Intensive Remedial Instruction for Children with Severe Reading Disabilities:

Immediate and Long-term Outcomes From Two Instructional Approaches

Joseph K. Torgesen, Ann W. Alexander, Richard K. Wagner, Carol A. Rashotte, Kytja K. S. Voeller, and Tim Conway

Abstract

Sixty children with severe reading disabilities were randomly assigned to two instructional programs that incorporated principles of effective instruction but differed in depth and extent of instruction in phonemic awareness and phonemic decoding skills. All children received 67.5 hours of one-to-one instruction in two 50-minute sessions per day for 8 weeks. Both instructional programs produced very large improvements in generalized reading skills that were stable over a 2-year follow-up period. When compared to the growth in broad reading ability that the participants made during their previous 16 months in learning disabilities resource rooms, their growth during the intervention produced effect sizes of 4.4 for one of the interventions and 3.9 for the other. Although the children's average scores on reading accuracy and comprehension were in the average range at the end of the follow-up period, measures of reading rate showed continued severe impairment for most of the children. Within 1 year following the intervention, 40% of the children were found to be no longer in need of special education services. The two methods of instruction were not differentially effective for children who entered the study with different levels of phonological ability, and the best overall predictors of long-term growth were resource room teacher ratings of attention/behavior, general verbal ability, and prior levels of component reading skills.

One of the most daunting and clearly defined current challenges for both researchers and practicing educators is to develop, disseminate, and implement methods for teaching reading that will help all children acquire adequate reading skills. Clearly, the demands for literacy in our society are increasing very rapidly (Snow, Burns, & Griffin, 1998), and state legislatures and other governing bodies are instituting accountability standards in literacy that reflect those increasing demands. In recognition of this trend, and of the broad and serious consequences of reading failure for children's overall development, organizations such as the National Institutes of Health and the U.S. Office of Education, along with a number of private foundations, have provided substantial funds over the past two decades for research on reading, reading development, and the nature of reading disabilities in children and adults. One of the goals of this research, as defined by the National Institute of Child Health and Human Development (Lyon, Alexander, & Yaffee, 1997), is to investigate the conditions that need to be in place for all children to acquire *adequate reading skills* in school.

The present study was designed to contribute information about the conditions that need to be in place to remediate the reading difficulties of children with serious learning disabilities (LD). We acknowledge at the outset that the interventions examined in this study may not be immediately practical for broad-scale implementation in public schools. However, we agree with Schulte (1996) that the first goal of intervention research with older children should be to discover interventions that accomplish the task of remediation. When this goal is achieved, the next set of questions might concern methods for implementing these interventions in the schools. The research reported in this article is designed to answer three questions related to the remediation of reading difficulties:

1. Can either of two carefully designed instructional approaches accelerate reading growth sufficiently to bring the reading skills of children with severe reading disabilities into the average range?
2. Are there significant differences in the effectiveness of two instructional approaches, both of which contain many elements of effective instruction but differ broadly in the

specific instructional activities they emphasize?
3. Are the two methods differentially effective for children with different cognitive, linguistic, and demographic characteristics?

Existing knowledge related to each of these questions will be considered briefly in turn.

There is evidence from a variety of sources that typical public school interventions for children with reading disabilities can most accurately be characterized as stabilizing their degree of reading failure rather than remediating, or normalizing, their reading skills (Kavale, 1988; Schumaker, Deshler, & Ellis, 1986). For example, in a carefully controlled longitudinal study, McKinney (1990) found that resource room placements for children with reading disabilities produced no gains in word-level reading skills relative to nondisabled readers during a 3-year period in elementary school. The children with reading disabilities were placed in special education with an average standard score of 92, and after 3 years of special instruction, their standard score for word-level skills was 90. The children actually experienced a significant relative decline in their standing on a test of reading comprehension, falling from an average score of 94 to a standard score of 88 three years later.

Recently, Hanushek, Kain, and Rivkin (1998), using a very large sample from the Texas Schools Microdata Panel, showed that typical special education placements during the fourth- and fifth-grade years of elementary school accelerated reading growth by only .04 standard deviations over the rate the children had been achieving in their general education classroom placements. Although this represents a positive accomplishment for special education, it is hardly sufficient to normalize the reading skills of children with severe reading disabilities in any reasonable period of time.

A discussion of recently developed inclusion models of intervention (Zigmond, 1996) reached much the same conclusion as has been reported for widely used resource room, or pull-out, models of intervention. Across three different intervention sites (Zigmond et al., 1995), children with LD, as a whole, experienced little movement in reading ability relative to non-LD children in their classrooms. Although they kept pace with normal reading growth during the interventions, they did not significantly close the reading gap that got them identified as learning disabled in the first place.

These data indicate that although the reading instruction provided by special education is more effective than general education classroom instruction for children with reading disabilities, current instruction in many special education placements is not sufficient to *accelerate* reading growth so that there is reasonable hope for these children to achieve average-level skills in a reasonable period of time. Furthermore, most well-controlled intervention studies do not fully address questions about the conditions that need to be in place to remediate reading disabilities, because the interventions are not powerful enough to produce large effects on the reading skills of the children being studied. For example, one excellent and widely cited study (Lovett, Borden, Lacerenza, Benson, & Brackstone, 1994) examined the relative effectiveness of several carefully contrasted interventions. Although the study produced useful information about critical elements of reading instruction for children with severe reading disabilities and showed that core reading deficits were amenable to improvement through direct instruction, at the conclusion of the study, the children's reading skills still fell in the severely disabled range. The children in the two strongest interventions began the study with an average standard score (M = 100, SD = 15) on a measure of word-reading ability of 64.0, and at the conclusion of the study their score was 69.5, with pre- and posttest scores on a measure of reading comprehension being 66.4 and 70.8, respectively. Although one might argue that continued application of the successful instructional techniques from this study would eventually produce complete remediation of these children's reading disabilities, in the absence of direct evidence we simply do not know if this assumption is correct.

Swanson (1999) recently reported a comprehensive meta-analysis of intervention research with children with LD that found average effect sizes, using standardized reading measures, of .62 for word recognition and .45 for reading comprehension. These data are valuable because they show that we understand many of the elements of effective instruction for children with reading disabilities, but they are also misleading in that they do not provide information about the rate of normalization of reading skills. Instead, they describe the advantage in reading growth for children in an experimental condition relative to a control condition. They demonstrate that some instructional techniques are more effective than others, but they do not provide information about the extent to which the reading skills of the children in the most effective condition approached normal levels at the end of the intervention or follow-up period.

In the present study, we have described the reading growth of children in our sample in terms of changes in their standard scores on a variety of reading and nonreading measures. These data indicate the extent to which the children have changed positions within the distribution of reading ability of a large normative sample, and they also provide evidence about the extent to which their reading skills differ from average readers at the conclusion of the study. Several other studies have taken a similar approach, and they have begun to produce evidence that with the right instructional conditions, it is possible to produce very large effects on the reading skills even of children who have experienced several years of reading failure as a result of severe reading disabilities (Alexander, Anderson, Heilman, Voeller, & Torgesen, 1991; McGuinnes, McGuinnes, & McGuinnes, 1996; Truch, 1994; Wise, Ring, & Olson, 1999).

Our current understanding of the most common form of reading disability suggests that for children with reading disabilities to achieve adequate reading skills, they must receive more intensive, explicit, and systematic instruction in word-level skills than is typically provided in schools (Clark & Uhry, 1995; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Torgesen, 1998a; Vellutino et al., 1996). For example, there is now overwhelming evidence that most children with reading disabilities experience a major bottleneck to reading growth in the area of skilled word identification (Share & Stanovich, 1995; Torgesen, 1999). Compared to nondisabled readers, these children exhibit two kinds of word-level reading problems when they are reading text. First, when they encounter a word they are not familiar with, they tend to place too much reliance on guessing the word based on the context of the passage (Briggs, Austin, & Underwood, 1984; Simpson, Lorsbach, & Whitehouse, 1983), which produces a high rate of word-level errors in their reading. Their phonemic analysis skill, or ability to use phonics to assist in the word-identification process, is usually severely impaired (Bruck, 1990; Siegel, 1989). Second, children with reading disabilities encounter many more words in grade-level texts that they cannot read "by sight" than do average readers. Compared to children of the same age who are learning to read normally, the number of words that children with reading disabilities can recognize fluently and easily as orthographic units is usually quite limited (Manis, Custodio, & Szeszulski, 1993).

Current theories about the growth of word-reading ability (Ehri, 1998; Share & Stanovich, 1995) suggest that phonemic decoding skills play a critical supporting role as children begin to acquire the orthographic reading skills that enable relatively fluent and effortless identification of words in text. Thus, a primary limitation in skilled use of the alphabetic principle to decode unfamiliar words has been referred to as a "core characteristic" of the most common type of reading disability (Siegel, 1989; Stanovich, 1988). Although some children may have problems in acquiring a sight vocabulary that are relatively independent of their limited phonemic decoding skills (Bowers, Golden, Kennedy, & Young, 1994; Olson, Wise, Johnson, & Ring, 1997), early difficulties in phonemic decoding consistently account for a substantial proportion of the variance among children in the growth of their fluent word recognition skills (Wagner et al., 1997). Reading comprehension is often limited in children with reading disabilities because of difficulties with accurate and fluent word recognition, and because they have missed opportunities to acquire reading comprehension strategies (Brown, Palincsar, & Purcell, 1986).

Another major discovery from the research on reading within the last two decades has been that the word-reading difficulties of children with reading disabilities are caused primarily by weaknesses in their ability to process the phonological features of language (Liberman, Shankweiler, & Liberman, 1989). These weaknesses have been demonstrated on a variety of nonreading tasks, including measures of phonological awareness, verbal short-term memory, speed of access to phonological information in long-term memory, and some forms of speech perception (Stanovich & Siegel, 1994; Torgesen, 1995). In particular, individual differences in phonological awareness and rapid automatic naming ability have been shown to exercise unique causal influences on the rate at which children acquire important early word-reading skills (Wagner et al., 1997).

The phonological weaknesses of children with the most common form of reading disability require that they receive reading instruction that is more phonemically explicit and systematic than other children's. However, there are many ways to accomplish this aim, and there is little consensus about the nature and balance of specific instructional activities for children with severe reading disabilities. In this study, we contrasted two instructional approaches, both of which were phonemically explicit and systematic but varied in method of instruction and in depth and extent of phonemic decoding practice. The Auditory Discrimination in Depth (ADD; Lindamood & Lindamood, 1984) program was designed to directly attack the phonemic awareness problems of children with reading disabilities by helping them discover articulatory cues to the number, identity, and order of phonemes in words. It emphasizes instructional activities that teach children to "feel," as well as hear, the individual sounds in words. As implemented in this study, the vast majority of time in this program was spent building phonemic/articulatory awareness and applying this awareness to solving decoding problems with individual words. In contrast, the Embedded Phonics (EP) program, as developed for and implemented in this study, provided explicit instruction in phonemic decoding strategies (letter-sound knowledge and blending) within a direct instruction framework. Phonemic awareness was stimulated during spelling and writing activities, and word identification strategies were practiced extensively while the participants read text. The EP program provided much more practice than the ADD program in reading and comprehending meaningful text, while the ADD program provided more explicit (down to the articulatory level) and extended practice on phonemic awareness and phonemic decoding skills than the EP program.

Both of the programs incorporated principles of instruction that have generally been found to be successful with children who have LD (Swanson, 1999). That is, both programs provided ample opportunities for guided practice of new concepts, were taught using one-to-one tutoring methods, provided systematic cuing of appropriate strategies, and taught children to segment and blend the sounds in words. In some discussions of preliminary reports of this study (Allington & Woodside-Jiron, 1999; Swanson, 1999), the EP condition has been mistakenly considered to be an instructional control to test the advantages of explicit and systematic instruction in word-level skills (as implemented in the ADD condition) for children with reading disabilities. The EP condition does provide a control group that can help determine whether instruction at the articulatory level is necessary for children with severe reading disabilities to improve their phonemic awareness and grow in reading ability, but it is not a control for explicit word-level instruction per se.

Rather, this study contrasted two theoretically viable instructional strategies for children with severe reading disabilities. The goal was to determine whether two approaches that both contain explicit instruction in word-level skills but vary systematically in their depth of instruction in phonemic awareness and extent of practice in decontextualized phonemic decoding skills would affect specific reading skills in different ways. The methods' overall effectiveness was assessed by contrasting the children's progress during the study with their progress in special education placements before the experimental instruction began.

The last question we will address concerns individual differences in the way that children in our sample responded to the two interventions. We will be concerned not only with identifying child characteristics that predict stronger and weaker growth during and following the interventions, but also whether these characteristics are differentially important, depending on the nature of the intervention. For example, Foorman et al. (1998) showed that the most phonemically explicit of several interventions was particularly beneficial for children who began the study with the lowest levels of phonemic awareness. The most explicit condition produced better reading outcomes than the less explicit instructional conditions for the entire sample, but it was particularly beneficial for children with pronounced weaknesses in phonological processing. Another study of the prevention of reading disabilities (Torgesen, Wagner, Rashotte, Rose, et al., 1999) found reading growth to be significantly and uniquely predicted by entering levels of phonological ability, classroom teacher ratings of attention and behavior, and socioeconomic status (SES). In the latter study, no significant interactions were found between child and treatment characteristics, perhaps because the treatments were more similar to one another than in the study by Foorman et al., or because the sample size limited power to detect significant aptitude by treatment interactions.

Consistent with the two previously described prevention studies, Vellutino et al. (1996) found phonological variables, but not general verbal ability, to be significant predictors of growth in word-level reading skills. This finding is consistent with an emerging consensus that discrepancy between level of general intelligence and word-level reading skills should not be used as one of the core defining characteristics of children with reading disabilities (Fletcher et al., 1994; Lyon, 1995). That is, children with similar levels of phonological ability seem to respond similarly to explicit instruction in word-level reading skills, regardless of variability in general intelligence within the normal range.

Thus far, only one intervention study (Wise et al., 1999) has reported results that are somewhat inconsistent with this conclusion. That study contained instructional contrasts that varied along some of the same dimensions as those used in the present study, and Wise et al. found that individual differences in growth on measures of word-level reading skills were significantly predicted by age, general intelligence, and initial levels of phoneme awareness. When the predictors were combined into a single multiple regression, the strongest predictors were age and phonemic awareness, with level of general intelligence uniquely predicting growth on one of the two measures of real-word recognition used in the study. As in the study reported by Torgesen, Wagner, Rashotte, Rose, et al. (1999), the study by Wise and her colleagues did not find a clear pattern of interactions between the entering aptitudes, or abilities, of the students and their response to the different instructional methods.

Method

Participants

Sixty children between the ages of 8 and 10 who were previously identified as learning disabled were recruited for participation in the study. In this age range, the state of Florida requires that children who qualify for special education as learning disabled must demonstrate a discrepancy of at least 1 standard deviation between their scores on a standardized test of reading and their full scale score on an intelligence test. This criterion allows a wide range of variation in levels of intelligence.

Each year for 3 years, we selected from LD classes in three elementary schools a sample of 20 children who met the following criteria:

(a) They were identified by their teachers as having serious difficulty acquiring word-level reading skills;
(b) their average standard score on two measures of word-level reading (Word Attack and Word Identification from the Woodcock Reading Mastery Test–Revised [Woodcock, 1987] was at least 1.5 *SD*s below average for their age;
(c) their estimated verbal intelligence was above 75; and
(d) they performed below minimum required levels for their grade on a measure of phonological awareness (the Lindamood Auditory Conceptualization Test), as de-

scribed in the test manual (Lindamood & Lindamood, 1979).

We excluded children from our sample who were adopted; who showed evidence of an acquired neurologic disease; who had experienced a perinatal encephalopathic event; who had sensory deficits (hearing loss greater than 20 dB, visual acuity of at least 20/40 in the better eye); who showed evidence of chronic medical illness; who showed some form of severe psychopathology; or for whom English was a second language. We allowed variables such as SES, race, gender, and co-morbidity of ADHD to vary in a manner consistent with the larger population from which the sample was selected.

Materials and Procedure

Children identified as eligible for the study were randomly assigned to one of two groups. One of these groups (ADD) received the Auditory Discrimination in Depth Program, and the other group (EP) received an instructional program we developed called Embedded Phonics. Characteristics of the children in the two instructional conditions are provided in Table 1. We employed a two-group design for several reasons. First, we did not use a normal intervention control group because we were able to establish baseline, or preintervention rates of growth, for these children from assessments made previously by the schools. A normal intervention, or no-treatment, control group actually has little meaning in a study such as this because of the extremely high intensity of intervention we were able to provide. We also did not employ a treatment, or attentional, control group in this study because it would have been unethical to consume such a large part of the children's day with an intervention that was not focused on their primary reading difficulty. Any treatment control condition involving reading instruction would have to provide instruction with a reasonable probability of successfully affecting reading growth and thus would have simply been a third treatment group.

We completed training for 10 participants a year in each of the two treatment groups, so that a total of 30 children received training in each of the groups. Because a limited number of experienced educational therapists were available to deliver the training, it was necessary to aggregate the sample over a period of several years. All training took place in a room provided on school grounds. Treatment was provided on a 1:1 basis in two 50-minute sessions (separated by a brief break) each day of the week. This training substituted for the time the children would normally have spent in their learning disabilities resource room. Training was provided over a period of 8 to 9 weeks, until 67.5 hours of instruction were accomplished.

At the conclusion of the intensive phase of training, each child received generalization training for the next 8 weeks. The teacher who worked with the child during the intensive phase went into the LD class for one 50-minute session each week and worked with the child using classroom materials. The work done during this phase of training focused on helping the child apply the skills learned in the intensive training to tasks in the LD class. It also allowed the LD teacher to learn how to assist the child further by becoming acquainted with his or her new reading skills. Children who concluded their intensive training close to the end of the school year had this follow-up training extended into the first several weeks of the next school year.

The teachers who administered each program all had at least 1 year's experience teaching children with reading disabilities using that method or one very similar to it. The teachers who taught the ADD program were all drawn from those working at a clinic where the program had been used for the previous 5 years. Teachers who taught the EP program were drawn either from the same clinic (one staff member had several years' experience with the Reading Recovery method) or from a pool of individuals working in other private clinics who had experience using direct and synthetic phonics approaches in teaching children with reading disabilities. One teacher was trained in both methods and served as a substitute, or backup, teacher when the a child's primary

TABLE 1
Participant Characteristics

Variable	Instructional condition: ADD	Instructional condition: EP
n	30	30
Age (in months)	117.6 (10.5)	117.6 (12.6)
Grade	4.1 (.8)	4.0 (.9)
Full Scale IQ	96.2 (9.9)	95.6 (10.3)
Verbal IQ	92.2 (8.5)	93.0 (12.3)
Word Attack[a]	67.8 (12.3)	69.4 (8.5)
Word Identification[a]	67.8 (8.6)	66.5 (9.1)
Phoneme Awareness (LAC)	54.7 (15.6)	47.6 (14.3)
Gender ratio	22M/8F	21M/9F
Racial balance	18White/12Black	21White/9Black

Note. LAC = Lindamood Auditory Conceptualization Test.
[a]Subtests from the Woodcock Reading Mastery Test–Revised.

teacher was not available. Those teaching the EP program received 10 hours of preservice training in the particular sequence of instructional activities used in the Embedded Phonics program employed in this study. Separate weekly staff meetings were held for teachers in each instructional program to discuss any instructional issues that needed clarification, and also to provide consultation about any behavioral problems that occurred. Over the 3 years in which interventions were provided, five different teachers taught the ADD curriculum and five teachers taught the EP curriculum.

Pretesting of all children took place during the 2 to 3 weeks prior to the beginning of treatment. The pretest battery included

1. two measures of phonological awareness—Phoneme Elision from the Comprehensive Test of Phonological Processes (CTOPP; Wagner, Torgesen, & Rashotte, 1999) and the Lindamood Auditory Conceptualization Test;
2. two measures of phonological coding in working memory—Nonword Repetition and Memory for Digits from the CTOPP;
3. two measures of rate of access to phonological information in long-term memory—Rapid Digit Naming and Rapid Letter Naming from the CTOPP;
4. eight measures of reading skills—Word Attack, Word Identification, and Passage Comprehension from the Woodcock Reading Mastery Test-Revised (Woodcock, 1987); Phonemic Decoding Efficiency and Sight Word Efficiency from the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999); and Reading Accuracy, Reading Rate, and Reading Comprehension measures from the Gray Oral Reading Test-III (Wiederholt & Bryant, 1992);
5. two measures of other academic skills—the Spelling subtest from the Kaufman Test of Educational Achievement (Kaufman & Kaufman, 1985) and the Calculation subtest from the Woodcock-Johnson Psychoeducational Battery–Revised (Woodcock & Johnson, 1989);
6. measures of expressive and receptive language skills from the Clinical Evaluation of Language Fundamentals–Third Edition. (Semel, Wiig, & Secord, 1995);
7. a full scale IQ test—the Wechsler Intelligence Scale for Children–Revised (Wechsler, 1974)—if one had not been given by the school district within the last year;
8. three teacher checklists to measure behaviors associated with attention deficit, filled out by the child's LD resource room teacher—the Multigrade Inventory for Teachers (Agronin, Holahan, Shaywitz, & Shaywitz, 1992), The IOWA Conners Teacher Rating Scale (Loney & Milich, 1982), and the Swanson, Nolan, and Pelham (SNAP) rating scales (Atkins, Pelham, & Licht, 1985);
9. two questionnaires that were filled out by parents to assess home reading environment, family SES, and medical history; and
10. a physical/neurological examination that included assessment of fine-motor functions.

The last two sets of measures (9 and 10) were administered during the year in which training occurred. In addition, we consulted school records to obtain information about performance on individually administered standardized tests of reading that had been previously administered to each child. (The instruments used in the pretest are described in detail in an appendix available from the first author.)

All children were administered the same measures of phonological awareness, phonological short-term memory, rapid naming, reading, other academic skills, and expressive and receptive language during the 2- to 3-week period immediately following the end of the intensive training period. The same tests were also administered at 1- and 2-year intervals following the posttest in order to monitor growth in reading and language skills for an extended period following the intensive instruction.

Description of the Interventions

As mentioned in the introduction, both of the interventions in this study provided explicit and systematic instruction in word-level reading skills, but differed in their methods of teaching and in the relative amounts of time spent on various types of instructional activities. The ADD curriculum stimulated phonemic awareness via articulatory cues and spent almost all the instructional time building phonemic/articulatory awareness and individual word-reading skills. In contrast, the EP program stimulated phonemic awareness through writing and spelling activities, taught phonemic decoding strategies directly, and spent a much greater percentage of instructional time in reading and writing connected text. Each program will be described in turn.

Auditory Discrimination in Depth

It should be noted here that this program has recently been revised and is now called The Lindamood Phoneme Sequencing Program for Reading, Spelling, and Speech (Lindamood & Lindamood, 1998). The new program is not significantly different from the ADD curriculum, which was the one used in this study. As outlined in the instructor's manual (Lindamood & Lindamood, 1984), the ADD program has three major goals. The first is to provide a basis for accurate discriminations among phonemes by teaching the distinctive kinesthetic, auditory, and visual (mouth form) features associated with all of the common phonemes of the English language. Kinesthetic and visual features are taught to help make the phoneme more concrete, and to allow children to both "hear" and "feel" phonemic contrasts and identities in spoken patterns.

The second goal of the program is to teach children to use their knowledge of the distinctive features of phonemes to monitor and represent sequences of sounds in spoken syllables. This is done by engaging them in a series of problem-solving activities that allow the children to use concrete objects (mouth-form pictures and/or colored blocks) to represent sound sequences. A third important goal using these problem-solving activities is to teach children self-monitoring skills. The teacher is encouraged to use techniques (socratic questioning, proleptic exchanges) that allow children to discover methods by which they can correct themselves on these activities. Children are then taught to transfer these same self-monitoring and self-correcting strategies to the spelling and reading of individual words.

Because a "language" for talking about phonemes is taught in this program, all children began at the beginning, regardless of differences in reading skill or skill with phonics. Instruction began with teacher–student exchanges designed to help children become aware of the specific mouth movements associated with each phoneme. As part of this instruction, they also learned labels for each phoneme that are descriptive of place and manner of articulation (e.g., "lip popper," "tip tapper"), and they learned to associate each sound with a picture showing the mouth making the sound (mouth-form pictures). The children worked initially with a group of 10 consonants that were divided into 5 "brother pairs" that had the same articulatory gestures but differed in voicing (e.g., /p/ and /b/ or /t/ and /d/). When children had achieved reasonable mastery of these consonant sounds, they began learning to distinguish vowel sounds around a "vowel circle" that represented differences among vowel sounds in terms of mouth shape (smiling, open, or round) and tongue position. Following these initial discovery and labeling activities with individual phonemes, the children engaged in an extensive series of problem-solving exercises that involved representing sequences of phonemes with either mouth-form pictures or colored blocks. The purpose of this training was to help the children learn to focus on mouth movements as an additional cue to the identity, number, and sequence of sounds in syllables; it also enabled them to learn to represent these sequences with concrete visual objects.

As they learned to label each phoneme with a descriptive name, the participants were also taught to associate specific letters with each phoneme. Once they became facile at representing sequences of sound with concrete objects, it was a natural transition to begin to represent them with letters. Children learned first to encode (spell) syllables with letters using small plastic tiles and then learned to decode (read) syllables by blending the separate phonemes together. Much of this beginning work with spelling and decoding simple patterns (CV, VC, CVC combinations) included the use of nonwords, in order to reinforce the habit of "feeling" and "hearing" the individual sounds in words. Activities in which words were spelled or read in chains were used extensively to illustrate the ways in which words can change when only one phoneme is different. For example, the child might be asked to show what needed to change in order to make "pop" from "pot," and then change "pop" to "top."

The children continued to be introduced to additional consonant and vowel phonemes, with an emphasis on acquiring awareness of the articulatory gestures that are uniquely associated with each phoneme, until all 44 English consonant and vowel phonemes were introduced. The ADD program provided extensive practice in reading and spelling individual words and nonwords that followed regular patterns. However, children were also explicitly taught the principle that some words "don't play fair," so that parts of them simply have to be memorized. At the same time that children were building facility at spelling and reading regular words of increasing complexity, they also received instruction and practice in reading words that occur with high frequency in printed English (Fry, Kress, & Fountoukidis, 1993). Typically, the children were asked first to sound out words from this list and then were shown the parts of the words that needed to be pronounced differently. They then saw these words repeatedly in word drills until they could recognize them fluently. The children spent about 10% of their instructional time on sight word practice.

The children were also taught a number of simple phonics rules to help them deal with print conventions in reading and spelling, and they were taught specific strategies for dealing with multisyllable words. About 95% of the instructional time in this condition was spent in stimulating phonemic/articulatory awareness and in building skill at decoding and encoding individual words. The other 5% of the time was spent reading from the *Poppin Readers* (Smith, 1992) and the *Early Literacy Series* (Hannah, 1993), which have been specially produced to provide decodable text as children progress through the ADD program. When reading text, the children were cued through appropriate questioning to use their skills at "feeling" the sounds in words to check that the words they pronounced matched those on the printed page.

Embedded Phonics

This program was designed to provide direct, explicit instruction in word-level reading skills while providing extensive opportunities to read and write meaningful text. Children were initially given an informal assessment of their knowledge of letter-sound correspondences, blending skills, and sight word vocabulary. The content of instruction was then tailored to each individual child's needs, but the time spent on each kind of instructional activity was roughly the same for all children. Because a standard pattern of instructional activities was followed for all children in this condition, it will serve as the basis for description of the treatment.

Instruction was divided into two 50-minute sessions each day. In the first session, the following sequence of activities took place.

Ten Minutes—Introduction and Practice in Reading Sight Words. Words were selected from the same list that was used for children in the ADD program. As words were introduced for the first time, the child attempted to decode them, and both the parts that "played fair" and those that did not were identified. A phonemic guide to the pronunciation of each word appeared below the word to aid in identification, and this guide was gradually faded as the child acquired facility in identifying each word. Words were practiced repeatedly until children could pronounce them correctly within 1 second over 3 successive days.

Five Minutes—Spelling Newly Introduced Sight Words. This activity was designed to draw attention to the spelling patterns within new words and to stimulate phonemic awareness through a questioning strategy that encouraged children to "stretch" words and listen for all their sounds as they were spelled. In the case of words that involved irregularities in sound–symbol correspondences, the irregular parts were pointed out as parts of the word that just needed to be memorized.

Ten Minutes—Word Games for Fluency with Sight Words. A variety of games was used to provide repeated practice in correctly identifying words from the sight word list.

Ten Minutes—Phonics Minilesson. During this time, children were directly taught the information required for phonemic decoding and spelling. They were taught the most common spelling variations of all the consonant phonemes and all the vowel phonemes. In addition, children who required it received direct help and modeling in the skill of blending sounds together to form words. They were also taught a small number of phonetic rules/patterns (e.g., signal *e*, *r*-controlled vowels, inflected endings, and syllable patterns) that are helpful in decoding real words. Participants were able to practice these basic phonemic decoding skills by using word chains in a manner similar to the ADD program.

Fifteen Minutes—Oral Reading in Trade Book or Basal. Depending on the level of the children's phonemic decoding skill and the extent of their sight vocabulary, they practiced reading in either a graded series of trade books or a basal series that had a highly controlled vocabulary (the HBJ Bookmark Series). While reading the text, the children were encouraged to read words accurately while focusing on meaning. Word-level errors were corrected in two ways: If the error involved a phonemic decoding principle that had been taught, the children were cued to correct their error through a series of leading questions, asking them to notice the specific way the phonemes in the word they said differed from the word on the page, if the error involved an obvious violation of the context of the passage, the children were asked to think about whether the word they said made sense in the sentence. They were then encouraged to sound out as much of the word as possible and then find a word that "sounded like that and made sense in the sentence." The error-correction procedures used while the children read text were designed to build their skills in using cues from both the letters on the page and the meaning in the passage in identifying words. As the passages and stories were read, the teachers consistently emphasized meaning by asking the children specific questions, asking them to summarize what was just read in a sentence or paragraph, and asking them to predict what might happen next.

The second 50-minute daily session contained the following activities.

Ten Minutes—Sight Word Practice. Sometimes practice was provided on small groups of words that were just introduced, and other times it involved larger, cumulative samples of all words that had been taught. A combination of word-card drills and word games was used.

Five Minutes—Spelling. Children practiced spelling both sight words and words that could be spelled phonemically.

Twenty Minutes—Reading in Basal or Trade Book. This activity was similar to that in the first session, although sometimes it varied, with the child first reading the passage silently while noting and discussing difficult words, then reading it orally for fluency and accuracy.

Fifteen Minutes—Writing Activities Using Sight Words. Children were asked to compose and write meaningful sentences containing words from their practice list of sight words. Here, the emphasis was on the meaning of the words, proper use in the sentence, and correct spelling.

From these descriptions, it is clear that the two instructional methods were different in a number of important ways. First, the amount of time spent on reading and writing connected text varied substantially between the two conditions. The ADD group spent only 5% of their time applying their word-level skills to reading and comprehending text. In contrast, the EP group spent 50% of their time in meaningful activities with connected text. The ADD group spent 85% of their time learning and practicing phonemic decoding skills with individual words whereas the EP group, in the phonics minilessons and spelling practice, spent 20% of their time practicing on broadly similar activities. Whereas the ADD group spent 10% of their instructional time on learning and practicing recognition of high-frequency sight words, the EP group spent 30% of their time on this activity. Finally, the ADD group received instruction in phonemic awareness that taught children to use both kinesthetic and auditory cues for the identity of phonemes in words, whereas the EP group received only indirect training in phonemic awareness through instruction in "phonics" and phonemic spelling.

Although the complexity of each instructional program precludes hypotheses attributing instructional effects to single variables, the overall differences between the programs provide a contrast between one approach that placed primary instructional emphasis on building skills in phonemic awareness and phonemic decoding (ADD) and another approach that taught those skills while placing more emphasis on their application while reading meaningful text (EP). Again, we were interested primarily in questions about the ultimate level of reading skill attained by children in both programs, the relative effectiveness of one approach versus the other, and the predictability of individual children's responses to the interventions based on their entering characteristics.

Results

Although 60 children received remedial instruction, 10 of them moved to another community before they received the 2-year follow-up test. Six children were lost from the EP condition and four from the ADD group. The remaining sample was not significantly different from the original sample on any of the major pretest or demographic variables. The findings to be reported here are based on the 50 children (26 ADD, 24 EP) for whom complete data were available. Within this sample, there was a high degree of comorbid attention-deficit disorder. For example, 81% of the ADD group was diagnosed with either attention-deficit disorder or attention-deficit/hyperactivity disorder, while 71% of the EP group had similar diagnoses. These diagnostic judgments were made by Dr. Alexander and Dr. Voeller, both highly experienced clinicians, on the basis of observations during pretesting and information from the teacher survey forms. Of the children diagnosed with attention-deficit disorder, approximately half received some form of stimulant medication during the course of the intervention or follow-up. However, medication condition during pre- and posttesting was not carefully controlled in this study, nor was medication consistent over the course of the treatment. Thus, medication must be regarded in this study as a variable that may have introduced error variance, rather than one whose effects were systematically examined.

Table 2 reports pre-, post-, and follow-up testing on all the reading variables for children in both conditions. As can be seen from this table, the effects of intervention were both substantial and stable over the course of the 2-year follow-up period, and outcomes were very similar for both instructional methods. We analyzed growth during the intensive treatment period, separately from growth during the follow-up period in order to isolate specific effects of the treatment versus long-term outcomes and maintenance of instructional gains.

During the treatment period, a series of 2 (ADD vs. EP) × 2 (pre- vs. posttest) repeated-measures ANOVAs showed that growth in standard scores was statistically significant ($p < .01$) for all the reading measures, with $Fs(1, 47)$ ranging from a high of 309.2 for Word Attack to a low of 7.6 for rate on the GORT-III. The only measures for which the two treatment groups showed different rates of growth from pre- to postest were Word Attack from the Woodcock Reading Mastery Test–Revised, $F(1, 47) = 8.4$, and the Rate $F(1, 47) = 4.2$, and Accuracy, $F(1, 47) = 4.4$, measures from the Gray Oral Reading Test–III. The time × treatment interaction for the Comprehension score from the GORT-III just failed to reach the .05 probability level, $F(1, 47) = 3.6$, $p = .06$. In all of these cases, children in the ADD condition improved more than those in the EP condition.

Growth during the follow-up period was examined using 2 (ADD vs. EP) × 3 (posttest vs. 1-year vs. 2-year) repeated-measures ANOVAs. During the follow-up period, the groups showed a decline in standard scores on one test, gains in standard scores on three tests, and stable performance on four others. The children lost ground relative to normal growth on the Word Attack test (this means not that their actual skills declined but, rather, that they did not keep pace with normal growth during this period), $F(2, 94) = 6.6$, $p < .01$, with most of this decline being shown by children in the ADD group. On the other measure of phonemic decoding skill (Phonemic Decoding Efficiency), the children's performance was stable during the follow-up period. In contrast to their decoding skills, the children's standard scores on measures of sight word vocabulary increased during the 2 years following intervention. Changes on both the Word Identification, $F(2, 94) = 9.1$, and the Sight Word Efficiency, $F(2, 94) = 27.9$, test were statistically reliable, $p < .01$. The children's standard scores for the Passage Comprehension test from the WRMT-R also increased significantly during the follow-up period, $F(2, 94) = 8.6$. Standard scores on the GORT-III Accuracy, Rate, and Comprehension measures did not show significant change during the follow-up period. For none of the reading variables was there a significant treatment × time interaction, nor were the two groups significantly different from one another on any of the measures during the follow-up period.

One of the most striking findings from the data in Table 2 was the large difference between gains on measures of word-reading accuracy and comprehension and those on the rate mea-

TABLE 2
Outcomes for Reading Measures at All Measurement Points

	Group							
	ADD				EP			
Measure	Pre	Post	1-year	2-year	Pre	Post	1-year	2-year
Word Attack	68.5	96.4	90.7	91.8	70.1	90.3	87.0	89.9
SD	11.8	7.0	9.3	12.5	9.2	8.3	8.9	10.4
Word Identification	68.9	82.4	82.7	87.0	66.4	80.5	78.2	83.9
SD	8.3	11.2	9.6	12.1	8.7	9.6	11.3	12.2
Passage Comp.	83.0	91.0	92.8	94.7	82.2	92.0	91.5	96.9
SD	19.4	9.0	8.0	8.9	11.0	19.8	10.8	11.5
Phoneme Decoding Eff.	74.3	83.3	81.6	84.3	75.7	83.7	80.6	82.7
	6.4	4.8	7.4	7.5	6.1	5.8	9.1	10.7
Sight Word Effic.	69.7	74.5	79.3	82.1	67.3	72.7	74.4	77.8
	6.3	5.8	6.4	6.5	7.5	7.8	9.6	9.5
Gray Accuracy[a]	73.8	89.4	93.7	91.3	77.5	87.5	90.8	90.4
SD	8.8	12.4	12.3	15.5	10.3	13.4	14.8	14.7
Gray Rate[a]	71.3	75.4	75.0	72.7	71.5	72.1	72.1	70.7
SD	5.9	8.2	9.3	9.5	8.4	7.9	13.2	12.9
Gray Comp.[a]	73.3	85.6	90.2	87.9	79.4	86.0	88.1	87.2
SD	10.8	10.0	10.0	11.8	12.3	10.4	12.2	15.1

Note. Preliminary reports on the immediate posttest and 1-year follow-up for some of these measures were presented in earlier discussions of this study that appeared as part of a special issue in *Learning Disabilities: An Interdisciplinary Journal* (see Torgesen, Wagner, Rashotte, Alexander, & Conway, 1997), and in book chapters in *Word Recognition in Beginning Reading* (see Torgesen & Burgess, 1998), *Specific Reading Disability: A View of the Spectrum* (Torgesen, 1998a), *Language Basis of Reading Disabilities* (Torgesen, 1998b), and *Perspectives on Learning Disabilities* (Torgesen, 1999).
[a]To be consistent with the other measures, standard scores from the Gray Oral Reading Tests–III (Wiederholt & Bryant, 1992) were transformed to have a mean of 100 and standard deviation of 15.

sures. On the rate measure from the *GORT-III*, which was our only measure for text processing rate, the children showed almost no change in their standing relative to average readers. They began the study almost 2 standard deviations below average, and they were at roughly the same point at the conclusion of the follow-up period. It must be emphasized, however, that this *does not* mean that they did not become more fluent readers. To illustrate the absolute level of gains in fluency, we calculated a words-per-minute score on the two most difficult passages they read at the pretest and compared this to their rate for passages of the same level of difficulty at the 2-year follow-up test. For the most difficult passage at pretest, rate changed from 38 to 101 words per minute, with a corresponding drop in errors from 10 to 2. On the next most difficult passage, their rate changed from 42 to 104 words per minute, with a drop in errors from six to one. Thus, for passages that had a constant level of difficulty, the children's reading rate more than doubled from pretest to end of the follow-up period.

To establish that the reading growth attained during the intervention period for these children was significantly different from that of the preintervention period, we used standard scores on the Broad Reading Cluster from the Woodcock-Johnson Psycho-Educational Battery–Revised (Woodcock & Johnson, 1989) that were obtained before, during, and following the intervention period. The Broad Reading Cluster is composed of scores on the Word Identification and Passage Comprehension subtests of the WJP-R, which are similar but not identical to the corresponding subtests of the WRMT-R. We obtained scores prior to the treatment from school records, with the average period elapsed between the school tests and our pretest being 16.6 months. During this 16-month time period, the children received remedial reading instruction in resource room settings. Although a variety of specific instructional methods were used in the children's resource rooms, they could all be characterized as providing direct instruction in basic reading skills, including phonics, sight word vocabulary, and comprehension. The teacher-to-student ratio in these resource rooms ranged between 1:8 and 1:18.

Standard scores for the Broad Reading Cluster for each group are plotted in Figure 1, and it is obvious from this figure that both treatments had a powerful effect on the rate of reading growth of children in the sample.

Growth slopes were calculated for each child by dividing the change in standard score within a given period by the number of months intervening between measurement points, and these slopes were analyzed in two planned contrasts. The first contrast compared growth during the preintervention period with that occurring during intervention and found that these slopes were significantly different, $F(1, 47) = 35.3$, $p < .01$. The second contrast found that the rate of growth during the preintervention period was not significantly different from the rate following the intervention. Effect sizes for the two interventions were calculated by dividing the difference between the slope during treatment and the slope prior to treatment by the pooled variability of the preintervention slopes. The effect size for the Broad Reading Cluster, using the "regular resource room intervention" as the control group, was 4.4 for the ADD group and 3.9 for the EP group.

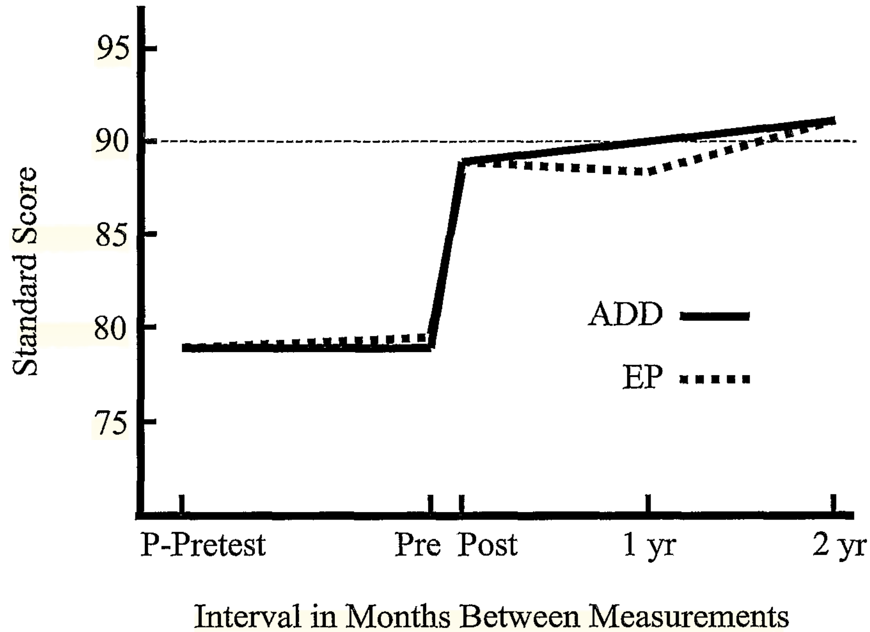

FIGURE 1. Standard scores on the broad reading cluster before, during, and following the intensive intervention.

Performance on nonreading measures of language, phonological processing, and other academic skills is presented in Table 3. These data were analyzed in a manner similar to the reading variables, and the first set of analyses showed that children in both groups improved significantly during the treatment phase in their expressive and receptive language skills, $Fs(2, 94) = 30.4$ and 42.2, respectively. Although the children's standard scores on these language measures continued to improve during the follow-up period, the change was not statistically reliable.

The phonemic awareness measures were analyzed separately, as standard scores were reported for one of the measures (Elision), and raw scores for the other (LAC). For the Elision measure, there was a significant and roughly equal effect of both interventions $F(1, 47) = 42.5$, with no significant changes in standard scores during the follow-up period. Scores on the LAC test also improved significantly during the intervention, $F(1, 47) = 245.3$, but there was also a significant time × treatment interaction, $F(1, 47) = 14.7$, with the ADD group improving the most on this measure. During the follow-up period, there was, again, a significant time × treatment interaction $F(2, 92) = 8.0$, with the ADD group's scores declining, whereas the scores for the EP group increased. Post hoc contrasts showed that the groups were not reliably different from one another on the LAC test 2 years following intervention.

Because the two naming tasks were very similar to one another, and the two memory tasks were also similar to each other, these measures were analyzed with multivariate repeated-measures analyses. For the naming measures, there was a significant multivariate effect of the intervention, Wilk's Lambda = 8.9, $p < .01$, and univariate treatment effects for both measures were also statistically reliable, $Fs(1, 47) = 8.3$ and 16.6 for digits and letters, respectively. No change was seen on the rapid naming measures during the follow-up period. The phonological memory measures also showed a significant multivariate effect of the intervention, Wilk's Lambda = 26.3, but univariate effects were reliable only for the nonword repetition measure, $F(1, 47) = 41.2$, $p < .01$. The same pattern was true for the follow-up period, with significant overall improvement on the memory measures, Wilk's Lambda = 4.2, but reliable improvement on only the measure of nonword repetition. For neither of these measures was there a significant time × treatment interaction.

There was no effect of the intervention on the calculation subtest, which indicates that the intervention did not have a generalized impact on academic performance; rather, its impact was restricted to the reading/language domain. Finally, slightly different outcomes were obtained for the two spelling measures. Standard scores on the Kaufman spelling test improved during the treatment phase, $F(1, 47) = 12.7$, $p < .01$, but this improvement was qualified by a significant time × treatment interaction, $F(1, 47) = 4.2$, $p < .05$. Follow-up analyses indicated that only the EP group improved their spelling scores. However, this effect was lost during the follow-up period, as the

TABLE 3
Outcomes for Nonreading Measures at All Measurement Points

Measure	ADD Pre	ADD Post	ADD 1-year	ADD 2-year	EP Pre	EP Post	EP 1-year	EP 2-year
Language								
CELF Total	76.3	85.9	87.9	89.7	81.0	86.6	89.0	89.9
SD	9.0	11.8	11.5	14.0	12.0	10.7	11.5	19.3
CELF Exp. Lang.	76.2	83.1	85.0	87.4	78.5	83.8	87.2	86.2
SD	9.1	10.0	8.3	13.3	11.5	10.8	10.8	10.5
CELF Rec. Lang.	79.7	90.7	92.9	93.8	85.0	91.5	92.4	95.2
SD	11.0	15.1	15.7	16.2	13.9	12.2	13.9	11.2
Phonological measures								
LAC	56.3	89.2	82.3	82.2	49.4	69.0	72.0	76.2
SD	13.5	10.1	14.0	14.1	15.3	17.3	13.9	13.4
Phoneme Elis.	88.8	101.0	97.9	97.9	84.2	97.9	94.4	98.8
SD	13.1	14.3	13.7	13.9	11.2	12.8	14.3	13.6
Digit Memory	88.8	92.0	98.1	97.1	88.3	92.7	99.0	102.7
SD	14.6	14.9	16.3	19.8	13.5	16.8	15.7	16.9
Nonword rep.	88.2	100.8	108.4	112.9	89.6	103.0	108.2	112.8
SD	15.3	16.1	15.6	13.7	18.0	19.0	17.4	14.0
RAN Digits	86.9	90.2	91.7	91.3	84.2	87.9	89.6	87.9
SD	10.7	10.3	12.4	13.2	9.3	11.9	11.1	13.0
RAN Letters	90.4	95.8	92.5	96.0	87.9	93.1	92.4	93.3
SD	10.7	10.3	10.6	11.3	8.3	11.5	10.2	12.4
Other Academic Skills								
Calculation	93.1	95.0	87.7	89.6	89.0	86.9	86.5	89.4
SD	14.4	14.1	13.5	12.4	11.8	11.5	10.1	11.1
Kauf. Spelling	75.6	77.5	76.7	76.2	74.4	80.0	74.0	75.3
SD	4.6	5.0	5.8	6.7	4.9	7.4	6.6	6.0
Developmental Sp.	1.15	1.52	2.25	2.91	1.21	2.04	1.78	2.83
SD	.78	.77	1.22	1.16	.83	1.22	1.25	.92

Note. CELF = Clinical Evaluation of Language Fundamentals (Semel, Wiig, & Secord, 1995); LAC = Lindamood Auditory Conceptualization Test (Lindamood & Lindamood, 1979); Phoneme Elision, Digit Memory, Nonword Repetition, RAN Digits, and RAN Letters are all subtests of the Comprehensive Test of Phonological Processes (Wagner, Torgesen, & Rashotte, 1999). Kauf. Spelling = Spelling subtest from the Kaufman Test of Educational Achievement (Kaufman & Kaufman, 1985). With the exception of the LAC test, for which raw scores are reported, all scores are standard scores based on a distribution with a mean of 100 and standard deviation of 15.

children's growth in spelling skill did not keep pace with that of the standardization sample, and the decline in growth rate was particularly pronounced for the EP group, $F(2, 94) = 4.5$, $p < .05$. For the measure of phonological spelling, there was also a significant effect of the intensive intervention, $F(1, 47) = 14.0$, but this effect was similar for both groups. During the follow-up period, the students continued to increase their raw scores on this measure, $F(2, 92) = 16.0$, but the rate of growth for the ADD group was stronger than for the EP group, $F(2, 92) = 3.3$, $p < .05$.

Findings about the remedial sufficiency of the instructional conditions in this study for children with severe reading disabilities vary both across types of reading skill and across individual children. If we take a standard score of 90 as the lower bound of the average range of reading ability, then the sample as a whole remained clearly deficient in reading rate both at the end of the intensive intervention and at the end of the 2-year follow-up period. In contrast, the accuracy of their phonemic decoding skills, word reading in text, and reading comprehension for short passages (WRMT-R) fell within the average range at the end of the 2-year follow-up period. Comprehension performance on longer passages (GORT-III) and untimed single word recognition fell slightly below the average range at the end of the follow-up period.

Apart from the remaining general deficiency in reading rate, the interventions were clearly not sufficient for a significant proportion of the children. Table 4 reports the percentage of children who attained standard scores below 90 at the pretest, posttest, and 2-year follow-up points. A little over one third of the sample were still performing below the average range on phonemic decoding skills and ability to read words accurately in text, whereas substantially more than half were still below average in their ability to identify real words without the aid of con-

TABLE 4
Percentage of Children in Each Instructional Group with Standard Scores Below 90

	Group					
	ADD			EP		
Measure	Pre	Post	2-year	Pre	Post	2-year
Word Attack	100	16	31	100	54	46
Word Identification	100	72	61	100	83	67
Passage Comp.	65	40	15	75	46	21
Phoneme Dec. Eff.	100	92	73	100	83	83
Sight Word Effic.	100	100	88	100	100	87
Gray Accuracy	92	40	35	79	62	35
Gray Rate	100	96	88	100	100	91
Gray Comp.	92	64	46	71	50	52

text. Thus, in terms of the word-level reading difficulties that are a core characteristic of children with severe reading disabilities, the interventions in this study were sufficiently powerful to "normalize" the skills of approximately one half to two thirds of the children, depending on the measure being considered.

Outcomes for the measures of reading comprehension were slightly better than for the word-level measures. On the Passage Comprehension subtest of the WRMT-R, about 80% to 85% of the children performed within the average range, while the corresponding percentage for the comprehension measure from the GORT-III was about 50%. Another way to evaluate comprehension outcomes is to compare the children's performance on a measure of general language comprehension with their performance on the reading comprehension measures. Our most current measure of general verbal ability at the end of the follow-up period was the receptive language measure from the CELF-R. If we assume that a goal of remedial reading instruction for children with severe reading disabilities is to help them acquire reading comprehension skills that are consistent with their general language ability, then children who score at or above their language ability on a reading comprehension measure can be considered to be successfully remediated. At the end of the follow-up period, 66% of the children obtained scores on the Passage Comprehension subtest within or above 1 standard error of measurement of their Receptive Language score, while the corresponding figure for the GORT-III comprehension measure was 53%.

To communicate the general growth trajectories for children who showed different levels of reading skill at the end of the follow-up period, the sample was divided into four groups, depending on their standard score on each of several reading measures. The Low, Mid–, Mid+, and High groups each contained approximately 25% of the children. Figure 2 reports standard scores at pre-, post-, 1-year, and 2-year follow-up tests for children in each of these four groups on the Word Attack and Word Identification measures from the WRMT-R, and on the Comprehension and Rate measures from the GORT-III. It is obvious from these graphs that the strongest divergence in growth rate among the groups took place during the follow-up period. For example, there was a significant time × group interaction, $F(3, 45) = 3.6$, $p < .05$, only for the Word Identification measure during the treatment period. However, the interaction between time and group was significant for all measures during the follow-up period, $Fs(6, 90) = 9.6$, 5.7, 6.3, and 7.3 for Word Attack, Word Identification, Comprehension, and Rate, respectively.

The sample in this study was too small for satisfactory analyses of individual growth curves using hierarchical linear modeling, as we did in an earlier study (Torgesen et al., 1999). We had originally taken measures of word-level reading skills and phonological variables at a midpoint during treatment in order to calculate the linear component of growth for each participant in the study. However, the reliability with which the slope parameter could be estimated was unsatisfactory (reliabilities ranged from .32 to .65) during both the treatment and the follow-up period. As an alternative to estimating the slope using hierarchical linear modeling, we calculated the reliability of gain scores for our reading measures and found them to be substantially higher. The average reliability of the gain scores for our reading measures from pre- to posttest was .76 (range = .68 to .86), and for the follow-up period (posttest to Year 2 follow-up) it was .81 (range = .70 to .89). For economy of presentation, we will present the growth analysis for only four of our eight reading measures (Word Attack and Word Identification from the WRMT-R, Comprehension and Rate from the GORT-III). Results for the other measures do not depart in meaningful ways from the results that are presented here.

The results presented in Table 5 address the question of which variables predicted growth on the four reading measures during the treatment and follow-up periods. For each predictor variable, the beginning point on the outcome variable (referred to as the *autoregressor*) was entered first into the equation, so the regression coefficients reported in Table 5 describe the degree to which a given variable predicted growth in the outcome measure with beginning performance level controlled. This was done to avoid spuri-

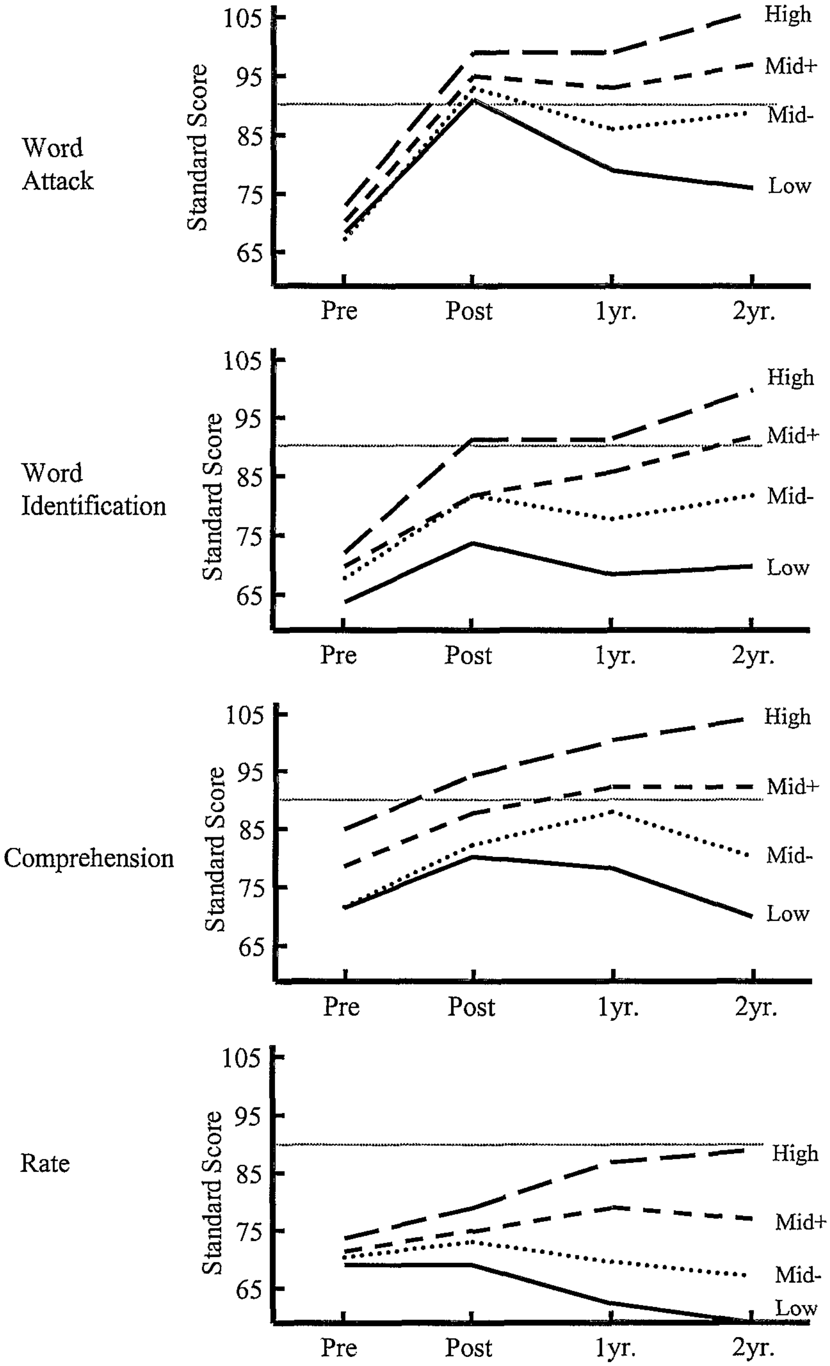

FIGURE 2. Growth on Word Attack, Word Identification, Comprehension, and Rate measures for children who fell within different quartiles on final outcomes for those measures.

ous relationships between predictor and outcome variables in cases where the predictor variables were highly correlated with the beginning score on the outcome variable. The variables used as predictors for the treatment period were taken from the pretest, whereas those used to predict growth during the follow-up period were taken from the posttest. Age is eliminated as a factor in all these analyses, as standard scores were used for all reading, phonological, and cognitive variables. The naming measure represented an average of naming rates for digits and letters, and the memory variable was an average of the two measures of phonological memory. Phonological awareness was represented by scores on the Elision test. The attention/behavior variable was constructed from the attention and activity categories of the Multigrade Inventory for Teachers and the attention items from the IOWA Connors Teacher Rating Scale. Average scores on each of these scales were transformed to within-sample *z* scores, and a unit weighted composite formed. For this variable, the higher the score, the greater the incidence of rated attention/behavior problems. The SES variable was derived from father's and mother's education and occupations and was scored according to criteria described in Hollingshead (1975). General verbal ability was represented by the Receptive Language score from the CELF-R, as it was the most current measure of verbal ability available and showed the most consistent pattern of relationships with the outcome variables. We also used reading scores as predictor variables in cases where a given score (e.g., Word Attack) could reasonably be construed as a component skill that might contribute to growth on another reading skill (e.g. Word Identification or Comprehension).

Although not reported in Table 5, the children's gender and race were not significantly related to growth on any of the variables. The values in Table 5 are standardized regression coefficients, so that it is meaningful to compare values both across predictor variables and across gain scores. The overall pattern of relationships indicates that other than the autoregressors, the predictor variables most reliably associated with growth were teacher ratings of attention, receptive language scores, and measures of other reading skills. There were very strong

TABLE 5
Prediction of Gain Scores During the Intervention Period and During the Two-Year Follow-Up period

	Intervention period				Follow-up period			
Predictor variable	WA	WID	Comp.	Rate	WA	WID	Comp.	Rate
Auto regressor	–.70**	–.18	–.61**	–.28*	–.40**	–.32*	–.22	–.36**
Age	.02	.26	.03	–.05	.14	.11	.19	.31*
Attention	.20	–.26	–.13	–.08	–.05	–.33*	–.46**	–.41**
PA	.16	.23	.12	.14	.13	.03	.19	.11
Naming	–.10	–.20	.03	.11	.10	.20	.20	.06
Memory	–.04	.06	.08	–.17	–.08	–.18	.02	–.12
Receptive Lang.	.11	.18	.06	.05	.25	.34*	.40**	.34*
SES	.05	–.08	.11	–.15	.11	.16	.21	.34*
Word Attack		.38*	.14	.00		.00	.24	.01
Nonword Effic.		.06	.04	.06		.17	.32*	.46**
Word Identification			.31*	.44**			.47**	.45**
Sight Word Effic.			.32*	.41**			.41**	.62**

Note. PA = phonological awareness; SES = socio-economic status; Nonword Effic. = nonword efficiency; Sight Word Effic. = sight word efficiency.
$^*p < .05$. $^{**}p < .01$.

negative relationships between entering score on a given variable and rate of growth on that variable, particularly during the intervention period for the Word Attack and Comprehension measures. This means that children who began the study (and the follow-up period) with the lowest scores showed the highest rate of growth on most of the reading variables. When we tested for interactions between predictor variables and treatment groups, we found no cases in which the variables reliably predicted growth differentially across groups.

To determine the degree to which individual differences on the predictor variables could explain variance in growth on the outcome measures, we conducted a series of analyses in which the variables significantly related to growth (including the autoregressor, regardless of whether it was significantly related to growth) were entered simultaneously into the regression analyses. The one exception to this entry rule applied to the reading measures used as predictors. Only the reading predictor variable with the highest individual relationship to growth was entered, to preclude the presence of too many predictors in the equation that were highly correlated with one another. The outcomes from these analyses are reported in Table 6. The standardized coefficient for each variable within the simultaneous regression is provided, along with the total percentage of variance accounted for by the regression equation. During the intervention period, growth was uniquely predicted only by the autoregressor and measures of lower level reading skills. For example, both strengths on the Word Attack test and relative weaknesses on the Word Identification test predicted growth in word identification. For the rate and comprehension measures, growth was uniquely and positively related to measures of real-word reading ability, in addition to being negatively related to the autoregressor. For the follow-up period in which a broader range of variables predicted growth in the individual analyses, the only variables that contributed uniquely to the prediction of growth were the autoregressors, attentional ratings, and, in the case of the rate measure, sight word efficiency.

Although the children's reading skills at the beginning of the treatment and follow-up periods tended to be *negatively* related to the amount of growth they experienced during each period, these beginning scores were *positively* related to the skill levels that were ultimately attained. That is, children who came into the study with higher reading skills tended to have higher reading skills at the end of the study. Furthermore, variables that most reliably predicted *rate of growth* (particularly attention and receptive language scores) were also those that tended to best predict the *ultimate skill level* the children attained. Table 7 reports standardized regression coefficients estimating the relationship between the same predictive variables used in the previous analyses and treatment outcomes for the intensive instructional and follow-up periods. These coefficients index the strength of relationships between the predictor and outcome variables with no other variables in the prediction equation.

TABLE 6
Simultaneous Prediction of Growth During Intervention and Follow-Up Periods

Word Attack		**Word Identification**		**Comprehension**		**Rate**	
				Intervention period			
Autoregressor	−.70**	Autoregressor	−.29*	Autoregressor	−.81**	Autoregressor	−.47**
		Word Attack	.38**	Sight Wd. Eff.	.32*	Word Ident.	.44**
Total R^2 = .49		Total R^2 = .17		Total R^2 = .43		Total R^2 = .24	
				Follow-up period			
Word Attack		**Word Identification**		**Comprehension**		**Rate**	
Autoregressor	−.40**	Autoregressor	−.49*	Autoregressor	−.45*	Autoregressor	−.63**
		Attention	−.26	Attention	−.37*	Attention	−.36*
		Recept. Lang.	.24	Recept. Lang.	.20	Recept. Lang.	.18
				Word Ident.	.12	Age	−.13
						SES	−.17
						Sight Wd. Eff.	.57**
Total R^2 = .16		Total R^2 = .19		Total R^2 = .31		Total R^2 = .63	

*$p < .05$. **$p < .01$.

Notable here are the significant relationships between individual differences in phonemic awareness at the beginning of each period (treatment and follow-up), and both word-level and comprehension outcomes. In contrast, the rapid naming tests were related to individual differences in reading rate at the end of the intervention period but were not related to individual differences on the rest of the reading outcomes. Performance on the Receptive Language measure was also a robust predictor of ultimate reading levels during both the treatment and the follow-up period. And as might be expected, beginning scores on component reading skills were also strongly predictive of the scores on the outcome measures at the end of each measurement period. As with the gain scores, when the correlations in Table 7 were calculated for each group separately, no significant patterns of difference existed for the two treatments.

In a manner similar to the analyses conducted for the gain scores, Table 8 reports the results from multiple regressions that included all the significant predictors from the bivariate analyses in Table 7. As was the case for the growth analysis, the only measures that contributed uniquely to the prediction of outcomes during the intervention period were initial status on the outcome in question (autoregressor) and component reading skills in word attack and sight word identification. Although phonological awareness and receptive language skills predicted outcomes reliably when entered by themselves, when considered simultaneously with the autoregressor and component reading skills, they did not contribute uniquely to the prediction of outcomes at the end of the intervention period. For the follow-up period, the most consistent predictors of ultimate outcomes were the initial status variables (autoregressors). Attentional behaviors explained additional unique variance for the comprehension and rate measures, and the sight word efficiency measure at posttest was the best predictor of reading rate at the end of the 2-year follow-up period.

Although we could find no evidence in the correlational analysis that children with different entering characteristics responded differently to the two interventions, the sample size in this study did not provide sufficient power for a reasonably sensitive test of this possibility. Because of the specific possibility that the ADD intervention, with its focus on phonemic awareness at the articulatory level and its extensive practice in phonemic decoding, may have been particularly helpful for children who entered the study with the most severe phonological disabilities (Wise et al., 1999), this hypothesis was tested in an extreme groups analysis. Pretest standard scores on the Word Attack and Elision tests were combined into an index of phonological weakness, and children who fell within the bottom and top 20% on this index within each treatment group were included in the analysis. Figure 3 shows

TABLE 7
Prediction of Posttest and Two-year Follow-Up Scores from Measures Taken at the Beginning of the Treatment and Follow-Up Periods

	Posttest score				2 year follow-up score			
Predictor variable	**WA**	**WID**	**Comp.**	**Rate**	**WA**	**WID**	**Comp.**	**Rate**
Auto regressor	.42**	.69**	.53**	.67**	.32*	.59**	.57**	.40**
Age	.05	.16	.15	–.36*	.09	.16	.21	.07
Attention	.18	–.24	–.19	–.01	.03	–.37*	–.45**	–.35*
PA	.33*	.33*	.22	.06	.36*	.35*	.34*	.18
Naming	–.16	.10	.19	.29*	.08	.23	.26	.10
Memory	.04	.10	.22	–.15	–.09	–.11	.09	–.18
Receptive Lang.	.25	.37**	.31*	.10	.33*	.55**	.51**	.39*
SES	.04	.13	.23	.02	.08	.19	.27	.26
Word Attack		.45**	.22	.03		.28*	.26	.05
Nonword Effic.		.31*	.25	.20		.36*	.41**	.54**
Word Identification			.52**	.56*			.59**	.53**
Sight Word Effic.			.54**	.50*			.55*	.64**

Note. PA = phonological awareness; SES = socio-economic status; Nonword Effic. = nonword efficiency; Sight Word Effic. = sight word efficiency.
*$p < .05$. **$p < .01$.

TABLE 8
Simultaneous Prediction of Outcomes at the End of the Intervention and Follow-Up Periods

Intervention period

Word Attack		Word Identification		Comprehension		Rate	
Autoregressor	.35*	Autoregressor	.59**	Autoregressor	.28*	Autoregressor	.44**
Phon. Aware	.20	Phon. Aware	.07	Recept. Lang.	.07	Age	–.13
		Recept. Lang	.05	Sight Wd. Eff.	.34*	Naming	.04
		Word Attack	.25*			Word Ident.	.36*
Total $R^2 = .21$		Total $R^2 = .56$		Total $R^2 = .36$		Total $R^2 = .55$	

Follow-up period

Word Attack		Word Identification		Comprehension		Rate	
Autoregressor	.20	Autoregressor	.39*	Autoregressor	.38*	Autoregressor	.07
Phon. Aware.	.02	Attention	–.10	Attention	–.31*	Attention	–.27**
Recept. Lang.	.25	Phon. Aware	–.17	Phon. Aware	.09	Recept. Lang	.07
		Recept. Lang.	.29	Recept. Lang.	.14	Sight Wd. Eff.	64**
		NonWord Eff.	.17	Word Ident.	.06		
Total $R^2 = .15$		Total $R^2 = .48$		Total $R^2 = .50$		Total $R^2 = .60$	

*$p < .05$. **$p < .01$.

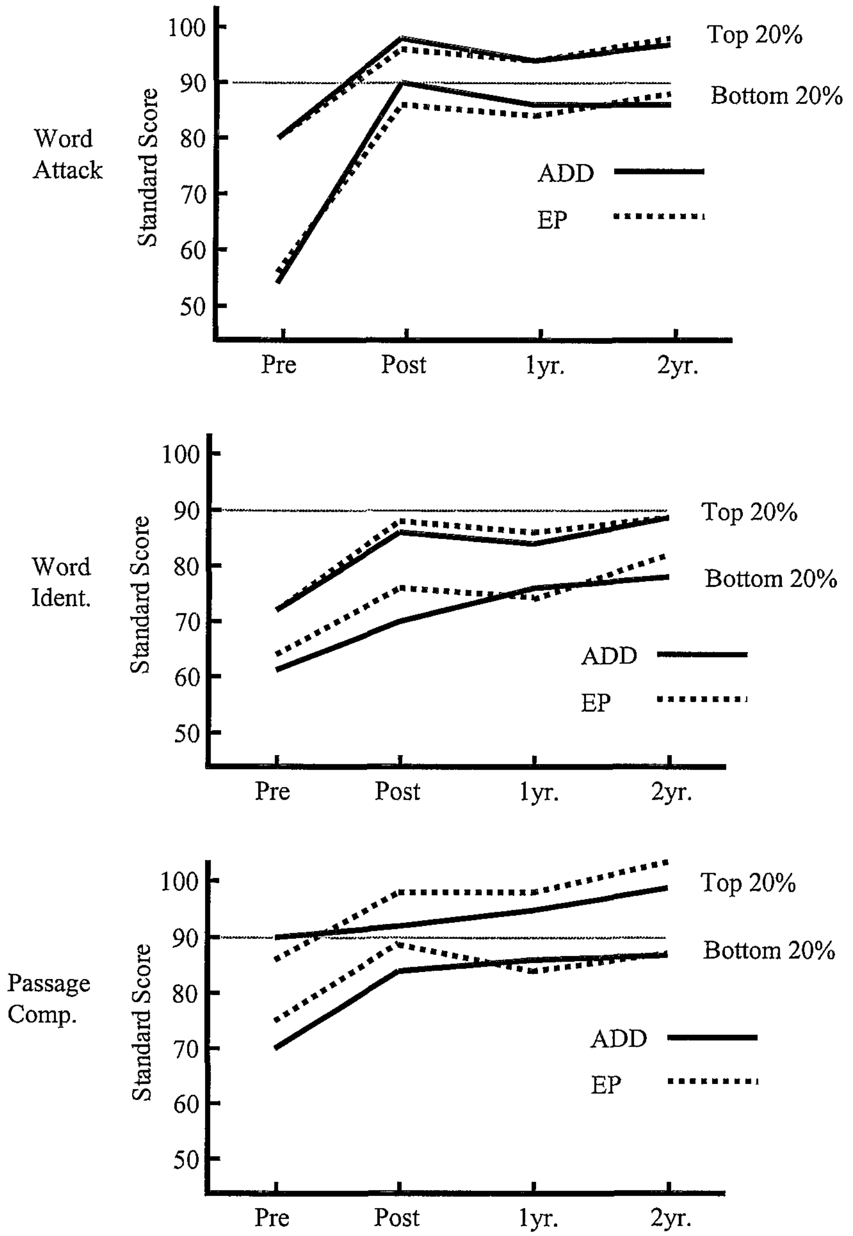

FIGURE 3. Reading growth of children who fell in the lowest and highest 20% in phonological awareness and phonemic decoding ability at pretest.

scores at all measurement points for the Word Attack, Word Identification, and Passage Comprehension measures from the WRMT-R. It is clear from this figure that there was not a striking advantage for the ADD intervention with children who entered the study with particularly weak phonemic awareness and phonemic decoding skills.

Approximately 40% of the children were judged to be no longer in need of special education services and were returned full-time to the general education classroom within the first year following the end of the intervention. Reports from the school district in which this study was conducted indicated that this rate was substantially higher than the normal 5% rate of similar-aged children who were returned to the general education classroom from learning disabilities placements over a similar period of time. A higher proportion of children from the ADD group (46%) than from the EP group (33)% were removed from special education, although this difference is not statistically reliable. Table 9 reports the characteristics and selected reading scores of children who were removed from special education as opposed to those who were not. From the information in this table, we can determine that the children who were returned to the general education classroom following our intervention tended to be older, have less severe attentional/behavior problems, have stronger general verbal abilities, and come from homes of higher socioeconomic status. In addition, children released from special education were better readers both before and after the intensive intervention, and their advantage in reading skill increased during the follow-up period. During the intervention period, the groups did not show differential growth on any of the reading measures. In contrast, during the follow-up period, significant interactions were found between time and group for all of the measures, $Fs(2, 92) = 4.2, 14.2, 9.9$, and 12.1 for the Word Attack, Word Identification, Comprehension, and Rate measures, respectively. At pretest, the groups were reliably different from one another on only the Word Identification test, but at the end of the follow-up period they were reliably different on all measures. The divergence of growth in reading skills between the groups during the follow-up period cannot be attributed to better reading instruction in the general education classroom environment, as no specific instruction in reading is provided in the upper elementary and middle school grades in this school district.

Discussion

Perhaps the most striking finding from this study was the size of the gains in reading achievement made by this sample of severely disabled readers, as

TABLE 9
Characteristics and Reading Outcomes for Children Who Were/Were Not Staffed Out of Special Education in Year Following Intensive Intervention

	Group							
	Returned to general education				Remaining in special education			
Measure	Pre	Post	1-year	2-year	Pre	Post	1-year	2-year
Word Attack	69.9	93.0	92.1*	96.7*	68.5	93.5	86.7	86.8
SD	13.5	8.5	7.5	9.3	8.4	8.3	9.8	11.2
Word Identification	71.4*	85.9*	84.9*	93.9*	65.2	78.2	77.3	79.5
SD	6.3	8.4	6.9	7.7	9.1	10.2	11.5	10.7
Gray Comp.[a]	80.5	88.6	93.4*	95.8*	74.0	83.5	88.9	81.7
SD	12.2	9.2	10.4	12.0	10.8	9.8	8.0	11.1
Gray Rate[a]	72.4	74.4	80.0*	80.3*	71.0	73.2	69.5	66.0
SD	5.9	7.6	10.5	9.0	7.9	8.6	10.1	8.8
CELF Rec. Lang.	87.3*	97.7*	99.9*	99.9*	80.3	86.7	88.5	90.1
SD	9.7	10.3	12.7	12.9	9.7	13.9	14.4	12.9
Age in months		121 (10.5)*				113.1 (10.7)		
Attention rating[b]		93.4 (9.3)*				103.2 (15.5)		
SES		41.3 (12.3)*				33.5 (11.5)		
Gender ratio		14M/5F				21M/9F		
Racial balance		14White/5Black				20White/10Black		

*This score was significantly ($p < .05$) higher than the corresponding score for the children who remained in special education.
[a]To be consistent with the other measures, standard scores from the Gray Oral Reading Test–III (Wiederholt & Bryant, 1992) were transformed to have a mean of 100 and standard deviation of 15. [b]For purposes of presentation in this table, the *z* scores for this variable were transformed to have a mean of 100 and a standard deviation of 15. High scores represent more attention/behavior problems.

well as the stability of those gains over the 2-year follow-up period. Both instructional methods produced very large alterations in the children's growth rates for broad reading ability when compared with the rate they had been growing during the previous 16 months' instruction in learning disability resource rooms. A direct comparison of growth rates on a measure of broad reading ability during the period before the intervention with those during the intervention produced effect sizes for the slopes of 4.4 for the ADD condition and 3.9 for the EP group.

We conclude that the ADD and EP instructional methods provided equally effective instruction for this sample of children because no differences existed between the groups on any of the important reading outcomes at the end of the 2-year follow-up period. Although the ADD condition did produce substantially stronger growth in accuracy of phonemic decoding skills during the intervention, these differences were not maintained during the follow-up period. The ADD group also showed significantly greater growth during treatment in the accuracy and fluency of word reading in text (GORT-III), but these differences were not large and did not maintain during the period following the intervention. The goal of any intensive intervention with children with reading disability is to produce large changes in reading ability that maintain over time, and the clear conclusion from this research is that the EP and ADD conditions were equally successful in achieving this goal. Although the ADD program produced slightly greater growth on three of our eight reading outcome measures during the intervention period, the overall pattern of growth in the treatment and follow-up periods indicates that the outcomes for the two methods were much more similar than different.

How do the growth rates obtained in this study compare with those reported by other investigators? McGuinnes et al. (1996) suggested a possible metric for reporting the effects of intensive interventions in terms of the number of standard score points gained per hour of instruction. Of course, this metric depends on the common use across studies of standardized measures that have the same standard deviation, but a number of studies have used measures similar enough to allow rough comparisons. In Table 10, we present values for this metric for measures of phonemic decoding (word attack), context-free word reading (word identification), and reading comprehension (passage comprehension). The studies by Wise et al. (1999), Lovett et al. (1994), Alexander, et al. (1991), and Truch (1994) all taught

TABLE 10
Gains in Standard Score Points Per Hour of Instruction for Three Measures of Reading Skill

		Phonemic Decoding	Word Identification	Passage Comprehension
Torgesen et al. (this study) 67.5 hours of 1:1	*ADD method*	.41	.20	.12
	EP method	.30	.21	.15
Wise et al. (1999) 40 hrs. sm grp + 1:1 computer practice		.31	.22	.14
Lovett et al. (1994) 35 hrs., 1:2		—	.13	.14
Alexander et al. (1991) 65 hrs., 1:1		.34	.23	—
Truch (1994) 80 hrs. 1:1		—	.21	—
McGuiness et al. (1994) 8 hrs., 1:1		2.57	—	1.7
Rashotte et al. (in press) 30 hours small group (4)		.50	.19	.32

Note. ADD = Auditory Discrimination in Depth; EP = Embedded Phonics.

children similar to those selected for the present study, whereas McGuiness et al. (1994) and Rashotte, MacPhee, and Torgesen (2000) worked with children of similar ages who were less severely impaired. It is apparent from Table 10 that the rates of growth we obtained in this study for phonemic decoding skills, word-reading ability, and reading comprehension are very similar to other studies of children with severe reading disabilities. The fact that we obtained larger overall gains than several other studies can be attributed to the number of hours the interventions lasted, not to substantial differences in the rate of growth among studies. The consistency in rate of gain across the first five studies in Table 7 seems remarkable, and it suggests that the high rates of growth obtained in our study should be generalizable to other settings, with other teachers implementing the interventions. Although it is true that the studies reported by Wise, Alexander, and Truch all used some variant of the Auditory Discrimination in Depth Program, the similarities in growth rate of the ADD and EP conditions in our study suggest that given the right level of intensity and teacher skill, it is possible to obtain these rates of growth via a variety of approaches to direct instruction in reading. We might even suggest that these rates could serve as a benchmark for "reasonable progress" in reading for students receiving remedial instruction in both public and private settings. As such, they are clearly much higher than is typically achieved in most current special education settings. Only one study reported growth rate values that were clearly out of range with the others (McGuiness et al., 1994) which suggests that the findings bear replication by other investigators.

Although the ADD and EP conditions both provided systematic and explicit instruction in the knowledge and strategies required for reading words accurately in print, the instructional activities and emphases within each program were substantially different. The children in the ADD condition spent 85% of their time on activities designed to stimulate phonemic awareness and build phonemic decoding skills using activities that did not involve connected text. In contrast, children in the EP group spent only 20% of their time on broadly similar activities and 50% of their time on reading and writing activities involving meaningful text. Furthermore, the ADD condition involved extensive instruction to build a basic and powerful level of phonemic awareness at the articulatory level, whereas the EP condition directed children's attention to phonemes at the auditory level as they engaged in reading and writing activities. Given these large differences in instructional emphases between the two approaches, the similarities in outcome between them are remarkable, and they stand in contrast to more substantial differences between similar instructional methods that were employed in an earlier study of the prevention of reading difficulties (Torgesen et al., 1999). In the latter study, children in the ADD condition obtained consistently higher scores on measures of phonemic decoding and word identification during the intervention, and those scores that have remained stable over a 2-year follow-up period (Torgesen, Wagner, & Rashotte, 2000).

There are a number of possible explanations for the differences in findings between these two studies, but the most likely one involves differences in levels of skill of the teachers in both studies. Whereas the present study employed highly skilled teachers who all had a number of years' experience teaching children with reading disabilities, the prevention study employed inexperienced teachers who were specifically trained for the study. In the present study, the teachers in the EP condition were able to make significant refinements to the children's phonemic awareness and word-reading strategies in the context of meaningful reading and writing activities. Their success in this area is most likely attributable to their expertise in error correction routines that guided children's attention to appropriate cues to word identity while reading, and to proper spelling strategies while writing. These error correction routines, or "scaffolded interactions," have been shown in other research (Juel, 1996) to be particularly important in stimulating growth of word-reading ability in children with reading disabilities.

Although the similarities in outcome between the ADD and EP conditions in this study were inconsistent with our earlier prevention study (Torgesen et al., 1999), they are entirely consistent with other remedial research that has examined whether instruction at the articulatory level is necessary for stimulating phonemic awareness and building phonemic decoding skills in children with severe reading disabilities. Wise et al. (1999) recently reported an investigation in which the presence or absence of instruction at the articulatory level was manipulated across methods that were carefully controlled on other important instructional dimensions. They found no significant differences in important reading outcomes across their methods that could be attributed to whether or not stimulation of articulatory awareness was part of the method. Thus, one conclusion that appears to emerge from both this and other research is that skillful instruction in phonemic awareness and phonics at the auditory level can be just as effective in building word-level reading skills as methods that work to build awareness of phonemes at the articulatory level. This is not to say that approaches such as Auditory Discrimination in Depth are not extremely effective with children who have reading disabilities. Rather, the conclusion is that there is more than one way to build phonemic awareness and word-level skills in these children.

Another striking finding from this study was the substantial difference in growth of reading accuracy as opposed to reading rate. One way to understand this difference comes from considering of the factors that contribute to reading fluency differences on measures such as the GORT-R. Perhaps the most important of these factors is the proportion of words in any given passage that can be read fluently as sight words. If too many words in a passage must be identified through phonemic analyses or guessed from context, reading rate will suffer. The children in our sample were able to improve their accuracy scores on the GORT-R because of their increased ability to apply phonemic analysis skills and other word-reading strategies while processing text. However, their reading rates remained severely limited because as the difficulty of the passages increased, they encountered an increasing percentage of words they could not recognize fluently as familiar orthographic units. This did not have a serious impact on their accuracy scores because they were able to decode these words using analytic strategies, but it did have a severe impact on their fluency scores because of the extra time it takes to analytically decode words rather than recognize them at a single glance as sight words.

This explanation highlights the challenge we face in helping older children who have been very poor readers for several years to "close the gap" in reading fluency with their normally reading peers. During the time they are allowed to remain poor readers, they miss out on an enormous amount of text exposure and word-reading practice compared to average readers (Nagy & Anderson, 1984). Because children add sight words to their lexicon primarily through reading them accurately in text a number of times (Ehri, 1998), this lack of word-reading practice places severe limits on the rate at which children with reading disabilities acquire the sight word representations that are the basis of fluent reading. Given the very low scores of the children in this study on the word identification and sight word efficiency measures at the beginning of the study, it is clear that the range of their sight word vocabulary was very limited compared to average readers'. To expect these children to entirely close the sight word gap either during the intensive intervention or in the 2 years following it is probably overly optimistic, particularly as normal readers are continuing to practice reading and thus are adding to their own sight word vocabularies at a high rate. In other words, the size of normally achieving children's sight word vocabularies is a rapidly moving target during the late elementary and middle school years. To close the gap in reading fluency during these years, our children would have to add sight words to their lexicons at a faster rate than nondisabled children, and that implies that they would be practicing reading *more* than their normally reading peers. To engineer conditions in which children for whom reading is still somewhat more difficult than average are willing to read more than average over an extended period of time is clearly a daunting task.

One piece of evidence that is consistent with a "practice" explanation of the fluency difficulties observed in this study is that we did not obtain the same disparity between growth of accuracy and rate of text reading in our studies of the prevention of reading disabilities. In one study, in which highly at-risk children received preventive instruction from kindergarten through second grade (Torgesen et al., 1999), and in another study, in which at-risk children received intensive instruction during first grade (Torgesen, Wagner, Rashotte, & Herron, 2000), only very small differences in standard scores were found for accuracy versus rate measures, either at the end of the intervention or at the end of a 1- or 2-year follow-up period. Apparently, if the word-reading skills of children at risk for reading failure can be maintained at normal levels during elementary school, their sight word vocabularies expand at normal rates and they do not develop serious discrepancies between reading accuracy and reading fluency.

Another factor that may have limited reading rate particularly on the GORT-R test was the emphasis both instructional methods placed on accuracy of word identification as a priority during reading. Given also the fact that the children knew they would be asked comprehension questions after reading the paragraphs, they may have artificially slowed their word-reading rates in order to maximize accuracy of both word reading and comprehension. In contrast, on the Sight Word Efficiency subtest of the TOWRE, on which the children obtained higher fluency standard scores than in the GORT-R, the children were specifically instructed to read as many words as they could during the time allotted. Thus, the difference in rate scores for the two types of measures may have arisen partially because instructions for the two tasks led children to make different choices in the trade-off between speed and accuracy in word identification.

A third potential explanation of the difference between outcomes for rate versus accuracy measures can be found in the comparison of final levels of performance on measures of phonemic awareness (an accuracy measure) and rapid automatic naming (a rate measure) at the end of the intervention period. The average standard score for the measure of phonemic awareness (Elision) was 99.4, whereas that for the rate measures was 91.8, with an effect size of .65. Thus, one factor that may have contributed to the enduring differences between accuracy and rate of word reading in the children in this study was fundamental limitations in processing rate for some of the children. Consistent with this explanation are the recurrent associations between measures of rapid automatic naming and measures of reading rate reported in the literature on reading disabilities (Wolf & Bowers, 1999), as well as their relationships in this study. For example, the concurrent correlations between naming measures and reading rate measures at immediate posttest for the sight word efficiency and Gort-III rate measures were .52 and .18, respectively, whereas analogous relationships at the 2-year follow-up point were .44 and .28. That naming measures were more highly correlated with single-word–reading rate than text-reading rate suggests that reading rate on the GORT-III reflected a more complex combination of factors than simple processing rate for individual words.

Another notable effect of the interventions was the growth they produced in receptive and expressive language skills. The children began the study with very low scores on these measures (average Total Language score = 78.6), but by the end of the 2-year follow-up period, their scores were very close to the average range (89.8). Most of this growth took place during the intervention, but there was also positive change in standard scores on both expressive and receptive measures during the follow-up period. The most likely explanation for the improvement during the intervention was the density of language stimulation occurring during teacher–student interactions. Both methods involved extensive teacher–student discussions, which frequently required the student to justify or explain answers to questions. Although the focus and content of these interactions were different between methods, they both were designed to increase students' awareness and control of their own cognitive processing during reading. The richness of the language interchange during this kind of instruction had a powerful, and unanticipated, effect on the children's ability to think about and respond to questions in the more general receptive and expressive language domains. Although growth in these skills was not as impressive during the follow-up period, the fact that growth slightly exceeded the normal growth rate might be attributed to a change in reading habits of many of the children in the sample. Longitudinal studies generally indicate that children with reading disabilities tend to show a relative decline, rather than an increase, in their verbal skills during the elementary and middle school period (Cunningham & Stanovich, 1998).

A final point about group-level effects involves the finding that the interventions produced significant changes in standard scores for the measures of phonological memory and rapid automatic naming. We are not aware of other studies in which significant improvements on these more basic measures of phonological ability have been shown to occur in the context of reading instruction that does not involve specific training in these areas. Although the children received fluency practice for recognizing sight words, neither of the interventions contained fluency-oriented practice at the phonemic level. At present, the most obvious explanation for the improvement on these measures involves either increased ability to focus and sustain attention, or the development of more precise or distinctive representations for phonemic elements in speech (Metsala & Walley, 1998). If the quality of children's phonemic representations developed as a result of the interventions, this could conceivably make it easier to access names of numbers and letters from long-term memory on the naming tasks and to remember digits and sequences of phonemes on the phonological memory tasks. The fact that children improved more on the nonword repetition task, on which they were required to repeat novel sequences of phonemes, than on the digit span task is consistent with this explanation. That the changes in performance were stable over the 2-year follow-up period argues in favor of a fundamental change in the quality of phonemic representation rather than a perhaps more transitory change in attentional habits.

An analysis of individual differences in response to the interventions used in this study indicated that they were insufficient to "normalize" the reading accuracy of approximately one third to one half the children in the sample, depending on the specific outcome measure being used. So, even in the context of very large average gains for the group as a whole, substantial numbers of the children with severe disabilities in this sample remained poor readers following the intervention. Because the children in this study fell within the bottom 2% of all children in terms of reading ability, we can estimate that if the interventions used in this study were available to all children with reading disabilities, only about 1% of children would remain severely impaired in reading accuracy upon leaving elementary school.

The growth trajectories presented in Figure 2 suggest that the ultimate outcomes from the intervention depended more on what happened during the 2-year follow-up period than during the intensive intervention itself. That is, those children who ended up with the lowest scores at the end of the follow-up period showed substantially different growth rates during the follow-up period than those who ended up with the highest scores, while the growth rates for these groups during the treatment period were much more similar. Unfortunately, we were not able to closely monitor the children's reading instruction or practice once they finished the intensive intervention, so it is not possible to specify all the factors associated with the differences in growth rate during the follow-up period. However, our analysis of growth using predictors from the pretest and immediate posttest shed a small amount of light on this question.

During the follow-up period, the most consistent predictors of growth were the children's attentional/behavior ratings by resource room teachers, their general verbal ability, and the levels they had attained on component reading skills. For example, approximately 30% of the variance in growth on the reading comprehension measure during the follow-up period was accounted for by the children's posttest scores on the comprehension and word identification measures, along with their verbal ability and behavior ratings. These same predictors, plus level of phonemic awareness at immediate posttest, accounted for fully half of the variance in children's ultimate scores on the comprehension measure at the end of the follow-up period.

A much smaller range of variables predicted growth and outcomes during the intervention period than during the follow-up period. This was at least partially due to the smaller variability among children in their growth during the intervention period; however, for two of the outcome measures (word attack and comprehension) it was also related to the strong negative relationship between entering level of skill and growth. For these measures, children who entered the study with the weakest skills tended to show the largest gains in their standard scores. One way to think about this outcome is in terms of the quality of instruction prior to entry into the study. Children who showed very large gains from pre- to posttest may have been those who received particularly weak instruction in basic reading skills, or for some reason were not able to profit from instruction in group settings. When exposed to the powerful one-to-one instruction in this study, they were able to particularly profit from the instruction. At the same time, however, the positive relationships between pretest and posttest scores on these measures indicate that the instruction we provided was not sufficient to eliminate the overall advantage in reading skill enjoyed by those who entered the study with the strongest skills.

The broad influence of general receptive language skills on growth and outcomes in reading was expected for reading comprehension but not for word-level reading skills. However, the data in Tables 5 and 7 indicate that children with higher scores on the receptive language measure did tend to show stronger growth on the word identification and rate measures during the follow-up period, and stronger outcomes on these measures during both the intervention and the follow-up period. Although general verbal ability (receptive language) did not contribute uniquely to the prediction of growth on either comprehension or word-level skills during the follow-up period, this was due to its significant relationships with other predictor variables in the simultaneous regressions, particularly with the autoregressors. Thus, we would argue that this study is consistent with other research with older children (Wise et al., 1999) in suggesting that general verbal ability does contribute to the ease with which children who have been instructed well in basic phonemic decoding skills can add words to their sight vocabularies.

Another way in which the findings from this study are consistent with those from Wise et al. (1999) is in the lack of differential outcomes across the two instructional conditions for children with different entering levels of phonemic awareness and phonemic decoding skills. In other words, neither study found that an instructional method involving stimulation of phonemic awareness at the articulatory level was *particularly* beneficial to children with the weakest phonemic decoding skills. Although our sample size provided limited power to detect significant aptitude/treatment interactions by traditional methods, there was not even a reasonable suggestion of such effects in the extreme group analyses described in Figure 3. Although the ADD condition produced stronger overall effects on phonemic decoding skills than the EP condition, the advantages of the former were similar for children entering the study with the strongest and weakest skills in this area.

Our findings concerning the factors related to growth in reading were also *inconsistent* with those reported by Wise et al. (1999) in one important way: Whereas Wise et al. found that younger children showed stronger gains than older children, we found no important outcome differences among our children that were related to differences in age. The one significant relationship between growth and age reported in Table 5 indicates that the older children tended to improve more in reading rate during the follow-up period than the younger children. The children in Wise et al.'s study covered a slightly larger age range than those in the present study (second through fifth grade vs. third through fifth grade), and they were also about two thirds of a standard deviation less impaired on word-level reading skills than the children we studied. Thus, as Wise and her colleagues suggested, because many of the older children in the study were reading above the third-grade level, the phonemically explicit instruction those authors provided may have been better matched to the instructional needs of the younger than the older children. Other remedial studies that have found no differences in growth related to age and covered a similar age range of children were reported by Lovett and Steinbach (1997) and by Rashotte et al. (2000).

Summary and Implications

The most significant outcome of this study was the demonstration of large, generalized, and stable changes in the reading ability of a sample of children selected because they had been unable to acquire adequate word-reading ability through instruction received in both general and special education classrooms. About half the children in our sample attained average-level reading skills by the end of the follow-up period. Given this demonstration of the power of intensive instruction, we would assert that one major task for the educational establishment is to find ways to deliver both the quality and the intensity of instruction that many children seem to require. This task clearly extends beyond the educators who are charged with providing direct instructional services. The larger political community served by our schools must also become involved to identify the additional resources that are needed. However, the finding that 40% of the children were returned full time to the general education classroom within 1 year following the intervention suggests that there may actually be some significant economies associated with increased quality and intensity of instruction for these children.

The fact that reading rate remained so impaired in most of the children suggests the need for further experimentation to develop interventions in this area, as well as the need to focus resources on the prevention of reading difficulties. If future studies, as well as longer term follow-up of early interventions, verify that reading rate problems can be eliminated for almost all children via preventive instruction (Torgesen et al., 1999), this would clearly be another powerful argument in support of more extensive preventive efforts with young children. However, it may also be possible to improve fluency outcomes in future studies, through more careful fluency-oriented training and practice than was provided in this study. We, and others, are trying to address some of these questions in the next round of intervention studies.

Another extremely important finding from this study was the fact that two quite different intensive instructional interventions produced essentially the same long-term outcomes for the children. It is important to emphasize that both of these interventions incorporated principles of instruction that have been shown to be particularly important for children with severe reading disabilities. That is, they both involved explicit instruction in phonemic awareness, phonemic decoding, and sight word recognition skills. They also included mastery-oriented progressions through essential skills, extensively scaffolded error-correction routines to establish appropriate word identification strategies, and many opportunities to practice with appropriate materials. However, the ADD and EP approaches varied substantially in their depths of instruction in phonemic awareness and in their emphases on various instructional activities. The fact that outcomes were so similar for the two methods suggests that within explicit "structured language" approaches that follow sound instructional principles, there may be considerable latitude for arranging components of instruction according to teacher and student preferences.

What neither intervention approach contained was a systematic approach to teaching reading comprehension strategies. Because other research (Mastropieri & Scruggs, 1997; Swanson, 1999) has consistently shown that large gains in comprehension result from such instruction, our next intensive intervention study will include a substantial amount of instruction specifically focused on helping children acquire and utilize effective comprehension strategies. Our hope is that this additional instruction will help even those children who remained relatively weak in their word-reading skill to improve their comprehension more substantially than in the present study.

ABOUT THE AUTHORS

Joseph K. Torgesen, *PhD, is a distinguished research professor of psychology and education at Florida State University.* ***Ann W. Alexander***, *MD, is a developmental pediatrician and director of the Morris Child Development Center in Gainesville, Florida.* ***Richard K. Wagner***, *PhD, is the Alfred Binet Professor of Psychology at Florida State University, and* ***Carol A. Rashotte***, *PhD, is a senior research associate in the Department of Psychology at the same institution.* ***Elaine Rose***, *MS, is an academic specialist and director of the Rose and Speech and Academic Center in Tallahassee, Florida.* ***Kytja K. S. Voeller***, *MD, is a professor of neurology in psychiatry at the University of Florida, and* ***Tim Conway***, *MS, is a PhD candidate in clinical psychology at the University of Florida. Address: Joseph K. Torgesen, Psychology Department, Florida State University, Tallahassee, FL 32306-1270 (e-mail: torgesen@psy.fsu.eduAbstract).*

AUTHORS' NOTE

The research reported in this manuscript was supported by Grant No. HD30988 from the National Institute of Child Health and Human Development, and by grants from the National Center for Learning Disabilities and the Donald D. Hammill Foundation.

REFERENCES

Agronin, M. E., Holahan, J. M., Shaywitz, B. A., & Shaywitz, S. E. (1992). The multigrade inventory for teachers. In S. E. Shaywitz & B. A. Shaywitz (Eds.) *Attention deficit disorder comes of age* (pp. 29–67). Austin, TX: PRO-ED.

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

Allington, R. L., & Woodside-Jiron, H. (1999). The politics of literacy teaching: How "research" shaped educational policy. *Educational Researcher, 28,* 4–13.

Atkins, M. S., Pelham, W. E., & Licht, M. (1985). A comparison of objective classroom measures and teacher ratings of attention deficit disorder. *Journal of Abnormal Child Psychology, 13,* 155–167.

Bowers, P. G., Golden, J., Kennedy, A., & Young, A. (1994). Limits upon orthographic knowledge due to processes indexed by naming speed. In V.W. Berninger (Ed.), *The varieties of orthographic knowledge: Theoretical and developmental issues* (pp. 173–218). Dordrecht, The Netherlands: Kluwer Academic.

Briggs, A., Austin, R., & Underwood, G. (1984). Phonological coding in good and poor readers. *Reading Research Quarterly, 20,* 54–16.

Brown, A. L., Palincsar, A. S., & Purcell, L. (1986). Poor readers: Teach, don't label. In U. Neisser (Ed.), *The school achievement of minority children: New perspectives* (pp. 105–143). Hillsdale, NJ: Erlbaum.

Bruck, M. (1990). Word-recognition skills of adults with childhood diagnoses of dyslexia. *Developmental Psychology, 26,* 439–454.

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

Cunningham, A. E., & Stanovich, K. E. (1998, Spring/Summer). What reading does for the mind. *American Educator, 22,* 8–15.

Ehri, L. C. (1998). Grapheme–phoneme knowledge is essential for learning to read words in English. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning reading* (pp. 3–40). Hillsdale, NJ: Erlbaum.

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

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90,* 37–55.

Fry, E. B., Kress, J. E., & Fountoukidis, D. L. (1993). *The reading teacher's book of lists: Third edition.* Englewood Cliffs, NJ: Prentice Hall.

Hannah, S. (1993). *Early literacy series, I & II.* Edmonton, Alberta Canada: The Centre for Literacy.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). *Does special education raise academic achievement for students with disabilities?* National Bureau of Economic Research, Working Paper No. 6690, Cambridge, MA.

Hollingshead, A. B. (1975). *Four factor index of social status.* New Haven, CT: Yale University, Department of Sociology.

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 31,* 268–289.

Kaufman, A. S., & Kaufman, N. L. (1985). *Kaufman test of educational achievement.* Circle Pines, MN: American Guidance Service.

Kavale, K. A., (1988). The long-term consequences of learning disabilities. In M. C. Wang, H. J. Walburg, & M. C. Reynolds (Eds.), *The handbook of special education: Research and practice* (pp. 303–344). New York: Pergamon.

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1–33). Ann Arbor: University of Michigan Press.

Lindamood, C. H., & Lindamood, P. C. (1979). *Lindamood auditory conceptualization test.* Austin, TX: PRO-ED.

Lindamood, C. H., & Lindamood, P. C. (1984). *Auditory discrimination in depth.* Austin, TX: PRO-ED.

Lindamood, P., & Lindamood, P. (1998). *The Lindamood phoneme sequencing program for reading, spelling, and speech,* Austin, TX: PRO-ED.

Loney, J., & Milich, R. (1982). Hyperactivity, inattention, and aggression in clinical practice. In M. Wolraich & D. D. Routh (Eds.), *Advances in developmental and behavioral pediatrics* (Vol. 3, pp. 1213–2147). Greenwich, CT: JAI Press.

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

Lovett, M. W., & Steinbach, K. A. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children? *Learning Disability Quarterly, 20,* 189–210.

Lyon, G. R. (1995). Towards a definition of dyslexia. *Annals of Dyslexia, 45,* 3–27.

Lyon, G. R., Alexander, D., & Yaffee, S. (1997). Progress and promise in research in learning disabilities. *Learning Disabilities: A Multidisciplinary Journal, 8,* 1–6.

Manis, F. R., Custodio, R., & Szeszulski, P. A. (1993). Development of phonological and orthographic skill: A 2-year longitudinal study of dyslexic children. *Journal of Experimental Child Psychology, 56,* 64–86.

Mastropieri, M. A., & Scruggs, T. E. (1997). Best practices in promoting reading comprehension in students with learning disabilities: 1976–1996. *Remedial and Special Education, 18,* 197–213.

McGuiness, C., McGuiness, D., & McGuiness, G. (1996). Phono-Graphix: A new method for remediating reading difficulties. *Annals of Dyslexia, 46,* 73–96.

McKinney, J. D. (1990). Longitudinal research on the behavioral characteristics of children with learning disabilities. In J. Torgesen (Ed.), *Cognitive and behavioral characteristics of children with learning disabilities* (pp. 115–138). Austin, TX: PRO-ED.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in be-*

ginning literacy (pp. 89–120). Mahwah, NJ: Erlbaum.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19,* 304–330.

Olson, R. K., Wise, B., Johnson, M., & Ring, J. (1997). The etiology and remediation of phonologically based word recognition and spelling disabilities: Are phonological deficits the "hole" story? In B. Blachman (Ed.), *Foundations of reading acquisition* (pp. 305–326). Mahwah, NJ: Erlbaum.

Rashotte, C. A., MacPhee, K., Torgesen, J. K., & Deacon, C. (in press). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly.*

Schulte, A. C. (1996). Remediation and inclusion: Can we have it all? In D. L. Speece & B. K. Keogh (Eds.), *Research on classroom ecologies* (pp. 203–210). Mahwah, NJ: Erlbaum.

Schumaker, J. B., Deshler, D. D., & Ellis, E. S. (1986). Intervention issues related to the education of learning disabled adolescents. In J. K. Torgesen & B. Y. L. Wong (Eds.), *Psychological and educational perspectives on learning disabilities* (pp. 329–365). New York: Academic Press.

Semel, E., Wiig, E. H., & Secord, W. (1995). *Clinical evaluation of language fundamentals–3.* San Antonio, TX: Psychological Corp.

Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: A model of acquisition and individual differences. *Issues in Education: Contributions from Educational Psychology, 1,* 1–57.

Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22,* 469–479.

Simpson, G. B., Lorsbach, T., & Whitehouse, D. (1983). Encoding and contextual components of word recognition in good and poor readers. *Journal of Experimental Child Psychology, 35,* 161–171.

Smith, L. (1992). *Auditory discrimination reading series–sequence B.* Independence, MO: Poppin Creations.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities, 21,* 590–604.

Stanovich, K. E., & Siegel, L. S. (1994). The phenotypic performance profile of reading-disabled children: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86,* 24–53.

Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities, 32,* 504–532.

Torgesen, J. K. (1996). A model of memory from an information processing perspective: The special case of phonological memory. In G. Reid Lyon (Ed.), *Attention, memory, and executive function: Issues in conceptualization and measurement* (pp. 157–184). Baltimore: Brookes.

Torgesen, J. K. (1998a). Instructional interventions for children with reading disabilities. In S. Shapiro, D. Accardo, & C. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 197–220). Parkton, MD: York Press.

Torgesen, J. K. (1998b). Assessment and instruction for phonemic awareness and word reading skills. In H. W. Catts & A. G. Kamhi (Eds.), *Language basis of reading disabilities* (pp. 128–153). Needham Heights, MA: Allyn & Bacon.

Torgesen, J. K. (1999). Phonologically based reading disabilities: Toward an integrated theory of one kind of learning disability. In L. Spear-Swerling & R. J. Sternberg (Eds.), *Perspectives on learning disabilities* (pp. 106–135). New Haven, CT: Westview Press.

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

Torgesen, J. K., Rashotte, C. A., & Alexander, A. (in press). The prevention and remediation of reading fluency problems. In M. A. Wolf (Ed.), *Time, fluency, and dyslexia.* Timonium, MD: York Press.

Torgesen, J. K., & Wagner, R. K. (1998). Alternative diagnostic approaches for specific developmental reading disabilities. *Learning Disabilities Research and Practice, 13,* 220–232.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency.* Austin, TX: PRO-ED.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2000). *Reading achievement outcomes two years following an intensive early intervention.* Unpublished manuscript, Florida State University, Tallahassee.

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: An Interdisciplinary Journal, 8,* 51–62.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., & Herron, J. (2000). *The effectiveness of teacher supported computer assisted instruction in preventing reading problems in young children: A comparison of two methods.* Unpublished manuscript, Florida State University, Tallahassee.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvin, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91,* 579–593.

Truch, S. (1994). Stimulating basic reading processes using Auditory Discrimination in Depth. *Annals of Dyslexia, 44,* 60–80.

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

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processes.* Austin, TX: PRO-ED.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing causal relations between phonological processing abilities and word-level reading as children develop from beginning to fluent readers: A five-year longitudinal study. *Developmental Psychology, 33,* 468–479.

Wechsler, D. (1974). *Wechsler intelligence scale for children: Revised.* San Antonio, TX: Psychological Corp.

Wiederholt, J. L., & Bryant, B. R. (1992). *Gray oral reading tests–III.* Austin, TX: PRO-ED.

Wise, B.W., Ring, J., & Olson, R. K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology, 72,* 271–304.

(continued on p. 78)