading (1989). Needham, MA: Silver Burdett and Ginn.

Slingerland, B.H. (1971). *A multi-sensory approach to language arts for specific language disability children: A guide for primary teachers*. Cambridge, MA: Educators Publishing Service.

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

Stanovich, K.E. (1994). Romance and reality, *The Reading Teacher* 47: 280–291.

Stanovich, K.E., Cunningham, A.E. & Cramer, B.B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability, *Journal of Experimental Child Psychology* 38: 175–190.

Tangel, D. & Blachman, B.A. (1992). Effect of phoneme awareness instruction on kindergarten children's invented spelling, *Journal of Reading Behavior* 24: 233–261.

Tangel, D. & Blachman, B.A. (1995). Effect of phoneme awareness instruction on the invented spelling of first grade children: A one year follow-up, *Journal of Reading Behavior* 27: 153–185.

Torgesen, J.K. & Burgess, S.R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal, correlational and instructional studies. In: J. Metsala & L. Ehri (eds.), *Word processing in beginning literacy* (pp. 161–188). Mahwah, NJ: Erlbaum.

Torgesen, J.K., Wagner, R.K. & Rashotte, C.A. (1997a). Approaches to the prevention and remediation of phonologically based reading disabilities. In: B. Blachman (ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 287–304). Mahwah, NJ: Erlbaum.

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

Vellutino, F.R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction, *Journal of Educational Psychology* 83: 437–443.

Vellutino, F.R. & Scanlon, D.M. (1987). Phonological coding, phonological awareness, and reading ability: Evidence from a longitudinal and experimental study, *Merrill-Palmer Quarterly* 33: 321–363.

Vellutino, F.R., Scanlon, D.M., Sipay, E.R., Small, S.G., Pratt, A., Chen, R.S. & 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: 607–638.

Williams, J.P. (1979). The ABD's of reading: A program for the learning disabled. In: L.B. Resnick & P.A. Weaver (eds.), *Theory and practice of early reading*, Vol. 3 (pp. 179–195). Hillsdale, NJ: Erlbaum.

Williams, J.P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending, *Journal of Educational Psychology* 72: 1–15.

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: Lexington Books.

Williams, J.P (1987). Educational treatments for dyslexia at the elementary and secondary levels. In: W. Ellis (ed.), *Intimacy with language: A forgotten basic in teacher education* (pp. 24–32). Baltimore: The Orton Dyslexia Society.

Williams, J.P. (1994). Twenty years of research on reading: Answers and questions. In: F. Lehr & J. Osborn (eds.), *Reading, language, and literacy: Instruction for the twenty-first century* (pp. 59–73). Hillsdale, NJ: Erlbaum.

Woodcock, R.W. (1987). *Woodcock Reading Mastery Tests-Revised*. Circle Pines, MN: American Guidance Service.

Address for correspondence: Benita A. Blachman, Reading and Language Arts, 200 Huntington Hall, Syracuse University, Syracuse, NY 13244-2340, USA
Phone: (315) 443-4755; Fax: (315) 443-9878; E-mail: blachman@sued.syr.edu