Journal of Experimental Child Psychology **72,** 271–304 (1999)
Article ID jecp.1999.2490, available online at http://www.idealibrary.com on IDEAL®

Training Phonological Awareness with and without Explicit Attention to Articulation

Barbara W. Wise, Jerry Ring, and Richard K. Olson

University of Colorado

One hundred twenty-two second- to fifth-grade (7- to 11-year-old) children with reading difficulties studied phonological awareness with or without explicit attention to articulation and with or without manipulation of sounds. They all studied identical phonics and read stories on the computer with speech and decoding support for difficult words. Regular-instruction controls received regularly scheduled language-arts or reading activities. After 40 h of training, children in all three trained conditions outperformed controls on all tests except math. Conditions that manipulated sounds showed advantages over the condition without explicit practice manipulating sounds, but only on the two measures of phonological awareness. Articulatory awareness training yielded no unique benefits during this training period. Individual differences in response to treatment related to initial levels of phonological awareness, naming speed, IQ, and grade. The similar outcomes of the three conditions suggest that specific variations in good phonological training may be less important than once thought for most children with reading difficulties.

Key Words: phonological awareness; articulation; reading; remediation; reading disabilities; computer-assisted instruction.

INTRODUCTION

Most researchers and clinicians in the field of reading disabilities concur that children with specific reading disability (SRD) have a language-based learning disability. Many agree that some training in phonological awareness and analysis is a necessary, but not sufficient, part of good reading instruction for these children. However, little or no consensus exists about exactly how that training should be done or about how much of it is needed. The current study was

We thank the principals, staff, and students of Coal Creek, Crestview, Heatherwood, Louisville, and Kohl Elementary Schools of Boulder Valley Schools for participating in this study. We thank our programmer Jennifer Restrepo and lab manager Laura Kriho, and our research assistants Bonnie Houkal, Mina Johnson, Martha Miller, Dale Peters, and Zoe Broomgard for their help testing the students. We especially thank our trainers Heather Burke, Robyn Krause, Sue Parrett, Luanne Sessions, and Joanna Stewart for their skills and efforts training the children with a guided discovery approach. We thank NICHD for supporting the research with grants HD 11683 and HD 22223 to Richard K. Olson and Barbara Wise.

Address reprint requests to Barbara W. Wise, Department of Psychology, University of Colorado, Boulder, Campus Box 345, Muenzinger Psychology Building, Boulder, CO 80309.

0022-0965/99 $30.00

designed to help answer the first of these questions about how the training should be done for different children. Before describing the study, the rest of the introduction discusses some of the research background that supports the above statements about causes and cures for SRD.

The primary reading deficit in most cases of SRD is in reading single words. These problems usually stem from difficulties in “phonological decoding,” or the ability to translate print into sound. Many children also have somewhat independent problems in “orthographic coding,” or the ability to remember word-specific spelling patterns. Without strong phonological decoding skills, children with SRD have trouble reading new words (Share, 1995). Poor decoding hinders comprehension directly when words are misidentified. But a struggle to decode words also hinders comprehension indirectly, by sapping attentional resources, leaving little available to understand the text (Perfetti, 1985). The deficits in phonological decoding are linked to deficits in “phoneme awareness,” or the awareness of speech sounds within *spoken* syllables, measured by having children count, delete, or reorder sounds within spoken syllables (Wagner, Torgesen, & Rashotte, 1994; Yopp, 1988). Several lines of evidence suggest that the relations between deficits in phoneme awareness, phonological decoding, and printed word recognition are correlational, causal, and partly biological (Lyon, 1995; Wise, 1991).

Many studies with early readers and prereaders have demonstrated that phonological skills can be improved with training (Ball & Blachman, 1991; Bradley & Bryant, 1983; Byrne & Fielding-Barnsley, 1993; Cunningham, 1989; Lundberg, Frost, & Peterson, 1988; Schneider, Kuspert, Mechtild, & Marx, 1997; Uhry & Shephard, 1993; Vellutino et al., 1996; Wise, Olson, & Lindamood, 1993). Some recent studies have focused specifically on trying to treat the underlying phonological deficit in children with or at risk for SRD (Felton & Brown, 1990; Foorman et al., 1997; Hatcher, Hulme, & Ellis, 1994; Kennedy & Backman, 1993; Lovett et al., 1994; Olson, Wise, Ring, & Johnson, 1997; Torgesen, Wagner, Rashotte, Alexander, & Conway, 1997; Vellutino et al., 1996; Wise & Olson, 1995). This research has shown that training these children in phoneme awareness and/or phonological decoding led to substantial improvement in those skills during and at the end of short- to moderate-length training periods. All of these studies showed gains in word recognition compared to non-reading-trained controls, and some of the studies also showed more modest differential gains for growth in standardized measures of word recognition, compared to other structured, but nonphonological, reading programs.

The above studies support that phonological training helps children with SRD. However, they do not specify how much or what kind of training is optimal, nor whether the optimal training program may vary for different children with SRD. Certainly awareness of speech sounds in syllables must, by definition, be included in phoneme awareness training. But how best to build that awareness, especially for children experiencing severe difficulties in the area, is still uncertain.

The present study was designed to compare the benefits from different elements that are included in some but not all types of phonological awareness training. This study compared three versions of phonological training: a combination condition included explicit instruction of articulatory awareness and explicit manipulation of sounds in syllables in analytic exercises with speech and print. The other two conditions left out either the instruction in articulatory awareness or the explicit manipulation of sounds in syllables. Yet all three conditions included the same amount of instructional time in small groups, individualized with the computer, and in total. The study also examined whether children with different initial profiles might vary systematically in their benefits from the different types of instruction. Before the design of the study is described, the next section frames its context by reviewing other training studies that used phonological manipulation and articulatory awareness and by discussing the theoretical support for training articulatory awareness for children with SRD.

Many instructional programs for beginning and remedial readers include some kinds of manipulation of sounds. In some programs, children use tokens, pictures, or blocks to represent sound changes in syllables (Ball & Blachman, 1991; Elkonin, 1973; Lindamood & Lindamood, 1975; Skjelfjord, 1976, 1987). Others use letters in spelling manipulation exercises (Ball & Blachman, 1991; Lindamood & Lindamood, 1975; Wise & Olson, 1995). Uhry and Shephard (1993) found strong benefits for phoneme awareness when first graders explicitly learned to manipulate sounds and letters in syllables, compared to similar children who studied the same words in reading and word family exercises but spent no time manipulating their sounds.

Many practitioners believe that “multisensory” work including all senses and some kind of kinesthetic feedback is stronger than sound-symbol work alone for children with SRD (see Clark & Uhry, 1995). This makes sense theoretically in terms of depth of processing. But what kind of kinesthetic feedback is stressed differs in different programs. Orton–Gillingham multisensory phonics associates kinesthetic manual movements with letters and sounds while using some oral-motor feedback (Clark & Uhry, 1995). The Auditory Discrimination in Depth (A.D.D.; Lindamood & Lindamood, 1975) method includes a systematic exploration of the articulation of phonemes (Truch, 1994).

Wise & Olson (1995) incorporated many of the Lindamoods’ concepts for their supplemental training in phonological awareness. They included training in articulatory awareness for several reasons. First of all, Montgomery (1981) had shown that children with SRD have difficulty accessing articulatory information, which suggested a reasonable underpinning for difficulties in hearing the order of sounds in syllables. Decades of work at Haskin’s laboratories had built a strong case that phonemes are perceived according to how they are produced (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Shankweiler, 1985). It seemed plausible that precise attention to and focused practice of speech movements could refine an “indistinct phonological representation” in the brain

(Snowling & Hulme, 1994). A final pragmatic benefit of developing an articulatory base for phonological work is that children can have access to their "language lab" base all their lives and in all situations. Therefore, the method seemed likely to help develop strong concepts and strategies that could transfer long-term to applied reading and spelling. While many educators think of this training as beneficial for early readers, studies suggest it helps poor readers of many ages with phonological problems (Alexander et al., 1991; Truch, 1994).

Wise and Olson (1995) decided to include articulatory and phonological training because of the above research and theory and because of results of their own prior studies of computer-supported reading. In Olson and Wise's (1992) studies, third- through sixth-grade children with SRD read stories on the computer, using a mouse to "target" any difficult word. Similar story-reading programs are described in the Method section of this paper. After 10–14 h of training over 3 months, the children gained significantly more in word recognition and phonological decoding than control children with SRD who received their regular classroom instruction. However, the trained children with lower levels of phoneme awareness gained only about half as much as trained children who began with relatively higher phoneme awareness.

Wise and Olson (1995) hypothesized that intensive training in phonological skills prior to and concurrent with the speech-supported accurate reading on the computer would produce more rapid growth both in phonological skills and in reading. Phonological training in that study included many of the articulatory concepts of the A.D.D. approach (Lindamood & Lindamood, 1975). Wise and Olson also included their own computer programs that trained the manipulation of phonemes/letter-symbols, spelling with speech feedback for errors, and nonword reading (phonological decoding). These programs are nearly identical to those in the current study and are described in the following Method section. Children in the phonological training condition were compared to trained control children, who spent all their individualized computer time reading stories accurately on the computer. Their small-group instruction was based on Palincsar and Brown's (1984) "reciprocal teaching" program, to balance small-group time and to improve their error detection in a very different way from the phonological training. At the end of about 50 half-hour sessions over a semester, children in both conditions made much stronger gains than regular-instruction controls had in previous studies. Children in the phonological condition gained about twice as much as children in the comprehension condition in phoneme awareness and phonological decoding, with more modest advantages in standardized and experimental measures of untimed word recognition.

A comparable experimental study by Torgesen et al. (1997) also included the Lindamoods' (1975) A.D.D. training method. Torgesen et al. trained a younger, at-risk sample for a longer 88-h training period through second grade, with different one-on-one teachers trained for each condition. Nevertheless, at the end of second grade, Torgesen et al.'s pattern of results was quite similar to that found by Wise and Olson. Children in the analytic phonological condition showed significant advantages in phoneme awareness and phonological decoding, with more modest and not quite significant advantages on a standardized measure of word recognition.

The phonological training programs in the Torgesen et al. (1997) and Wise and Olson (1995) studies were multifaceted. Both included training in articulatory awareness, and both programs included exercises in phoneme and letter manipulation. It is difficult to know exactly which components were primarily responsible for the observed advantages over the less explicitly phonological conditions.

Wise, Ring, Sessions, and Olson (1997) conducted a pilot study to try to separate the contributions of training in articulatory awareness and phoneme/letter manipulation for different children. The hypotheses and design partly overlap with those for the present larger and longer training study. One trainer in a single school administered two different computer-based phonological training programs. Second- to fifth-grade children with SRD attended the training during their regularly scheduled remedial reading or language arts classes. Five groups ($N = 24$) worked on articulatory awareness, phoneme/letter manipulation, and reading with speech support on the computer as in Wise and Olson (1995). Four groups ($N = 17$) learned phonological awareness without explicit articulatory work by identifying and manipulating syllables, onsets and rimes (e.g., pl/ant), and phonemes within spoken words (similar to activities in Catts & Vartiainen, 1993; Adams, Foorman, Lundberg, & Beeler, 1998). Small-group and computer instruction time were matched in both conditions. However, children in the sound-manipulation condition spent more time in spelling and reading work than did children in the articulatory condition, because the articulatory concepts, labels, and pictures took more time to learn than did the rhyme and phoneme manipulation. The main hypothesis of the study was that children with the most severe deficits in phoneme awareness would benefit the most from the articulatory awareness training. We also expected that children who had relatively better phoneme awareness might benefit as much or more from exercises that did not include training in articulatory awareness, since they could spend more time with print exercises.

After 60 half-hour training sessions over 5 months, gains from training in both conditions were quite large and were nearly identical. Only tentative support for the above hypothesis was found in interactions between initial level of phoneme awareness, treatment condition, and gains in phonological decoding ($p = .07$) and in the Wide Range Achievement Test (WRAT) test of word recognition ($p = .09$). However, a second standardized measure of word recognition showed no such trend. Wise, Ring, Sessions, and Olson (1997) suggested that the failure to find statistically acceptable evidence for the hypothesized interaction could have been due to the study's small sample size.

The present study aimed to test the relative benefits of training in articulatory awareness and phoneme manipulation with a much larger sample, using five trainers in five different schools. Methods were refined and clarified to strengthen and differentiate the conditions. Two conditions resembled the phonological awareness with and without articulatory awareness conditions described in the previous study. The study reported in this paper included two additional conditions: a condition that trained articulatory awareness and phonics without exercises in phoneme/letter manipulation, and a regular-instruction control condition.

METHOD

Subjects

Teachers in eight Boulder Valley schools recommended second- to fifth-grade students for the study whom they thought were of average intelligence but were among the lowest 10% of their classrooms' readers, based on difficulties in word recognition. Five schools served as training schools. Students in three other schools served as regular-reading-instruction controls. Control subjects were promised eligibility for training in the following year's program; thus, they could serve as regular-instruction controls for posttraining gains, but not for year follow-up analyses. The control schools were similar to the trained schools in schoolwide performance on fourth-grade standardized California Achievement Tests for reading ("Boulder County," 1997).

After parental permission was obtained, 307 second- to fifth-grade subjects were screened with the WRAT for word recognition (Jastak & Wilkinson, 1984). All students included in the study met the definition for SRD used for most of our studies: significant problems in word recognition (lower 10% in their classroom) despite intelligence in the normal range (either verbal or performance IQ at 85 or above), no primary sensory deficits or emotional problems, and English as their first language. Thirty-one regular-instruction controls and 142 students in the training schools were selected as the lowest readers on the WRAT who could also be scheduled in the study (see Table 1 for subject characteristics). Sixty-seven of the subjects were staffed for special education.

The children's degree of reading disability is apparent from the mean "severity of deficit" ratio presented for each condition in Table 1. This ratio is calculated by dividing the subjects' grade equivalent on the Peabody Individualized Achievement Test (PIAT) word recognition test by their expected age–grade equivalent on the national norms. The subjects' reading impairment is most apparent in comparison to the average grade equivalent over national grade equivalent for all children in the Boulder schools: Because the average performance in these schools is well above the national norms, the average ratio in the Boulder schools is 1.2, compared to a mean of .72 for our subjects. Pretest standard scores for the WRAT and PIAT tests in Tables 3 and 4 also indicate the degree of reading and spelling disability for these subjects relative to national norms. As in our previous studies, the PIAT standard scores tend to be higher than those for the WRAT tests of word recognition and spelling.

The subjects' grade-level range of second through fifth grade was selected for both theoretical and practical reasons. It is important to know if there are differential treatment effects for poor readers across the different grade levels. In

TABLE 1
Subject Characteristics

Condition	Combination	Sound manipulation	Articulation only	Control
N	37	42	43	31
Age	9.1	8.7	8.7	8.5
Grade	3.5	3.1	3.1	2.7[a]
IQ	103.7	103.7	103.1	not done
Severity of deficit[b]	.70	.74	.66	.73

[a] Controls significantly different than training sample, $p < .05$, since only one fifth grader participated as a control.

[b] "Severity of deficit" is calculated as the ratio of grade equivalent on the PIAT word recognition test over the national grade equivalent (average Boulder students score above 1.2 on this measure).

addition, it allows us to achieve a larger sample size of children with the most severe reading disabilities in each elementary school. Distributions of subjects in grades 2–5 were 13, 14, 3, and 1 for the control group; 11, 8, 7, and 11 for mouth manipulation; 16, 9, 12, and 5 for sound manipulation; and 17, 12, 8, and 6 for the mouth-no-manipulation group.

Trained students were scheduled in groups of three for 30-min training sessions during their usual remedial reading or language arts time. In this way, daily reading instruction time did not differ for treated or regular-reading instruction controls. Trained groups were pseudorandomly assigned to condition, with training conditions balanced across the five training schools. Eight subjects moved and 12 were dropped for other reasons (e.g., low IQ or scheduling problems). Thus 122 trained subjects completed the training and were included in the analyses. There were 89 male and 64 female students across the four conditions.

Apparatus

All computer-assisted instruction used IBM-compatible Pentium-based computers, with DECtalk speech boards for the Colorado programs (PAL, Non, Spello, and ROSS programs described below and designed in our laboratory). Sound-Blaster boards were used for some A.D.D. programs developed by Lindamood–Bell Learning Processes (1997) and used with the combination and articulation-only conditions. Children in the sound-manipulation condition practiced consonants and vowels in some programs from Lexia Learning Systems (1994).

Design

The study used a pretest, intervention, posttest, and year-later follow-up design, including midtraining testing on five measures to allow for more powerful growth curve analyses. Children in all conditions received equivalent time in instruction in small groups, on the computer, and one on one with the teacher. A planned difference in the design was that children in the articulation-only

TABLE 2
Design of Study

Conditions	Combination ($n = 37$)	Sound manipulation ($n = 42$)	Articulation only ($n = 43$)
Small groups	13.6 h	14.1 h	13.8 h
Phoneme awareness	Articulatory	Phonological: syllables, rimes, phonemes	Articulatory
Manipulation	Sounds, *pictures,* letters, in speech, reading, & spelling	Sounds, *blocks,* letters, in speech, reading, & spelling	*None* (more time associating mouth feelings, sounds, & letters)
Phonics	Consonants & vowels: name/sound, crazy R, open/closed syllable	Same	Same
Computer practice	A.D.D. 9 h	Lexia 2 h	A.D.D. 8 h
	Consonants & vowels	Consonants & vowels	Consonants & vowels
	Marvin	Matching & blending	
	PAL 2.4 h	PAL 3.2 h	0 h
	Non 0.9 h	Non 1.1 h	
	Spello 3.2 h	Spello 4.4 h	
	Vowel circle	Vowel keyword chart	Vowel circle
	Stories 10.3 h	Stories 12.7 h	Stories 15.7 h

condition spent the most time reading accurately in context on the computers. Children in the combination condition spent the least time reading in context, with more time spent practicing the articulatory concepts and doing the manipulation exercises. Time spent on all these activities in each condition is presented in Table 2.

Follow-up tests on several critical measures were given to 113 of the subjects in the experimental training groups in the spring of the following year, about 10 months after the end of training. The control subjects were not included in this follow-up testing because they were promised, and most received, training in different experimental conditions over the following year.

Assessment Measures

Word recognition. Subjects took the reading subtest of the WRAT, Level 1 (Jastak & Wilkinson, 1984), as a screening measure. Children read test words from a graded list on paper, until they missed 10 words in a row. On this and all other tests, the student’s final response was counted as correct or incorrect.

Other word recognition measures included the PIAT (Dunn & Markwardt, 1970). In the present adaptation of this test, words appeared one at a time, untimed, on the computer screen, and students progressed through the list until they missed five of the last seven items. An experimental time-limited word recognition test included a graded list of words displayed one at a time on the screen for 2 s each, but students took as much as time as they wanted to respond.

Phonological decoding. Three tests measured phonological decoding, by asking students to try to read or sound out items as if they were real English words. Students took experimental timed and time-limited tests of nonword reading, as well as the standardized Woodcock Word Attack Test (Woodcock, 1987). For the untimed nonword reading test, items appeared one at a time on the computer screen. Three practice items were followed by 44 test items, arranged in random order of difficulty. Students could take as long as they needed to answer, and, as in all other tests, their final response was scored correct or incorrect. Nonwords were scored correct by phonics rules (e.g., tive rhyming with hive) or by analogy to a real word (e.g., tive rhyming with give). On the time-limited nonword reading test, children were informed that the nonwords would only stay on the screen for 2 s. Items were presented in order of difficulty, and testing stopped after 7 consecutive errors. The two computerized nonword tests just described were designed in our laboratory and include 21 items with no consonant clusters, 50 items with two-consonant/phoneme clusters and, 6 items with three-consonant/phoneme clusters. The Woodcock Word Attack Test included 1 practice item and 50 nonwords of graded difficulty; only 7 items include even two-consonant clusters. Students read until they had missed 5 consecutive items.

Phoneme awareness. The Lindamood Auditory Conceptualization Test (LAC; Lindamood & Lindamood, 1979) and our own phoneme deletion test both measured phonological awareness. We used an “extended” second half of the LAC test, which included 18 items wherein children used colored blocks to indicate where sounds changed in syllables (e.g., “If this says ip, show me pip”). In the tape-recorded phoneme deletion test, devised in our laboratory and based on the Rosner Test of Auditory Analysis Skills (Rosner & Simon, 1971), the student was asked to repeat an item (e.g., “Say pran”) and then to say it again, leaving out a sound (e.g., “Say pran without the /r/”). After 6 practice items, the test presented 40 items. The first 8 items included deletion of the initial or final sound from CVC items. If a child missed the 8 initial items, testing was discontinued.

Nonword repetition. A tape-recorded nonword repetition test asked students to repeat 25 nonsense words which varied in terms of number of syllables and number of consonant clusters within syllables. This test was used as a measure of phonological short-term memory and of articulatory clarity.

Orthographic coding. A computer-administered orthographic choice test was developed for the Colorado twin study. It tested subjects' sensitivity to words' exact orthographic patterns, independent of phonological coding. It included 80 trials that required students to select the word from word–pseudohomophone pairs (e.g., rain rane).

Spelling. Spelling production was measured with the WRAT spelling test, Level 1 (Jastak & Wilkinson, 1984). Items increased in difficulty and testing stopped after 12 consecutive errors. The PIAT spelling test (Dunn & Markwardt, 1970) measured spelling recognition. In our own computerized version, computer speech pronounced a word, and students used the mouse to select the one of four items that correctly matched the spoken word. Testing stopped when the student made five errors in the last seven items.

Reading comprehension. In the PIAT test of reading comprehension (Dunn & Markwardt, 1970), children read sentences that increased in difficulty and then chose one of four pictures that best matched the meaning of the sentence. Testing stopped when the student missed five of the last seven items.

Arithmetic. The WRAT written test of arithmetic, Level 1 (Jastak & Wilkinson, 1984), was included as a test to control for Hawthorne effects, since this study included no math instruction. Testing was stopped after students had answered all the items they could, or after 10 min.

All the above measures were given as pretests and posttests. Four were also given as midtests, to be used in growth curve analyses. These were the tests of time-limited and PIAT untimed word recognition, untimed nonword reading, and phoneme deletion. These four tests were selected as stable measures of the skills we were most interested in, and included measures we have used in most of our past studies.

Two tests were included only for analyses of individual differences. The Colorado version of the Rapid Automatic Naming Test (RAN; Denckla & Rudel, 1976) measured speed of word retrieval for alphanumeric and graphic symbols. Students named as many letters, numbers, pictures, or colors as they could in 30 s, naming items of each kind of stimulus from a sheet of paper. The schools shared the Wechsler Intelligence Scale for Children–Revised (WISC-R; Wechsler, 1974) scores of the 60 of our students for whom they had testing. Testers from the study administered four subtests (Vocabulary and Similarities, and Picture Arrangement and Block Design) to 58 students who had no previous WISC testing. (Controls and 4 trained subjects were not tested, due to problems in scheduling.)

The study also measured children's daily performance in reading. Children took tests at the end of every month, at midterm, and at the end of the year on some of the words they had "targeted" during their reading. The monthly tests included 20 words, and the midtests and final tests included 40 words. The program constructed half as many nonsense words in each test by combining some onset and rhyming portions of different words (e.g., basket and spider would yield spasket and bider).

The Three Training Conditions

The three training conditions involved different combinations of aspects of phonological awareness training programs. Children in the combination condition explicitly learned how articulatory gestures relate to sounds and spelling–sound patterns and explicitly learned how to manipulate sounds in analytic spelling and reading exercises. Children in the sound-manipulation condition also learned explicitly to manipulate sounds, but had no explicit articulatory instruction. Children in the final articulation-only condition learned explicit articulatory awareness concepts, but never explicitly manipulated sounds in isolated exercises with words or nonwords.

The training in this study is rather involved and the design complex. Please refer to Table 2 to help keep track of how conditions do or do not differ. We will first discuss methods common to all three conditions. Next we will discuss activities common only to the two conditions in which children practiced explicitly manipulating sounds in exercises. Next we will describe the activities common only to the two conditions in which children learned explicit articulatory awareness. Finally we will describe elements unique to the sound-manipulation or to the articulation-only conditions.

Instruction Common to All Three Treatment Conditions

Motivation methods. In all conditions, the initial small-group session set up behavioral and academic goals for the program. The teachers explained that children would learn to become their own teachers and that they would discover what they needed to know by being guided with focused questions by their teacher. Children discussed how to help each other learn independently by giving each other hints, by asking good questions, and by respecting each other’s need and right to discover concepts themselves.

Behavioral goals were based on the concept of “participating politely.” Every small-group session began with the setting of behavioral and academic goals and ended with evaluating whether those goals had been met. Goals related to appropriate behavior, remembering previously taught concepts, using strategies, helping each other, and succeeding with the programs. Children earned tokens for meeting goals, and tokens could be traded every 2 weeks for small prizes. Children helped choose the goals, which were adjusted as easier goals were achieved.

Teachers and teaching style. The same teachers taught all three conditions. The teachers all knew articulatory phonetics, which they could use to understand children’s errors and to choose simple to complex patterns in spelling and sound-manipulation exercises away from the computer. Teachers used a “guided discovery” teaching style with all children. With this approach, teachers tried to guide children to discover concepts whenever they could, rather than telling them a concept. When a child made an error, teachers agreed with something that was correct in the child’s response and led the child to correction with a focused question. Teachers were trained during the summer before the study started. They had classes together, role-played and critiqued each others' performances, and cotaught and critiqued each other in a 6-week summer clinic. All this training was directed and supervised by the first author.

Research teachers were instructed to teach all three methods with equal enthusiasm and energy, with the idea that we really did not know yet which type of training would prove most helpful for any child. These teachers learned that the only way to find out if the methods differed would be to be as objective and equivalent as possible in delivering the instruction, at the same time keeping the crucial factors of the conditions (articulation and manipulation) as different as possible. Maintaining consistency among teachers and differences among conditions was supported by biweekly 2½-h training meetings with questions, problem solving, and role-playing, and by at least monthly visits from the first author. All these similarities were by careful design. The purpose of the study was to examine and compare benefits from explicit attention to articulation and to manipulation; not to measure the benefits of linguistically aware teachers or of guided discovery teaching, which could be worthwhile subjects of some other study.

Time on task. Children in all conditions spent equal time in small-group instruction and in individualized time on the computer. In all conditions, teachers spent one third of that computer time monitoring each student, helping the students use their instructed strategy to support the phonological exercises or the story reading. Teachers kept track of monitored time, to keep it as equivalent as possible among children and among conditions.

Structured phonics concepts. While the versions of phonological awareness were different, the actual grapheme–phoneme "phonics" rules that children in each condition learned were the same and were taught in the same order. Children in all conditions were taught sound–symbol relationships for consonant letters and digraphs. Children learned that vowel letters could say their "short sounds" in closed syllables (e.g., hop, bit, man) or their "names," as in hope, bite, and mane. Children learned "bossy e" as the first way of making a vowel say its name, and later learned "two vowels go walking" (e.g., main, bead, soap) and open syllables (e.g., she, so, hero, crazy) as the other two ways to make a vowel say its name. While the form of the charts was different, the mouth and nonmouth conditions all had vowel charts that children used to help them find vowel sounds. These charts got "decorated" with more spelling patterns as they were learned. Other phonics concepts taught in all conditions were "crazy r's," c- and g- spelling regularities, the pronunciations of y, and regularities about when letters would be doubled.

ROSS (Reading with Orthographic and Speech Support). The other program common to all conditions was ROSS reading. Time spent on ROSS itself was planned to be different, to keep overall instructional time the same. The articulation-only condition had the most ROSS story-reading time, and the combination condition had the least, due to all the other programs children in this condition did.

All students spent part of their training time choosing and reading stories of interest to them from 10 directories (graduated from primer to sixth grade). They were trained to “target” any word they found difficult. When a word was targeted, the program highlighted single-syllable words as onsets and rimes (e.g., pl/ant) and multisyllable words in syllables (e.g., plant/ing). After students made an attempt to sound out the word, the program pronounced the segments. At that point if the student requested, the program could pronounce the whole word. Students read silently when reading independently. When students read with teachers (about one third of the time), they read aloud. If they misread a word and did not correct themselves, teachers asked them to try the misread word again. Teacher feedback differed by condition. The program kept track of targeting rates on all days. Teachers showed students comparisons of their targeting on independent and monitored days and encouraged them to learn to target as many words when they read independently as when they read with the teacher. This was done to help students improve their detection of their own errors, thus decreasing the number of times they misread words without correction.

Methods Common to the Two Articulatory Conditions

Small-group instruction was initially identical for both conditions in which children explicitly learned articulatory awareness, which we will call combination and articulation-only conditions for brevity. The first lesson taught how the brain works during reading and why learning to use mouth feelings would help students learn to read and spell better. Children used mirrors and felt their faces and necks with their hands to discover the articulatory movements that produced different sounds and to associate these feelings with sounds, pictures, letters, and labels. They learned to organize consonant sounds by mouth actions, as in the A.D.D. manual (Lindamood & Lindamood, 1975). Lindamood “brother pairs” label sounds which have the same articulatory gestures but different voicing, and the labels and pictures are designed to highlight the gestures needed to make the phonemes. For instance, sibilants /s/ and /z/ are “skinny air” brothers, and bilabial stops /p/ and /b/ are “lip poppers.” Lindamood “cousins” are related less closely, only by “something about how” they are made (i.e., by their manner of production). The pictures and labels are designed to highlight the manner (e.g., nasals /m/, /n/, and /ng/ have a picture of a nose.)

Children then used their improving articulatory awareness to distinguish and compare vowel sounds, organizing these sounds into a “vowel circle” representing mouth shape (smily, open, or round) and tongue position in the mouth. Children’s initial small-group instruction was interspersed with computer exercises with programs under development at Lindamood–Bell Learning Processes. The programs showed animated mouth pictures, and children practiced associating the pictures, sounds, letters, and labels.

When children were 80% successful with the consonant concepts, they learned three vowel sounds, “ee,” “o” (ah), and “oo,” as the most extreme versions of each mouth shape. At this point, children in the combination small groups began manipulating mouth pictures to represent changes in simple two- and three-phoneme syllables. Children in the articulation-only condition learned new games and exercises to strengthen their knowledge of vowel sounds, in order to balance small-group time with the other conditions, but their games included no explicit manipulation of sounds. These activities are described below in the section about activities unique to each condition.

For the combination and the articulation-only conditions, the next two small-group sessions were spent discovering how vowel sounds can be organized into a "vowel circle" based on mouth shape and tongue position, and assigning letter symbols to the sounds and mouth positions. These associations were then practiced with the A.D.D. computer programs. These students used their vowel circle charts for support in all the programs they did. The vowel circle showed the letter symbols they had learned for vowel sounds, arranged according to tongue position and mouth shape, as they had learned in their small group and practiced on the computer. As students learned more vowel patterns, they were added to their vowel circles in the appropriate places.

Methods Common to the Combination and the Sound-Manipulation Conditions

Children in the combination and in the sound-manipulation conditions explicitly learned to manipulate sounds within syllables. However, the phonological awareness underpinnings, the appearance of the screen during the exercises, and the teachers' feedback and support for these activities were different by condition, as will be described later. Children in both conditions did some kind of sound-changing exercises in small-group exercises (e.g., "If this says boot, show me bat"), and both participated in spelling of "fair" words that illustrated the phonics concepts they learned. Children in both conditions also spent about half their individualized computer time on the following Colorado phonological exercises.

The PAL (Phonological Analysis with Letters) program. Letter-symbols in PAL were always pronounced in one way only (e.g., a_e was always as in "cake," "a" by itself was always as in "bat"). Letter-symbols were arranged on the computer screen to match the way children in the conditions learned consonants and vowels. The combination children saw the consonants arranged as brothers and cousins, described above. The stop-consonant brothers were in the first column (e.g., p b, t d, k g, ch j), the continuant brothers in the next column (e.g., f v, th *th*, s z, sh zh), and the cousins in the third column (e.g., m n ng, w wh h y, l r). Vowel letter-symbols were arranged as in the vowel circle described below in the section on the articulatory conditions. Children in the combination condition were encouraged to use vowel circles, to watch their mouths in mirrors, and to use all the articulatory concepts they had learned in groups for support in doing the PAL program.

The sound-manipulation children had consonants grouped as in the phrases of the alphabet song, with blanks where vowel letters would have been. Vowel letter-sound symbols appeared on the right of the screen arranged as Short Sound, Names, Others, and “Crazy-r’s,” as on their own “key-word” vowel chart described below. Children were encouraged to use their vowel charts to help them figure out sounds, but they never used mirrors.

In both conditions, the PAL program asked children to build a simple real or nonsense syllable and then to change one sound to match the change pronounced by the program. For example, the program might say, “Show me ’ook’.” The child used the mouse to click on letter symbols, which the program pronounced as the child selected them. The child arranged the letter-sound symbols in order and kept comparing the computer’s pronunciations of the attempt and the goal item until they matched. If the child failed three times to match the computer’s pronunciation, the program suggested he or she ask the teacher for help. After the child successfully constructed the syllable, the program might say, “If that says ’ook,’ show me ’koo’.” The child changed the syllable by adding, deleting, substituting, switching, or repeating a letter-sound symbol to match the change in the syllable pronounced by the program. Each PAL set consisted of six items. Item sets were arranged in levels of difficulty from two simple sounds to complex single syllables of five and six phonemes. The program advanced to higher levels of difficulty when children completed two stimulus sets with 90% accuracy; it decreased in difficulty if they scored below 80% on two consecutive sets. Children completed one stimulus set or exited the program after 5 min if the set was not completed.

Nonword reading program (Non). Children in both manipulation conditions used a nonsense word reading program designed in our laboratory. In the Non program, children chose one of four nonwords to match one pronounced by the program. The nonwords contained legal English orthographic patterns, as opposed to the Lindamood letter-sound symbols, so the screen appeared the same for both manipulation conditions. The program pronounced the nonwords as the children chose them. Children earned points for choosing the correct nonword, with fewer points on each succeeding attempt. Much as in PAL, the Non program automatically advanced and retreated in difficulty from CV to multisyllable levels, depending on the child’s performance. Children completed one file or exited after 4 min on Non.

Spello. After children in the manipulation conditions were 80% successful with the CVC level of PAL, they alternated between the Non or Spello program. This program encouraged children to manipulate letters and sounds in exploring real English spelling patterns in real words (Wise & Olson, 1992). Spello also advanced and retreated in difficulty from CVC to multisyllable words, according to the child’s performance. Sets included six items, and students spent up to 7 min or the time it took to complete one stimulus set on this program.

In Spello, the program pronounced a word and the child used the keyboard to spell it. The child could get phonological support at any time, by comparing the pronunciation of the attempt with the pronunciation of the spelling word. When the pronunciations of the attempt and of the spelling word matched, the child chose the Done button. The program then provided orthographic feedback by showing which letters were in the goal word and which were in the correct place. For instance, if the child had spelled “gane” for the goal word of “gain,” the program would indicate that *g*, *a*, and *n* were in the correct word, and *g* and *a* were in the right places. If the child did not spell the word correctly after three uses of the Done button, the program showed the correct spelling, and the child typed it in correctly. Thus, every word was spelled correctly at least once.

After children in the manipulation conditions were 80% successful at the CVC level of PAL, they began spending about half their time reading stories on ROSS, as described above, and half on the phonological exercises. Teacher support for misread words differed for each condition and will be described at the end of the Method section.

Methods Unique to the Combination Condition

After combination students learned the most extreme vowel sound for each mouth shape (e.g., “ee,” “ah,” and “oo”), they began practicing sound manipulation in small groups using mouth pictures to represent the sounds. For example, the trainer said, “Show me op,” and the child showed a picture of an open vowel followed by a lip popper picture (representing a bilabial plosive). Next the teacher said, “If this is op, show me pop,” and the child added a lip popper picture in front of the pictures already in place. After the child had learned the vowel circle, the same vowel pictures were used to represent any vowel in that family (e.g., ee, I, e, a, u, and a_e are all some form of fronted “smily” vowel). Two half lessons were devoted to manipulating mouth pictures for sound-changing exercises as described above, but using all vowel sounds. Then this condition began using the Colorado phonological programs described above. Combination children were encouraged to use their articulatory concepts, their vowel circles, and even their mirrors to watch their mouths to help them figure out the answers and correct their errors in the phonological programs. Teachers questioned these children about their responses in ways to encourage error checking using their mouth feelings. For instance, a child might have spelled “pate” when trying to spell “plate.” If the child were a combination student, the teacher would ask, “When you say ‘plate,’ what do you feel right after the popper?” Students were encouraged to use their mirrors to watch their mouth pronounce the word they were trying to spell, to help them notice the deleted sound.

The Marvin program. The combination condition was the only one to use the Lindamood–Bell program “Marvin,” since it involved both articulation and manipulation. Marvin appeared on the screen as a green blob with an animated mouth who read a nonsense word printed on the screen. Children decided whether Marvin’s pronunciation had “matched” by reading a nonword correctly, or if not, whether he had added, deleted, switched, or substituted a sound. This was the children’s favorite of the Lindamood–Bell programs; it was motivating and involved the children in thinking analytically about sounds. Students used Marvin at first a few times a week for a few minutes a day, and then about once a week to keep the articulatory concepts strong.

Teacher support in ROSS reading. When combination students failed to correct themselves or target a misread word by the end of a sentence, teachers questioned them about their mouth feeling to help them learn to correct their own errors. Teachers would cover the missed word and ask the student about what s/he had said at the point of contrast with the correct word (e.g., misreading "house" when the print said "home": "When you say house, what do you feel at the end?"). When students were still quite new and analytic with these skills, they used their mirrors to help them be aware of the "skinny" sound at the end of their pronunciation. Then the teacher would uncover the word, and the child would compare the /s/ that he or she said to the "nose"/m/ sound represented at the end of "home." Soon, as students improved, they no longer needed their mirrors to help them answer articulatory-based questions.

Methods Unique to the Articulation-Only Condition

At the same time that students in the combination condition started doing manipulation exercises, students in the articulation-only condition played games designed to strengthen their articulatory concepts. This was done to keep small-group instruction time equivalent yet avoid explicit instruction and practice in manipulation of sounds in small groups. The games included memory and matching games of pictures, labels, letters, and feelings. They included a "lip-reading" game where students silently demonstrated mouth positions to reflect tongue position and mouth shape for *isolated* vowel sounds. They tried to do this so clearly that all the other students in the group could guess what the vowel sound would be, and the group got to put a block onto an ever-growing tower. Small-group instruction continued to teach the same phonics concepts covered in the other two conditions, but children in this condition never practiced them in analytic spelling, segmentation, or sound-manipulation exercises.

When articulation-only children were 80% successful with placing vowel letter-symbols on the vowel circle, they began spending nearly all or all of their individualized computer time on ROSS and some on occasional review of consonant or vowel practice. Thus a planned difference in treatment was that the articulation-only children would spend the most time reading with ROSS.

Teacher support in ROSS. Teachers asked articulation-only students to say the vowel in each targeted word segment, using their vowel circle if needed. Then they learned to "use the colors" to help them sound out the word. It is important to note that children in this condition did practice segmenting and blending of words in stories but did not have explicit analytic sound-manipulation exercises outside of their story reading.

Methods Unique to the Sound-Manipulation Condition

Students in this condition learned that training in phoneme awareness meant learning to notice, listen to, manipulate, and play with sounds inside syllables. They also learned that improving their phoneme awareness would help their brain learn to read and spell better. The students learned to pay attention to and manipulate first syllables, rhymes, initial and final phonemes, and finally phonemes in all positions in single-syllable and multisyllable words. For example, they played games first counting, then deleting, and finally swapping syllables using colored squares to represent them. Next they learned to recognize rhymes and played rhyming games such as "Zip, zap, rhyme," where rhyming words are "zipped" from one student to another around a circle. Students then manipulated onsets and rimes in games like "*B*alking *T*ackwards," using colored strips and squares to help them keep track of the segments.

Next students learned to use colored squares or blocks to represent initial and final sounds in three-sound words and manipulated sounds at the beginnings or ends of words. Eventually they turned a square sideways to represent the vowel. Soon they could use squares to represent each phoneme in single-syllable words with up to six phonemes. They described phoneme manipulations including addition, deletion, switching, repetition, or substitution, as done in other phonological awareness programs (e.g., Ball & Blachman, 1991; Elkonin, 1973; Lindamood & Lindamood, 1975).

The sound-manipulation condition used a key word for vowel sounds and letters. The "names" of the letters were themselves key words for five vowels: a_e, ee, i_e, o_e, and u_e. "Short Sounds" had key words of apple, Ed, It, octopus, and up. "Other" sounds (neither "short sound" nor "name") used key words of oil, owl, awesome, ooooo (with a ghost saying that), and hook. Key words were chosen to minimize influence of the succeeding consonant (e.g., "Ed" raises /e/ less than does "egg"). They also were chosen so that unobtrusive picture cues could be drawn lightly right on the letter, to help children transfer the cued sounds into reading. For example, *a* had a green stem and two leaves on top to remind the child of "apple"; *u* had little arrows on the ascenders, to remind the child of "up."

Students in the sound-manipulation condition initially practiced consonant and vowel sounds for about an hour's total practice on rhyming and vowel-matching games from the Lexia Learning Systems program. As soon as children had been introduced to the "name" and "short sound" vowel concepts and their keyword pictures, they began practicing the Colorado phonological programs. Children progressed through these programs using the same criteria as the combination children. Lexia games were occasionally offered as motivators, as Marvin was for the articulatory condition. In all the programs, teachers questioned sound-manipulation students about what they *heard* and *noticed* about the sounds in a word, rather than about what they *felt* or *saw* in a mirror, as a teacher would ask a combination student. A sound-manipulation student who had read "fly" as "fry" would be asked, "When you say 'fry,' what do you hear after the /f/?" Students were encouraged to "stretch out the sounds" and to listen for all the sounds in order. Sound students never used mirrors, even though some other nonspecifically articulatory programs have used them (Lundberg, et al., 1988), because we wanted to avoid explicit attention to articulation in this program.

Students in all conditions spent about 14 h in small-group instruction and about 24 h working individually at the computer (see Table 2 for how that time was divided). Another 1½ h was spent near the end of training practicing reading in books, for a total training time of about 40 h. Another approximately 10 h of time was given to setting goals, giving and trading tokens, reviewing, and record keeping. We never measured the time spent on these goal-setting and administrative activities in our previous studies. Pretesting began in October, and training began in late October and continued to the start of May, when posttesting began. Thus, training spanned about 6 months.

RESULTS

General Points about Data Analyses

The data analysis for this study proceeded in three stages. These stages were: (1) to examine general gains from treatment, compared to the regular-instruction controls; (2) to compare the three types of phonological treatments to each other; and (3) to examine individual differences in response to treatments. For all analyses, a severity of deficit score (reading grade level/expected grade level) was used as a covariate because the four conditions differed (nonsignificantly) in mean severity estimates before training began (see Table 1). In each step the data were first analyzed by MANCOVA to test overall condition effects on all criterion variables. The multivariate tests were then followed by univariate analyses. A priori orthogonal contrasts were coded to test specific hypothesized condition differences in the first two series of analyses.

In the first stage of analyses, gain scores from pretest to posttest for the three treatment conditions were compared to the scores of regular-instruction controls on those measures administered to the controls. Recall that trained subjects' instructional time was taken from their regularly scheduled reading or language arts time, so that reading instructional time would be the same in all conditions. Of course, trained students had more personal attention in small groups and more individualization on the computers than would have been possible for most students in the regular-instruction control condition.

The second stage of analyses examined differential treatment effects. All analyses included school as a variable, to remove extraneous variance due to significant effects of between-school differences. Intervention conditions were coded for two planned contrasts. The first contrast examined differences between the manipulation versus articulation-only treatment conditions, to see what explicit analytic exercises in phoneme manipulation in isolated reading and spelling added to the instruction in articulatory awareness, phonics, and speech-supported reading practice of the articulation condition. The second contrast examined differences between the combination and the sound-manipulation conditions, to see what explicit instruction in articulatory awareness would add beyond well-structured training in phonological awareness and phonics. Gain scores were used as the dependent variables in analyses of intervention effects for all measures with data collected only pretest and posttest. Growth curve analyses were used for measures with additional data collected in the middle of the intervention. For these analyses, a linear growth model was first individually fit to each subject's data over the three testing points. The resulting estimates of the linear slope or growth over the intervention for each subject were then used as the dependent variables in subsequent between-subjects group analyses (Willett, 1989). Gain scores from pretest to follow-up test were used in all analyses of long-term treatment effects at follow-up testing.

The final stage of analyses examined individual differences in treatment effects and interactions with treatment condition. The MANCOVA test results are from a sample of 111 subjects as a result of missing data from 11 subjects on the Woodcock Word Attack. Subsequent univariate tests were conducted with all 122 intervention subjects. The covariates of interest included grade, full-scale IQ (WISC-R; Wechsler, 1974), initial levels on R.A.N. (Denckla & Rudel, 1976), and initial levels of phoneme awareness. The R.A.N. was defined as a composite of all four subtests. Initial level of phoneme awareness was defined as the average of *z*-scored pretest levels on the phoneme deletion and LAC tests. Grade was treated as a continuous variable. Age was included in all interaction analyses except those examining grade effects. Pretest was also included in the univariate interaction analyses to remove any autoregressive effects on all measures; that is, so that gain score differences were not influenced by a student's pretest level. The computed intercept was used as the autoregressor in any interaction analyses on growth curves. An alpha level of .05 was used to test for significance in all analyses.

1. Analyses of Treatment versus Control Conditions

Students in all conditions made highly significant gains relative to the regular-instruction control condition, on all reading-related tests (see Table 3). A significant overall 6 (Test) by 4 (Condition) MANCOVA on gain scores indicated reliable differences among the mean vectors for Condition (Wilks's $\Lambda = .48$, approximate $F(21, 405.4) = 5.56$, $p < .0001$). This test indicated a main effect of condition, and multivariate analyses showed that contrasts between control and treatment conditions were also significant (Wilks's $\Lambda = .61$, $F(7, 141) = 12.7$, $p < .0001$).

Univariate analyses next tested the overall effect of condition on each test. Table 3 presents means, standard deviations, significance levels, and standard effects for all measures. (Standard effect sizes were calculated in terms of the differences between treatment mean gain scores compared to the standard deviation of the control condition's gains.) Effects were significant for all measures of word recognition, for gains in PIAT standard scores ($F(3, 148) = 4.0$, $p < .01$, $r^2 = .08$), WRAT standard scores ($F(3, 148) = 6.6$, $p < .001$, $r^2 = .12$), and time-limited word recognition percentage correct ($F(3, 148) = 3.8$, $p < .02$, $r^2 = .07$). Standard score changes were used in the analyses because a change in standard score means more than just growth in grade equivalence. Positive standard score growth indicates the student has advanced more than the peer group has during the training period, measured against changes in national

TABLE 3
General Treatment versus Control Effects: Pretest and Posttest Scores with Standard Effect Sizes

Measure	Control (n = 31) Pre	Control (n = 31) Post	Treatment (n = 122) Pre	Treatment (n = 122) Post	Effect size
PIAT SS	84.3 (6.8)	87.9 (7.3)	83.2 (6.4)	91.6 (7.6)	.73**
WRAT SS	76.0 (7.0)	79.4 (8.6)	73.6 (7.7)	83.5 (9.7)	.94***
Time-limited word recognition RS	17.3 (12.7)	28.5 (14.5)	20.7 (15.1)	37.6 (19.3)	.98**
Nonword decoding %C	21.2 (20.4)	29.1 (16.1)	23.8 (17.6)	51.3 (20.1)	1.46***
Phoneme deletion %C	30.8 (23.5)	33.9 (19.7)	32.9 (20.9)	50.7 (23.8)	.92***
LAC, part II RS	4.6 (2.4)	5.1 (3.0)	4.7 (2.8)	9.0 (3.6)	1.73***
WRAT math GE[a]	2.9 (.6)	3.3 (.6)	3.2 (.6)	3.8 (.9)	.40 ns

Note. Standard deviations in parentheses. SS = standard score; RS = raw score; %C = percent correct; GE = grade equivalent. Effect sizes reflect mean differences in gain scores standardized against standard deviation of the control sample.

[a] Sample size = 30 for math test control.

** $p < .001$.

*** $p < .0001$.

standardized score units. Significance levels and effect sizes were even higher for main effects on phonological skills: phonological decoding as nonword reading gain scores ($F(3, 148) = 13.1$, $p < .0001$, $r^2 = .21$), and gains on phoneme awareness measured by phoneme deletion ($F(3, 148) = 9.8$, $p < .0001$, $r^2 = .17$) and the LAC ($F(3, 148) = 19.5$, $p < .0001$, $r^2 = .28$).

The specific contrasts between trained conditions versus regular-instruction controls were significant for all reading and language measures. The only test where this difference was not significant was on analyses of grade equivalent gains on the WRAT math test. This test was included as a control, since no math instruction was part of the treatment program. However, a trend did favor the trained children ($F(3, 147) = 3.24$, $p = .07$, $r^2 = .02$)

2. *Analyses of Differences among Treatment Conditions*

A significant overall 13 (Test) by 3 (Treatment) MANCOVA indicated that treatments differed on outcome measures (Wilks's $\Lambda = .54$, $F(26, 182) = 2.5$, $p < .001$). Subsequent univariate analyses indicated that the treatment effects differed significantly only on phoneme deletion growth curves ($F(2, 114) = 3.07$, $p = .05$, $\eta^2 = .05$), LAC test gain scores ($F(2, 114) = 9.43$, $p < .0001$, $\eta^2 = .14$), and orthographic coding ($F(2, 114) = 4.2$, $p < .05$, $\eta^2 = .07$). No other univariate tests were reliable on any other pretest or posttest measures. Performance also did not differ by condition on any of the ROSS reading measures. Table 4 presents means and standard deviations for all treatment conditions for all measures used in pretests and posttests.

A multivariate analysis of contrasts was conducted to make comparisons between conditions. The contrast of manipulation versus articulation-only was significant (Wilks's $\Lambda = .64$, $F(13, 91) = 3.85$, $p < .0001$). Planned orthogonal contrasts indicated that the two manipulation conditions versus the articulation-only condition differed on phoneme deletion ($F(1, 114) = 4.32$, $p < .05$, $r^2 = .036$), the LAC test ($F(1, 114) = 18.5$, $p < .0001$, $r^2 = .14$), and orthographic coding ($F(1, 114) = 7.72$, $p < .01$, $r^2 = .06$). Significance levels and effect sizes in Table 4 relate to this contrast. The manipulation conditions showed advantages on measures of manipulation of sounds in syllable, consistent with their training. The articulation-only condition, with more hours of accurate reading in context, showed an advantage on orthographic coding. The combination versus sound-manipulation contrast was not reliable (Wilks's $\Lambda = .84$, $F(13, 91) = 1.29$, $p = .23$). No univariate test indicated any reliable difference between the two manipulation conditions.

3. *Individual Difference by Treatment Analyses*

Analyses examined whether effects of treatment might differ according to a student's initial profile. Theory and results from previous studies led us to suspect that effects might vary according to children's initial levels of intelligence (measured by IQ) and rapid naming ability. We had special interests in how effects might vary according to a child's grade and initial level of phoneme awareness.

Full-scale IQ. IQ had a significant main effect on outcome measures in the MANCOVA (Wilks's $\Lambda = .76$, $F(13, 82) = 2.01$, $p < .05$). Univariate tests revealed that effects of IQ were reliable on gains in word recognition as measured by the PIAT, WRAT, and time-limited test ($F(1, 104) = 4.1$, $p < .05$, $r^2 = .037$; $F(1, 104) = 9.22$, $p < .01$, $r^2 = .08$; and $F(1, 104) = 5.98$, $p < .02$, $r^2 = .05$ respectively). IQ also had a significant effect on phonological coding as measured by untimed nonword reading ($F(1, 104) = 5.19$, $p < .025$, $r^2 = .048$) and nonword repetition ($F(1, 104) = 9.6$, $p < .01$, $r^2 = .08$), and on gains in PIAT comprehension standard score ($F(1, 104) = 25.1$, $p < .001$, $r^2 = .19$). Interestingly, IQ had no effect on gains in phoneme awareness measures. The regression weights indicated that full-scale IQ was positively correlated with growth on all other measures. The multivariate analysis of covariate by treatment interactions was not significant, indicating effects of IQ did not vary across conditions.

Rapid automatic naming. The RAN did not have a reliable effect in multivariate analyses (Wilks's $\Lambda = .84$, $F(13, 87) = 1.22$, $p = .28$). Even though the multivariate test was not significant, univariate analyses were conducted on the two time-limited measures, since they share speed processing requirements similar to those of the RAN. The RAN had a reliable effect on gains in time-limited word recognition ($F(1, 109) = 7.67$, $p < .01$, $r^2 = .065$), and it showed a trend on time-limited nonword decoding ($F(1, 109) = 3.11$, $p = .08$, $r^2 = .027$). The regression weights suggest that RAN performance at pretest correlated positively with time-limited outcome variables. However, these anal-

TABLE 4
Treatment Effects: Pretest and Posttest

	Combination ($n = 37$)		Sound manipulation ($n = 42$)		Articulation only ($n = 43$)	
Condition	Pre	Post	Pre	Post	Pre	Post
Phoneme awareness						
LACii RS	4.9 (2.8)	10.1 (3.4)	4.7 (3.0)	10.0 (3.5)	4.5 (2.7)	7.0 (3.2)
Phoneme deletion %C	35.3 (23.4)	53.6 (25.7)	32.4 (22.0)	54.4 (22.2)	31.4 (17.4)	44.7 (22.8)
Phonological decoding						
Timed nonword reading %C	15.7 (14.4)	40.9 (23.5)	17.4 (17.6)	40.9 (19.6)	10.8 (12.1)	31.5 (19.6)
Untimed nonword reading %C	26.9 (17.4)	53.7 (21.5)	26.2 (19.1)	54.7 (18.9)	18.8 (15.5)	45.9 (20.4)
Woodcock SS	81.1 (9.0)	93.8 (10.8)	83.5 (8.3)	95.9 (7.6)	80.9 (10.1)	91.4 (10.4)
Word recognition: untimed						
PIAT SS	83.3 (7.3)	90.8 (8.0)	84.2 (5.7)	93.2 (6.9)	82.0 (6.0)	90.8 (7.8)
WRAT SS	73.6 (7.6)	82.4 (10.2)	75.4 (6.0)	85.0 (8.7)	71.7 (8.9)	82.9 (10.3)
Time-limited RS	22.5 (15.5)	39.0 (20.9)	21.4 (15.5)	37.7 (18.6)	18.6 (14.4)	36.3 (18.7)
Nonword repetition %C	57.9 (15.4)	71.7 (12.5)	54.2 (16.5)	69.0 (13.1)	53.8 (13.7)	67.0 (15.1)
Spelling						
PIAT SS	81.5 (6.9)	87.3 (7.7)	86.4 (6.7)	88.7 (7.0)	84.7 (7.0)	88.2 (7.1)
WRAT SS	74.6 (8.7)	82.0 (9.4)	77.8 (7.0)	84.7 (7.6)	74.8 (9.8)	80.1 (9.4)
Orthographic coding %C	59.2 (13.3)	65.0 (13.4)	60.6 (13.3)	64.6 (12.4)	56.4 (12.1)	66.2 (12.1)
Reading Comprehension						
PIAT SS	89.8 (7.6)	95.3 (9.8)	94.0 (7.2)	95.8 (7.5)	92.5 (6.3)	95.2 (7.4)
WRAT math GE	3.3 (0.8)	3.9 (0.8)	3.3 (0.7)	3.8 (1.0)	3.2 (0.7)	3.8 (0.9)

Note. Standard deviations in parentheses. SS = standard score; RS = raw score; %C = percent correct; GE = grade equivalent.

yses should be interpreted with some caution given that the multivariate analysis was not reliable. The multivariate analysis of covariate interactions with treatment condition was not significant (Wilks's $\Lambda = .86$, $F(26, 174) = .71$, $p = .65$). Neither univariate analysis of time-limited measures indicated covariate by treatment interactions.

Grade. A significant MANCOVA indicated reliable main effects of grade (Wilks's $\Lambda = .56$, $F(13, 88) = 5.31$, $p < .0001$). Subsequent univariate analyses, which included an autoregressor, found that grade significantly affected PIAT standard score and time-limited word recognition growth curves ($F(1, 110) = 15.37$, $p < .001$, $r^2 = .12$ and $F(1, 110) = 16.42$, $p < .001$, $r^2 = .13$, respectively) and WRAT standard score gains ($F(1, 110) = 10.08$, $p < .01$, $r^2 = .08$). Grade also had a reliable effect on Woodcock Word Attack standard score gains ($F(1, 103) = 10.41$, $p < .01$, $r^2 = .09$) and showed a strong trend on nonword decoding growth curves ($F(1, 110) = 3.53$, $p = .06$, $r^2 = .03$). The youngest children tended to gain the most from the training, and gains on all measures decreased as grade increased. A multivariate test of grade with treatment condition was significant (Wilks's $\Lambda = .64$, approximate $F(26, 176) = 1.7$, $p < .05$). Univariate analyses that included the autoregressor found small but reliable interactions on the PIAT standard score growth curves ($F(1, 110) = 3.37$, $p < .05$), nonword2 decoding growth curves and nonword1 decoding gain scores ($F(1, 110) = 3.7$, $p < .05$ and $F(1, 110) = 5.5$, $p < .01$, respectively), and nonword repetition gain scores ($F(1, 110) = 3.5$, $p < .05$). Regression weights indicated that the gains in the combination condition on these variables were less affected by grade/reading level than gains in the other conditions. However, in all other variables gains were the greatest for younger/poorer readers in all conditions. In our previous studies, phonological training also benefited younger students the most (Wise, Ring, & Olson, 1999).

Phoneme awareness. A significant MANCOVA indicated main effects of initial levels of phoneme awareness on criterion variables (Wilks's $\Lambda = .79$, $F(11, 89) = 2.12$, $p = .026$). Univariate analyses showed reliable effects of phoneme awareness on WRAT standard score gains ($F(1, 109) = 15.03$, $p < .001$, $r^2 = .12$), time-limited word recognition growth curves ($F(1, 109) = 10.69$, $p < .01$, $r^2 = .089$), untimed nonword decoding growth curves ($F(1, 109) = 15.38$, $p < .001$, $r^2 = .12$), time-limited nonword decoding gain scores ($F(1, 109) = 13.77$, $p < .001$, $r^2 = .11$), WRAT spelling standard score gains ($F(1, 109) = 11.22$, $p < .01$, $r^2 = .09$), nonword repetition ($F(1, 109) = 3.9$, $p = .05$, $r^2 = .03$), and PIAT comprehension standard score gains ($F(1, 109) = 9.53$, $p < .01$, $r^2 = .08$). Standardized regression weights indicated positive correlations between initial levels of phoneme awareness and gains on all the above measures, showing that children with lower initial phoneme awareness tended to achieve less than did children who began with relatively higher skills. However, growth in phonological awareness itself, in phoneme deletion and LAC tests, did not show any effect of initial levels of phoneme awareness. Multivariate analysis of initial phoneme awareness covariate by treatment interaction was not reliable (Wilks's $\Lambda = .75$, $F(22, 178) = 1.24$, $p = .22$). This shows effects of children's initial levels of phoneme awareness on outcome measures did not vary according to treatment condition, as we had originally hypothesized.

Combined and independent effects of IQ, grade, and phoneme awareness. The combined and independent variance associated with these three pretest variables was analyzed by multiple regression for standard score gains on the PIAT and WRAT tests of word recognition, after controlling for subjects' pretest scores. For the PIAT word recognition linear slope across pretests, midtests, and posttests, the adjusted R^2 was .42, and η^2 was .235 ($p < .05$) for grade, .014 ($p > .05$) for IQ, and .084 for phoneme awareness. For pretest to posttest gains on the WRAT, the adjusted R^2 was .277, and η^2 was .174 ($p < .05$) for grade, .057 ($p < .05$) for IQ, and .053 ($p < .05$) for phoneme awareness. Thus, there were consistently strong and independent effects for grade and weaker but significant independent effects of pretest phoneme awareness on gains in these two measures of word recognition. The influence of phoneme awareness is probably stronger than indicated by this analysis because the subjects' pretest scores on word recognition, which are highly correlated with phoneme awareness, were controlled in the analysis. The independent effect of IQ on gains in word recognition was significant for the WRAT but not the PIAT for unknown reasons. As noted in the previous description of the univariate effects of IQ on gains, the strongest influence of IQ was noted for standard score gains on the PIAT measure of reading comprehension ($r^2 = .19$).

4. *One-Year Follow-Up Results*

Tests administered 10 months after the end of training to 113 of the original 122 trained subjects included PIAT word recognition, WRAT word recognition, time-limited word recognition, untimed nonword decoding, phoneme deletion, and the LAC measure of phoneme awareness. The means and standard deviations at pretest, end of training, and follow-up are presented in Table 5. A 6 (Test) by 3 (Treatment) MANCOVA indicated that the overall differential treatment effect on gain scores was not quite significant (Wilks's $\Lambda = .83$, $F(12, 198) = 1.59$, $p = .098$). However, the MANCOVA trend suggested that univariate analyses on each variable might be informative. The only significant univariate contrast was for the LAC test of phoneme awareness ($F(2, 111) = 3.45$, $p = .035$), suggesting stronger gains at follow-up for the two training conditions that included manipulation. However, there were no trends suggesting that this advantage transferred to any of the reading measures.

A second main and very positive result from the follow-up testing indicated that gains in standard scores on the PIAT and WRAT tests of word recognition at the end of training were largely maintained 10 months later. The nature of standard scores (relative position of scores to grade-level peers) means that trained students continued to gain as much as students at the higher standard-score rank they had achieved by the end of training. The standard score gains at

TABLE 5
Pretest, Posttest, and Follow-Up Means by Condition

Measure	Test	Combination (*n* = 35)	Sound manipulate (*n* = 38)	Articulation only (*n* = 40)
LACii RS[a]	Pre	4.8 (2.7)	5.0 (3.0)	4.7 (2.7)
	Post	9.8 (3.3)	10.3 (3.3)	7.2 (3.1)
	Follow-up	10.1 (3.2)	9.6 (2.7)	8.4 (2.8)
Phoneme deletion %C	Pre	34.2 (23.5)	33.4 (20.7)	32.3 (17.6)
	Post	52.6 (26.0)	56.0 (20.4)	46.1 (22.5)
	Follow-up	57.3 (25.6)	60.3 (20.3)	53.6 (23.0)
Untimed nonword reading %C	Pre	25.1 (16.0)	27.5 (18.6)	19.4 (15.4)
	Post	52.5 (21.5)	56.7 (17.5)	48.7 (18.1)
	Follow-up	55.8 (23.0)	59.3 (17.1)	55.9 (21.2)
Word recognition				
PIAT SS	Pre	83.5 (7.5)	84.5 (5.8)	81.8 (5.7)
	Post	91.1 (8.1)	93.5 (6.9)	91.0 (7.9)
	Follow-up	89.8 (9.0)	90.6 (7.4)	90.7 (8.6)
WRAT SS	Pre	73.3 (7.7)	75.9 (5.2)	72.2 (8.6)
	Post	82.1 (10.5)	85.6 (8.6)	83.6 (10.1)
	Follow-up	82.3 (11.8)	85.2 (7.8)	84.9 (10.2)
Time-limited RS	Pre	21.9 (15.7)	22.5 (15.4)	19.3 (14.5)
	Post	38.5 (21.4)	39.5 (18.2)	37.5 (18.6)
	Follow-up	50.0 (23.2)	48.9 (17.9)	48.5 (18.3)

Note. Standard deviations in parentheses. SS = standard score; RS = raw score; %C = percent correct; GE = grade equivalent.

[a] Articulation-only condition *n* = 39.

follow-up are still substantially larger than those for the control group at the end of the experimental training period. Follow-up data were not available for the control group because the schools required their inclusion in new experimental training conditions over the following year. Nevertheless, the PIAT and WRAT standard-score gains from pretest to follow-up for the three trained groups indicated that most of the initial training effects on word recognition were maintained over the following year.

DISCUSSION

One important finding of this study is that children in all three phonological training conditions made impressive and lasting gains. At the end of training, they did very much better on all tests of reading and all tests of phonological skills than did the regular-instruction controls who received the same amount of reading instruction time in the regular classroom. The accurate phonologically supported reading practice combined with phonological awareness training led to gains in phoneme awareness, phonological decoding, and untimed word recognition as large as or larger than in our previous studies with phonological training

(Olson et al., 1997; Wise & Olson, 1995; Wise, Ring, & Olson, 1997). Children enjoyed the training, and most seemed to use it well at least while working in the small groups or at the computer. Of course, it would be difficult for a classroom teacher to provide the level of intensive support in the classroom that one can provide in a group of four children with the support of one talking computer per child. But adding talking computers to classrooms would greatly extend the amount of individualized support these children could receive in a classroom setting.

Another important result was that children in the three trained conditions performed remarkably similarly to each other after about 50 h of training. There were very few significant differences among trained conditions in this study. All three types of phonological awareness and phonics instruction with computer-speech-supported reading were highly, and nearly equivalently, effective on reading measures in posttests and tests of words studied in the computer stories.

The current study found no significant differences at all in main effects between the combination and sound-manipulation conditions. Because this is similar to the Wise, Ring, Sessions, and Olson (1997) pattern of results, the lack of main effect did not surprise us. We will consider whether our expectation about interactions were met when we later consider the individual difference results.

In contrast, we were quite surprised that the differences between the manipulation versus articulation-only conditions were so few. Children who explicitly manipulated sounds did gain consistently and significantly more in both measures of phonological awareness than children who spent no time manipulating sounds. On one measure of phonological awareness, they showed a continued advantage 10 months after training had ended. However, their gains in phonological awareness did not transfer to similarly better performance on reading measures over the no-manipulation children either at the end of training or 10 months after training had ended.

There was one other difference related to the manipulation contrast. The articulation-only condition, with the greatest time reading stories accurately, performed better on the orthographic coding task. Recognizing specific spelling patterns of words does depend to a large extent on reading experience (Olson, Forsberg, & Wise, 1994; Stanovich, West, & Cunningham, 1991), so the result of greater gains on this test is compatible with the children's extra reading practice. However, this condition did not show an advantage on time-limited word recognition nor on PIAT spelling recognition, as conditions with extra reading practice have done in our previous studies. Thus the result may have been due to chance.

We had expected that the articulation-only condition would gain less not just in phonological awareness, which did happen, but also in phonological decoding (nonword reading) and perhaps in untimed word recognition, which did not happen. We expected this partly because of a study by Uhry and Shephard (1993). They had found impressive gains in phonological decoding, word recognition, and spelling for children who had manipulated sounds in spelling-type exercises compared to children who had studied the same words but without sound manipulation. Also, our previous studies had found that children who had made significantly greater gains in phoneme awareness had also shown similar advantages in phonological decoding. We were therefore surprised especially at the growth in phonological decoding as measured by nonsense word reading for the articulation-only children. They performed equally as well as the children in the manipulation conditions on phonological decoding, and their scores were as high as in phonologically trained conditions in previous studies.

Let us consider why children in our articulation-only condition made such impressive gains in phonological decoding (nonword reading) relative to their somewhat lower phonological awareness gains and to the similarly high nonword reading gains of the children in the manipulation conditions. The lower gains in phonological awareness make sense, since the other children's training included a strong emphasis on manipulating sounds in syllables. To consider the gains in nonword reading, recall first that the articulation-only children did learn extensive articulatory awareness. They also learned all the detailed phonics that the children in other conditions did, with much more time spent practicing associating vowel sounds with mouth feelings and letters, while the children in other conditions practiced their sound manipulations. Thus children in this condition practiced far more analytic segmentation than did children in any other extra reading practice conditions in our previous studies (Olson et al., 1997; Wise & Olson, 1995; Wise, Ring, & Olson, 1997).

A second factor relating to why the articulation-only children gained so much in nonword reading resulted from the type of sounding out we encouraged during ROSS reading. We had tried to ensure that all children received equally energetic and enthusiastic support from trainers while they were reading the speech-supported stories. When children in either manipulation condition targeted words, they were encouraged to figure out each segment in the word. If these children misread a word without correcting it, the teacher at first covered the word and asked the student, "When you say [what the child said], what do you feel [or hear (for the Sound condition) at the point of contrast]?" to encourage them to check and correct their errors. For instance, if the child read mane as name, the teacher might say, "When you say name, what do you feel [or hear] at the beginning?" The child would answer, and then the teacher would uncover the word so the child could correct his or her error.

To balance this rather extensive feedback, teachers added extra attention to vowels in the articulation-only condition. When these children targeted a word, trainers asked them first to figure out the vowel in each highlighted segment, using their vowel circle charts if needed. Then they sounded out each segment, before asking for the speech support. Thus, this condition did include segmenting and blending of the vowel and consonant sounds in the story reading. If these children failed to target a misread word, they just went back and used their same sounding out strategy on that word when the teacher pointed it out. This condition's careful attention to reading word segments was tantamount to reading nonwords. In previous ROSS studies, children in the extra reading conditions (which were nonphonological) were not encouraged to be nearly so analytic and accurate with the segments before asking for speech support.

Besides the expectation of more benefits in main effects from sound manipulation, the main hypothesis of the study was that treatment effectiveness would differ according to the child's initial phoneme awareness and reading level. Let us now examine these individual differences in response to treatments. First we will discuss main effects of differences in initial profile, and then we will consider the interactions in which we had so much interest. In the current study, individual difference analyses did show some interesting general differences in performance and effects of training response to all treatments. Children with faster word-retrieval abilities as reflected in the RAN test did gain more on time-limited variables. Children with higher IQs and with higher phoneme awareness gained more on nearly all reading measures, but showed no advantage on phoneme awareness gains. The specific training in phonological awareness led to equivalent gains in that skill regardless of IQ or initial phoneme awareness, but decoding, word recognition, and comprehension gains did still correlate with initial profiles on these skills. Thus, the training seemed equally effective for children of different abilities for improvements in phoneme awareness, but the skill transferred better to reading for children with higher IQs and with higher phonological skills.

Effects of grade and reading ability were interesting in this as well as in our other studies (Wise, Ring, & Olson, 1999). For nearly all measures, including those with standard scores, younger children tended to gain the most. It is conceivable that scaling differences on the nonstandardized time-limited word recognition or nonword tests may have made larger gains easier at lower levels of these tests. However, covarying pretest performance out of each analysis should have reduced this as a problem. Also, ceiling effects should not affect the standardized test scores where we found most of the grade/reading level effects. These grade/reading level effects could be due to a number of factors. It is possible that the older children have become more rigid in their reading strategies and are less amenable to change. It is also possible that the precise and analytic phonics instruction is more applicable at lower reading levels; few of the older children were reading below third-grade level. At lower grade levels, children may read more slowly and be more willing or able to apply the strategies in their simpler reading. Supporting evidence comes from analyses of our previous study, which compared phonological training to a trained control condition with more time reading accurately in context. This study found that the older children in the extra reading practice condition gained most on time-limited word recognition (Wise et al., 1999).

Contrary to our hypothesis, treatment effects varied little or not at all according to initial levels of phoneme awareness or reading. We did not find the interactions with initial levels of phonological awareness that we were expecting. We had thought that children with especially poor phoneme awareness might gain extra benefits from the concrete sensory foundation of the training in articulatory awareness. This expectation was based on our postulation that the articulatory training might be especially powerful for refining an imprecise phonological system in children with especially severe deficits (Snowling & Hulme, 1994). Our pilot study had also lent some support to this hypothesis, although only with trends and with a small sample (Wise, Ring, Sessions, & Olson, 1997). We did find interactions with grade/reading level on two nonword reading tests, the nonword repetition test, and the PIAT standardized test of word recognition. Children's gains in the combination condition were less affected by their grade/initial reading level than in the other conditions, in that the younger/poorer readers did not gain more from the training than the older readers, as they did in the other conditions. However, these effects were small and were not evident on any other measures, so they may be spurious.

One of the goals of this study was to add power to address this question about the benefits for the most severely deficient readers. Unfortunately, the children's average reading deficits in the current study were less severe than in the pilot sample (Wise, Ring, Sessions, & Olson, 1997). For a post hoc analysis, we tried to select children from the current five-school sample to match the selected sample in the pilot study's single school. When we did this, we found only the same small number of extremely low readers in both samples, despite the large differences in the size of the total samples. Even so, the current lowest and highest five children in each manipulation condition did not replicate the previous interactions. The lowest children and the highest children performed equivalently on all reading measures, whether or not they had had specific articulatory training.

It is possible that such an interaction might be found with longer training and with a more severely deficient population. To investigate the question more fully, we hope to repeat the study with a clinic population that includes more children with severe deficits. It is also quite possible we would see this interaction in a situation without the speech-supported reading. The speech support eliminates the need for a child to sound out words. The fact that the computer can provide accurate speech and decoding support for every word needed in the text makes this type of study an ideal platform for studying the possibility of the inductive learning of phonology from accurate word reading. On the other hand, a major benefit of training in phonological awareness and decoding is that it teaches children to decode words accurately and correct their own reading errors. In effect, good phonological training teaches children to become their *own* "talking computers." The ability of the computers to allow a child to read every word accurately without knowing how print maps to sound cannot tell us how the child's reading and spelling would develop without that support. Thus the potential differences among treatments and the amount of training needed to transfer phonological skills to independent reading would surely be different for children learning to read without the support of talking computers. This should be an interesting and important topic for further study.

Nevertheless, the evidence suggests that for this length of training and in this speech-supported context, most students benefit equivalently from good phonological training with or without explicit articulatory work. Since gains from each kind of training seem equally strong on most measures, it appears at this point that the choice of programs can best be left to the skill and preference of the teacher.

SUMMARY AND CONCLUSIONS

Training in phonological awareness and analysis, integrated into speech-supported reading on computers, led to large and highly significant gains for children with reading disabilities compared to the gains of children with similar reading problems who spent their reading instructional time back in regular class. It is very important to stress how well children did with all three versions of the phonological training. This finding lends support to the consensus that phonological awareness work prior to and melded within a well-structured approach to reading is helpful for children with SRD.

The study also showed surprisingly few treatment differences, except for advantages in phoneme awareness from learning to manipulate sounds in syllables. Neither the contrast between manipulation and articulation-only nor that between explicit articulation or simpler syllable, rhyme, and phoneme manipulation made a difference on measures of reading. Aspects that were common to all three conditions in this study included phonological awareness, much phonics, and much stress on sounding out and self-correcting errors in order to read accurately in context. The study suggests these elements are important for the gains relative to regular-instruction controls, but does not specify exactly how those phonological skills should be trained. It suggests instead that it may be less important than many people think, *exactly* how that work is done for the bulk of children in a school setting. These results are empowering for teachers. They suggest that teachers should learn about language, reading, and children’s learning strengths and weaknesses; and then tailor the methods they learn to meet the needs of students and to account for the teachers’ own strengths, knowledge, and experience.

This study has limits in interpretability, because it did not include a nonphonologically trained comparison condition. However, these comparisons have been made in previous studies (Olson, Wise, Ring, & Johnson, 1997; Wise & Olson, 1995; Wise, Olson, & Ring, 1999). Other current important questions in research on remediation for reading disabilities ask how much of this training is needed and how these skills can be extended better into fluent word recognition, spelling, and writing long after training ends. These are topics of current and future studies by ourselves and by others.

REFERENCES

Adams, M. (1994). *Beginning to read.* Cambridge, MA: MIT Press.

Adams, M., Foorman, B., Lundberg, I., & Beeler, T. (1998). *Phonemic awareness in young children.* Baltimore, MD: Paul H. Brookes.

Alexander, A., Anderson, H., Voeller, K., & Torgesen, J. (1991). Phonological awareness training and remediation of analytic coding deficits in a group of severe dyslexics. *Annals of Dyslexia,* **31,** 193–207.

Ball, E., & 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.

Boulder County back to school [Special supplement]. (1997, August 10). *Boulder Daily Camera,* pp. 10–29.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection. *Nature,* **301,** 419–421.

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

Catts, H., & Vartiainen, T. (1993). *Sounds abound.* East Moline, IL: Linguisystems.

Clark, J., & Uhry, J. (1995). *Dyslexia, theory and practice of remedial instruction (2nd ed.).* Baltimore: York Press.

Cunningham, A. (1989). Phonemic awareness: Development of early reading competency. *Reading Research Quarterly,* **24,** 471–479.

Denckla, M. B., & Rudel, R. G. (1976). Rapid "automatized" naming of pictured objects, colors, letters, and numbers by normal children. *Cortex,* **10,** 186–202.

Dunn, L. M., & Markwardt, F. C. (1970). *Peabody Individualized Achievement Test.* Circle Pines, MN: American Guidance Service.

Elkonin, D. B. (1973). U.S.S.R. In J. Downing (Ed.), *Comparative reading* (pp. 551–580). New York: MacMillan.

Felton, R., & Brown, I. S. (1990). Phonological processes as predictors of specific reading skills in children at risk for reading failure. *Reading and Writing: An Interdisciplinary Journal,* **2,** 39–59.

Foorman, B., Francis, D., Winikates, D., Mehata, P., Schatschneider, C., & Fletcher, J. (1997). Early interventions for children with learning disabilities. *Scientific Studies of Reading,* **1,** 255–276.

Frith, U. (1997). Brain, mind, and behaviour in dyslexia. In C. Hulme & M. Snowling (Eds.), *Dyslexia: Biology, cognition, & intervention* (p. 12). London; Whurr.

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.

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

Kennedy, K., & Backman, J. (1993). Effectiveness of the Lindamood Auditory Discrimination in Depth program with students with learning disabilities. *Learning Disabilities Practice,* **8,** 253–259.

Lexia Learning Systems. Lexia, (1994). Lincoln, MA: Author.

Liberman, A. M., Cooper, F. S., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review,* **74,** 731–76.

Liberman, I. Y., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial & Special Education,* **6,** 8–17.

Lindamood, C., & Lindamood, P. (1975). *Auditory Discrimination in Depth (A.D.D.).* Allen, TX: DLM Teaching Resources.

Lindamood, C., & Lindamood, P. (1979). *Lindamood Auditory Conceptualization Test.* Hingham, MA: Teaching Resources Corporation.

Lindamood–Bell Learning Processes. (1997). Lindamood–Bell CD-Rom Software. San Luis Obispo, CA: Author.

Lovett, M., Borden, S., DeLuca, T., Lacerenza, L., Benson, N., & Brackstone. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer-of-learning. *Developmental Psychology,* **30,** 805–822.

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

Lyon, G. R. (1995). Toward a definition of dyslexia. *Annals of Dyslexia,* **15,** 3–30.

Montgomery, D. (1981). Do dyslexics have difficulty accessing articulatory information? *Psychological Research,* **43,** 235–243.

Olson, R. K., Forsberg, H., & Wise, B. W. (1994). Genes, environment, and the development of orthographic skills. In V. Berninger (Ed.), *The varieties of orthographic knowledge I: Theoretical and developmental issues.* Dordrecht, The Netherlands: Kluwer Academic.

Olson, R. K., & Wise, B. (1992). Reading on the computer with orthographic and speech feedback. *Reading and Writing,* **4,** 107–144.

Olson, R. K., Wise, B., Conners, F., Rack, J., & Fulker, D. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities,* **22,** 339–348.

Olson, R. K., Wise, B. W., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on post-training word recognition. *Scientific Studies of Reading,* **1,** 235–253.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activity. *Cognition and Instruction,* **2,** 117–175.

Perfetti, C. (1985). *Reading ability.* New York: Oxford Univ. Press.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly,* **27,** 28–53.

Rosner, J, & Simon, D. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities,* **4,** 384–92.

Schneider, W., Kuspert, P., Roth, H., Mechtild, V., & Marx, H. (1997). Short- and long-term effects of training phonological awareness in kindergarten: Evidence from two German studies. *Journal of Experimental Child Psychology,* **66,** 311–340.

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

Skjelfjord, V. J. (1976). Teaching children to segment spoken words as an aid in learning to read. *Journal of Learning Disabilities,* **9,** 39–48.

Skjelfjord, V. J. (1987). Phonemic segmentation: An imporant subskill in learning to read: II. *Scandinavian Journal of Educational Research,* **31,** 81–98.

Shaywitz, S. (1996, November). Dyslexia. *Scientific American,* pp. 98–104.

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London, B,* **346,** 21–27.

Stanovich, K., West, R., & Cunningham, A. (1991). Beyond phonological processes: Print exposure and orthographic processing. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Liberman.* Hillsdale, NJ: Erlbaum.

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: A Multidisciplinary Journal,* **8,** 51–61.

Truch, S. (1994). Stimulating basic reading *processes* using Auditory Discrimination in Depth. *Annals of Dyslexia,* **24,** 60–81.

Uhry, J., & Shephard, M. (1993). Segmentation/spelling instruction as part of a first-grade reading program. *Reading Research Quarterly,* **28,** 218–233.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denkla, 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.

Wagner, R., Torgesen, J., & Rashotte, C. (1994). The development of reading related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology,* **30,** 73–87.

Wechsler, D. (1974). *Examiner's manual: Wechsler Intelligence Scale for Children–Revised.* New York: Psychological Corporation.

Willett, J. B. (1989). Questions and answers in the measurement of change. *Review of Reseaarch in Education,* **15,** 345–422.

Wise, B. (1991). What reading disabled children need: What is known and how to talk about it. *Learning and Individual Differences,* **3,** 307–321.

Wise, B. W., & Olson, R. K. (1992). Spelling exploration with a talking computer improves phonological coding. *Reading and Writing,* **4,** 145–156.

Wise, B., & Olson, R. (1995). Computer-based phonological awareness & reading instruction. *Annals of Dyslexia,* **45,** 99–122.

Wise, B. W., Olson, R. K., & Lindamood, P. (1993, April). *Training phonemic awareness: Why and how in computerized instruction.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Wise, B. W., Ring, J., & Olson, R. K. (1997). Phonological awareness training with and without the computer. In C. Hulme & M. Snowling (Eds.), *Dyslexia: Biology, cognition, & intervention.* London: Whurr.

Wise, B. W., Ring, J., & Olson, R. K. (1999). *Individual differences in gains and retention of gains from computer-assisted remedial phonological training.* Manuscript in preparation.

Wise, B. W., Ring, J., Sessions, L., & Olson, R. K. (1997). Phonological awareness with and without articulation: A preliminary study. *Learning Disability Quarterly,* **20,** 211–225.

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

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

Received May 1, 1998: revised December 14, 1998