

Assessment and Early Instruction of Preschool Children at Risk for Reading Disability

Barbara Hindson, Brian Byrne,
Ruth Fielding-Barnsley, Cara Newman,
and Donald W. Hine
University of New England

Donald Shankweiler
University of Connecticut and Haskins Laboratories

Preschool children at familial risk for reading disability were assessed on cognitive and linguistic variables and compared with preschoolers without familial risk. Risk children displayed performance profiles resembling those of older children with reading disability. Each group received intensive instruction in phonemic awareness and structured book reading. Instructed risk children made somewhat smaller gains than the nonrisk and required more teaching sessions to reach criterion. Rhyme and phoneme awareness predicted instruction outcome levels, and vocabulary and verbal short-term memory predicted number of teaching sessions to criterion. In kindergarten, the nonrisk group outperformed the risk group on reading and spelling, although the risk group reached grade-appropriate levels. At-risk children can be helped by appropriate preschool instruction, but they require more sustained teaching than nonrisk preschoolers.

Keywords: reading disability prevention, familial risk, preschool intervention, phonological awareness, pre-reading assessment

Developmental reading disability is one of the most common of the childhood disorders and also one of the most unremitting. Follow-up studies of disabled readers during childhood (e.g., Satz, Taylor, Friel, & Fletcher, 1978; Waring, Prior, Sanson, & Smart, 1996) and from childhood to adolescence (e.g., Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Shaywitz et al., 1999) all have reported the strong persistence of reading problems during the school years. Against this background, the prompt identification of children likely to develop reading difficulties becomes important because evidence indicates that early intervention can enhance the prospects for literacy growth (Bus & van IJzendoorn, 1999; Byrne, Fielding-Barnsley, & Ashley, 2000). A more complete picture of the cognitive and behavioral characteristics of children who are at risk is also needed to better design and deliver appropriate interventions. In addition, the observation that many children fail to fully benefit from remediation indicates that the

intervention process needs to be understood more comprehensively. A picture of the characteristics that predict responsiveness to early intervention is needed, as is tracking of the short- and long-term efficacy of early intervention.

In this article, we address aspects of each of these issues. We report a longitudinal intervention study of preschool children carrying a familial risk for reading disability. We recruited three successive cohorts of family-at-risk (FAR) children in their final preschool year and, in the first phase of the project, compared them on a variety of cognitive, linguistic, and behavioral measures with children deemed not to be at similar family risk (nonFAR). A subset of both types of children received a more intensive battery of tests exploring language processing in greater detail. Variables were selected from those known to be deficient in older children with established reading difficulties and from others known to predict subsequent reading ability when assessed at the preschool stage. The children then received individualized intervention designed to provide foundation skills for reading development, with pre- and posttesting of these skills and some related ones. A smaller wait-list control group of FAR children was also recruited to determine the effectiveness of the preschool intervention for at-risk children. In this article, we report the results of the preintervention measures and of the intervention itself and outline our findings about the predictors of responsiveness to the intervention. We also report follow-up data from the children's kindergarten year.

FAMILY RISK FACTOR

It has been recognized for a century that reading disability tends to run in families (Thomas, 1905). Since that time, many studies have confirmed familial transmission of reading disability. Scarborough (1989) summarized eight of these and reported estimates

Barbara Hindson, Brian Byrne, Ruth Fielding-Barnsley, Cara Newman, and Donald W. Hine, School of Psychology, University of New England, Armidale, New South Wales, Australia; Donald Shankweiler, Department of Psychology, University of Connecticut, and Haskins Laboratories, New Haven, Connecticut.

This research was funded by the Australian Research Council Grant A79602700. The participation of Donald Shankweiler was supported in part by National Institute of Child Health and Human Development Program Grant HD-01994 to Haskins Laboratories. We are grateful to both agencies for support.

We acknowledge the contributions of the many children and their parents involved in the project and of the preschool staff who assisted us.

Correspondence concerning this article should be addressed to Brian Byrne, School of Psychology, University of New England, Armidale, New South Wales 2351, Australia. E-mail: bbyrne@pobox.une.edu.au

of reading disability in offspring of families with affected parents ranging from 23.0% to 62.0%, with an average of 38.5%. The most recent studies by Pennington and Lefly (2001) and Snowling, Gallagher, and Frith (2003) reported estimates of 34.0% and 66.0%, respectively.

The familial nature of reading disability is consistent with reading disability as a heritable disorder, and family studies of reading disability have confirmed a role for genetic endowment (see Olson & Byrne, 2005, and Pennington & Olson, 2005, for reviews). The behavior-genetic evidence for heritability of reading disability is complemented by direct studies at the molecular level. Regions of Chromosomes 2, 3, 6, 15, and 18 have been implicated in dyslexia by linkage (S. E. Fisher & DeFries, 2002), and recently, evidence for a susceptibility gene on Chromosome 6 has been found (Cope et al., 2005). The mechanisms mediating between genotype and phenotype are as yet unknown.

SELECTION OF VARIABLES FOR INCLUSION IN THE PREINTERVENTION ASSESSMENT

Many of the variables that we elected to assess prior to the instructional phase were already known to be predictors of reading ability (when administered at preschool or close to school entry). They were included to see how well they predicted responsiveness to the intervention. This list includes processes close to reading such as letter knowledge and other measures of print familiarity (Byrne, Fielding-Barnsley, Ashley, & Larsen, 1997; Gallagher, Frith, & Snowling, 2000; Scarborough, 1998) and phonemic awareness (Byrne et al., 1997; Elbro, Borstrom, & Petersen, 1998; Scarborough, 1998). It also includes various aspects of memory for phonetically based material such as digit span and word repetition (Elbro et al., 1998; Gallagher et al., 2000), naming speed and accuracy, rhyming ability, and vocabulary (Byrne et al., 1997; Lefly & Pennington, 1996; Locke et al., 1997; Scarborough, 1998). Measures of syntax have also been associated with risk status or subsequent reading outcomes (Gallagher et al., 2000; Locke et al., 1997; P. Lyytinen, Poikkeus, Laakso, Eklund, & Lyytinen, 2001; Scarborough, 1998), although some researchers have raised questions about the source of apparent syntactic deficits in older reading-disabled children, preferring to attribute these deficits to processing limitations rather than to genuine gaps in syntax (Shankweiler & Crain, 1986; Shankweiler, Crain, Brady, & Macaruso, 1992). This issue is also relevant to younger at-risk children.

We elected to include some variables about which evidence is more equivocal. One is a measure of speech perception. Although some studies failed to find any differences between children from reading-disabled and nondisabled families (Lefly & Pennington, 1996; Scarborough, 1990), others found them (Elbro et al., 1998; Leppänen et al., 2003; Locke et al., 1997; Van der Leij, Lyytinen, & Zwarts, 2001; see also H. Lyytinen et al., 2004, for an extensive review). Differences in measures or in samples may have contributed to the varying outcomes. The evidence on speech production is also equivocal. There are studies that failed to detect differences (e.g., Locke et al., 1997) and others that found them (Elbro et al., 1998; P. Lyytinen et al., 2001; Scarborough, 1998). Using samples of natural language production, Scarborough (1998) found that at 30 months of age, children from reading-disabled families were less accurate than controls in their pronunciation of consonants and that phonological production abilities at this age were strongly

predictive of outcome reading status in the sample. Lyytinen et al. (2004) observed shorter utterances in children at risk than in controls. As with perception, methodological and sample differences could account for the apparent inconsistencies.

We also included two variables not so far studied in preschool children. Word identification point was included because it is believed to discriminate older reading-disabled from nondisabled readers (Metsala, 1997; Walley, 1993). Articulation rate was included because of its close connection with verbal working memory (Hulme, Thomson, Muir, & Lawrence, 1984) and because of the many observations that rapid naming is compromised in the reading disabled (see Wolf & Katzir-Cohen, 2001, for a review).

We assessed nonverbal abilities (Block Design, Raven's Colored Progressive Matrices, and visual matching) and we collected parent and teacher ratings of temperament because of the pervasive evidence of comorbidity between reading skill and aspects of the attention-deficit/hyperactivity syndrome, including evidence for pleiotropy at a quantitative trait locus on Chromosome 6p (Willcutt et al., 2002).

CLASSIFICATION OF FAMILIES

The classification of families into risk and nonrisk categories involved issues of judgment, for several reasons. One of these is that there is evidence that risk status is continuous rather than categorical (Snowling et al., 2003), as might be expected from a characteristic with multiple genetic sources, each possibly of small or modest effect (S. E. Fisher & DeFries, 2002). Second, the basis for classification is open to debate. Relying on self-report runs the risk of misclassification: Unpublished evidence from Scarborough indicated that 46.8% of the adults with self-reported reading problems were not identified as disabled readers by test-based classification methods and that 25.4% of the adults who were classified by tests as reading disabled had not also self-reported. On the other hand, ignoring self-report may exclude from risk status individuals who, although genetically compromised, have been successfully remediated. Third, if test results are used, the cutoff for risk is somewhat arbitrary—too strict and genuine cases are excluded, too lenient and the proportion of false positives increases. Thus, we considered families on a case-by-case basis, although relying primarily on testing, as detailed below.

Because there are no Australian norms for adults on reading tests as far as we are aware, we conducted a small-scale norming study on the tests we elected to use, the Word Identification, Word Attack, and Passage Comprehension subtests from the Woodcock Reading Mastery Tests—Revised (WRMT-R; Woodcock, 1987) and the Fast Reading subtest from the Stanford Diagnostic Reading Test (Karsen & Gardner, 1984). The Word Attack and Fast Reading tests were particularly useful because both nonword decoding, with its basis in phonemic awareness, and rate can remain diagnostic of earlier reading problems after a degree of remediation (Bruck, 1990; Carver, 1997; Shankweiler et al., 1999; Wimmer, Mayringer, & Landerl, 1998).

We assembled a convenience sample of 198 adults (114 women and 84 men); mean age was 35 years 4 months (range: 16 years 10 months to 70 years 7 months, $SD = 12$ years 4 months). Years of schooling completed ranged from 8 to 12 ($M = 11.18$ years). Occupational status ranged from professionals such as lawyers (3.5%) to sales staff and machine operators (each 9.6%), with

approximately equal numbers of intermediate status occupational classes of paraprofessionals such as nurses, tradespersons, and clerical workers (each around 25.0%). These classifications are based on the Australian Standard Classification of Occupations (ASCO; Castles, 1989).

Age and years of schooling, but not sex, accounted for significant and separable variance in the dependent measures, a total of 29.8% for the three WRMT-R tests treated as a composite, 16.4% for Word Attack considered alone, and 19.4% for Fast Reading. In the majority of cases in which we used test results for classification, a self-report of reading difficulties and a WRMT-R composite, Fast Reading, or Word Attack score 1.0 *SE* or more below expectation based on age and school years in at least one parent served to classify the family as FAR, a criterion identical in magnitude to that used by Snowling et al. (2003). If scores on the WRMT-R composite and Fast Reading tests were both above -1.0 *SE* and the Word Attack was not below -1.0 *SE*, the person was classified as reading disabled if both the former scores were below -0.5 *SE*. This criterion was invoked in four cases. In all other cases in which results were available, we classified the families as nonFAR.

In six families, no objective test results were available because of absence or refusal of the identified parent. In each case, the available information pointed to severe reading difficulties, often accompanied by problems in other family members. We included these in the FAR group.

We used the same criteria to classify families not self-declaring as containing a reading-disabled parent, resulting in a small number being classed as FAR. In some cases, we were not able to test both parents. In these cases, we elected to classify the child as nonFAR if the tested parent was classified as nondisabled and the other parent either self-reported as having no known present or earlier reading problems or was reported as such by the tested parent. There were 17 such cases out of a total of 68 families classified as nonFAR.

We could have used stricter criteria for classification, although the ones we did use resulted in nonoverlapping samples of reasonable size. As is shown in the results below, the substantial differences between FAR and nonFAR children on many of the variables assessed tend to confirm that we had formed groups clearly distinguished by a reading risk factor.

PHASE 1: PREINTERVENTION ASSESSMENT

In this section, we report on the preintervention assessment of the two groups of children, those with a family history of reading difficulties and those without. We recruited successive cohorts of preschoolers over a 3-year period. A subset of 19 of the at-risk children formed a wait-list group for the intervention, but in this section, we report their initial scores as part of the at-risk group as a whole. At the request of their parents, another small subset of 6 at-risk children received only initial assessment—they did not participate in the intervention. These are also included in this section.

Method

Participants

Families

The participating families were residents of Sydney, Australia, recruited through advertisements and notices placed in schools and preschools. Only

families in which English was the first language for both parents and children were included. We were able to classify 166 families, 98 as at risk and 68 as not at risk, under the criteria reported earlier. Of the at-risk families, 40 were identified because the mother was classified as reading disabled on the tests, 34 because of the father, and 24 because both parents had reading problems. In 3 of the 98 at-risk families, 2 children participated in the project, recruited in separate cohorts, for a total of 101 at-risk children.

Parental occupational status was assessed using the second edition of the ASCO rating scale (Castles, 1994). It classifies occupations into nine major groups: managers and administrators (Level 1); professionals; paraprofessionals; tradespersons and related workers; advanced clerical and service workers; intermediate clerical, sales, and service workers; intermediate production and transport workers; elementary clerical, sales, and service workers; and laborers and related workers (Level 9). Both parents' occupations were assessed in two-parent families, and whichever was the higher was taken as the family's overall occupational status. In one-parent families, the occupation of the residing parent was used. Among the families classified as at risk, 26.7% of households had no employed parent, mostly single-parent households. The analogous figure for families classified as not at risk was 13.2%.

In both groups, the full range of occupations was represented, although the mean value of 5.69 for the at-risk parents represents lower occupational status than the mean value of 4.33 for the control parents, Mann-Whitney $U = 1,645.5$, $p = .01$, $\chi^2(1, N = 166)$ approximation = 6.00. This difference, combined with the greater proportion of at-risk families with no employed parent, points to the leaner economic prospects that accompany low levels of reading in adults. Despite this, in both groups, there was a wide spread of occupational levels.

Children

In this phase, we report on a total of 169 children. In the at-risk group, there were 53 boys and 48 girls, total 101, mean age at initial testing 54.6 months. In the not-at-risk group, there were 27 boys and 41 girls, total 68, mean age 55.5 months. The age difference was not significant, $t(167) = 1.07$, $p = .29$, nor was the sex imbalance, $\chi^2(1, N = 169) = 2.68$, $p > .05$. The subsample of children receiving the more extensive battery of assessments comprised 49 at-risk children (27 boys and 22 girls, mean age 55.0 months) and 41 controls (13 boys and 28 girls, mean age 55.9 months). In this case, the sex imbalance was significant, $\chi^2(1, N = 90) = 4.91$, $p < .05$, although, as shown in Results, below, there were very few sex effects evident in the data as a whole.

Materials

Reliability and Validity

For tests that were administered pre- and postintervention, we report the correlations for the two occasions as lower bound estimates of test-retest reliability, lower bound because of the time between test occasions, around 3 months (see Method for Phase 2, below), and because some children responded better to the intervention than others (see Results and Discussion for Phase 2, below). We supplemented these values with a measure of internal consistency, Cronbach's alpha, where appropriate. In the light of generally acceptable reliabilities for the tests that we created or modified specially for this project, we report published reliabilities for commonly used tests such as the Wechsler Preschool and Primary Scale of Intelligence—Revised (WPPSI-R; Wechsler, 1989).

Tests Administered to All Children

Phoneme identity. Phoneme identity was measured using a 20-item test developed by Byrne and Fielding-Barnsley (1993b). In this test, the child was required to indicate which of two orally presented words (e.g., *beak* or

pool) started the same as a target word (*pig*). All of the words were pictured, with the target word at the top of the stimulus page and the other two words beneath it. Before the test began, the examiner administered three practice items with feedback. The first two practice items were designed to introduce the terminology of the test, in particular the phrase "starts the same," using syllable-level initial sound identity, for example, *dollar* and *dolphin* or *garden* and *garbage*. The lower bound reliability estimate was a relatively low .53. This value was probably affected by the large numbers of children scoring around chance on the first administration (see Results, below, for chance-level performance of the FAR group), confirmed by a Cronbach's alpha for this test occasion of just .43. At postintervention, both groups of children scored above chance, with a Cronbach's alpha value of .70.

Rhyme recognition. Rhyme recognition was measured using a 10-item test developed by Byrne and Fielding-Barnsley (1991a). In this test, the child was required to indicate which of three orally presented words (e.g., *leg*, *car*, and *bike*) rhymed with a target word (*star*). All of the words were pictured. Test occasion correlation = .64; Cronbach's α = .75.

Letter knowledge. Letter knowledge was assessed in multiple-choice recognition format. For each letter, the child was shown a row of four lowercase letters and asked to point to the one that matched the letter name said by the examiner. In previous research, we have found the recognition paradigm to be more sensitive than name recall (Byrne & Fielding-Barnsley, 1991). Test occasion correlation = .79; Cronbach's α = .91.

Knowledge about print. Emerging knowledge about print was measured using the 24-item Concepts About Print test developed by Clay (1975). This test assessed concepts such as why people read, how a book is manipulated, and the differences between pictures and text. Test occasion correlation = .66; Cronbach's α = .69.

Receptive vocabulary. The measure of receptive vocabulary used was the Peabody Picture Vocabulary Test—Revised (PPVT-R; Dunn & Dunn, 1981). Half of the children received Form L, and the other half received Form M. Test occasion correlation = .81.

Expressive vocabulary. We used the Hundred Pictures Naming Test (HPNT; J. P. Fisher & Glenister, 1992), consisting of 100 line drawings of everyday objects that the child was required to name. Test occasion correlation = .83.

Block Design. Nonverbal ability was measured using the Block Design subtest of the WPPSI-R (Wechsler, 1989). Published test-retest reliability = .77.

Raven's Colored Progressive Matrices. This test's norms apply only to children 5.5 years and older, and therefore, raw scores were used. For the same reason, published reliability is not applicable. Cronbach's α = .52.

Temperament questionnaire. The Rutter Child Behavior Scale (Rutter, Tizard, & Whitmore, 1970) as used in the Australian Temperament Project (Prior, Sanson, & Oberklaid, 1989; Sanson, Prior, & Oberklaid, 1985) was completed by one parent in each family and by the child's preschool caregiver. This is a 30-item questionnaire focusing on aspects of temperament such as hostility, anxiety, hyperactivity, and distractibility. Cronbach's alpha for the parent version was .85; for the teacher version, it was .90.

Tests Administered to a Subset of Children

As described above, a subset of children, mainly the first cohort recruited, was assessed on a fuller battery of linguistic and cognitive measures. This subset consisted of 49 at-risk children and 41 controls.

Initial phoneme segmentation. The test was based on a task developed by Fox and Routh (1975). The child was asked to "say a little bit" of phrases and words to finally indicate that he or she could isolate the initial phoneme of a single-syllable word. In the 24 one-syllable words (4 practice and 20 test words) used for the initial phoneme segmentation component, the initial consonant phoneme was either a fricative (*/f/*, */v/*, */s/*, */z/*, or */ʃ/*) or a lateral (*/l/* or */r/*). (This test was subsequently dropped from further analysis on account of poor distribution; see Results, below.)

Phoneme blending. This 10-item test was presented using a puppet who "says his words in a funny way." The child's task was to indicate which of three pictured words, each three phonemes in length, was the one spoken by the puppet in segmented form (e.g., b-u-g). The phonemes were delivered at the rate of one per second. Cronbach's α = .76.

Word span. Word span was assessed using three sets of words. These consisted of one set of six phonologically dissimilar one-syllable words (*bus*, *clock*, *hand*, *horse*, *girl*, and *spoon*), one set of six phonologically similar one-syllable words (*hat*, *bat*, *cat*, *rat*, *mat*, and *tap*), and one set of six phonologically dissimilar three-syllable words (*telephone*, *aeroplane*, *butterfly*, *piano*, *banana*, and *kangaroo*).

The phonologically dissimilar words were those used by Gathercole and Baddeley (1990). The phonologically similar one-syllable words and the phonologically dissimilar three-syllable set were used by Roodenrys, Hulme, and Brown (1993). From each of these sets, three lists from two to six words long were generated at each length. The words were read at the rate of one per second, and if a child failed to correctly repeat two out of three lists at a given length, testing was halted. Span was scored as the maximum length achieved for each of the three list types. Cronbach's α = .78.

Sentence Memory. We used the Sentence Memory subtest from the WPPSI-R. This test consisted of 12 sentences ranging from 2 to 18 words in length that the child was asked to repeat verbatim. Published test-retest reliability = .76.

Nonword repetition. Nonword repetition was assessed using a test developed by Gathercole, Willis, Emslie, and Baddeley (1992). The test consisted of 40 nonwords, constructed so that there were equal numbers of 1-, 2-, 3-, and 4-syllable items. All of the items were constructed so that they were phonologically legal phoneme sequences in spoken English. Each nonword was spoken aloud by the examiner, who covered her mouth while producing them. Cronbach's α = .72.

Word repetition test. A set of words was selected to provide a test of word recognition that was broadly comparable with the test of nonword repetition. As with the test of nonword repetition, it comprised 40 items ranging from one to four syllables in length. As shown in Table 1, performance for both groups on this test was close to ceiling, with a Cronbach's alpha value of .46.

Speech discrimination. Speech discrimination was assessed using a test developed by Bishop (1985) in which children judged whether words were the same or different. Stimuli for the test consisted of 36 pairs of consonant-vowel-consonant (CVC) words. One third of the pairs were identical. One third differed within fricative and affricative classes (examples: *leaf*, *leave*; *badge*, *batch*; *shop*, *chop*). One third differed within nasal and plosive classes (examples: *duck*, *dug*; *mock*, *knock*; *tail*, *nail*). In half of the different pairs, the initial phoneme contrast was different, and in the other half, the final phoneme contrast was different. In Results, below, we report just the data for successful discrimination of the 24 different items. Cronbach's α = .86.

Articulation. Articulation was assessed using the Sounds in Words subtest from the Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 1986). The test permitted assessment of the child's spontaneous production of all of the most frequent consonants of English in their most common positions as well as 11 consonant blends. Overall performance was over 90.0%. Cronbach's α = .59, probably reflecting the ceiling effect and restricted variance.

Word identification point. Word identification point was assessed using a set of 16 CVC words selected from the database of 906 words compiled by Luce (1986). These 16 words were selected to vary on two major dimensions, neighborhood density (high density [HD] vs. low density [LD]) and frequency (high frequency [HF] vs. low frequency [LF]).

Measures of neighborhood density, the total number of words that differed from the target word by a one-phoneme substitution, were available from previous unpublished research (Metsala, 1993). A word was defined as coming from an HD neighborhood if it had more than 11

Table 1
Performance on Tests Administered to All Children Prior to Intervention

Test	FAR		nonFAR		<i>t</i>	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Phoneme-identity recognition (maximum = 20, chance = 10)	10.26	2.92	11.78	3.09	3.20 ($p < .01$)	0.49
Letter recognition (chance = 6.5)	10.04	5.64	13.76	6.00	4.04 ($p < .01$)	0.62
Rhyme recognition (maximum = 10, chance = 3.3)	5.11	2.66	6.99	2.63	4.50 ($p < .01$)	0.71
Concepts About Print (maximum = 24)	5.55	3.05	7.57	3.20	4.11 ($p < .01$)	0.63
PPVT-R	93.99	14.25	104.78	12.77	4.73 ($p < .01$)	0.84
Confrontation naming (maximum = 100)	73.77	11.31	81.88	8.26	5.03 ($p < .01$)	0.98
WPPSI-R Block Design	11.82	2.72	13.74	2.65	4.35 ($p < .01$)	0.72
Raven's Colored Progressive Matrices	11.67	3.59	12.96	3.15	2.37 ($p < .05$)	0.41
Temperament rating: teacher (maximum = 90)	42.33	10.57	37.54	9.04	3.02 ($p < .01$)	0.53
Temperament rating: parent	44.98	7.20	42.02	5.78	2.68 ($p < .01$)	0.51

Note. FAR = at-risk group; nonFAR = group not at risk; PPVT-R = Peabody Picture Vocabulary Test—Revised; WPPSI-R = Wechsler Preschool and Primary Scale of Intelligence—Revised.

neighbors and as coming from an LD neighborhood if it had fewer than 9 neighbors. Word frequency was based on Kucera and Francis (1967). A word was defined as HF if it occurred more than 40 times per million and as LF if it occurred fewer than 10 times per million. All words were selected to be familiar to 4-year-olds based on subjective age-of-acquisition ratings available from previous research (Metsala, 1993). Care was taken to select only words estimated to be acquired before 4.5 years. Examples of word types are *big* (HF, HD), *watch* (HF, LD), *bug* (LF, HD), and *nail* (LF, LD).

The stimulus words, as well as two practice items, were read by the experimenter and stored on computer. Gated versions of each word were prepared using the program SoundEdit Pro 1.0 (1992). The first gate for each word was 100 ms in duration, and each subsequent gate was an additional 60 ms from word onset (e.g., Gate 2 = 160-ms duration, Gate 3 = 220-ms duration, etc.), until the last trial, on which the complete word was presented.

The test stimuli were recorded on digital audiotape. Following Metsala (1993), each word was recorded with a 6-s interval between gates to allow the child to guess the target word. The number of gates varied from 9 to 15 depending on the length of the word. The test stimuli were recorded in two random orders to form Tape A and Tape B. Each tape had the same two gated words as practice trials. Half of the children received Tape A, and the other half received Tape B.

The experimenter introduced the game by explaining that the children would hear some words on the tape recorder. They would hear only the very beginning of each word and then would hear a little more of the word, until finally, they would hear the whole word. They were told that each time they heard one of these "little bits," they had to guess aloud what the word might be.

Following Walley (1988), the identification point for a word was defined as the number of gates present for the trial on which the child correctly identified it. If the child had not correctly responded by the last trial for a word, the identification point was scored as the number of gates present on the last trial plus one. Cronbach's $\alpha = .41$.

Articulation rate. Articulation rate was assessed using the task devised by Gathercole and Baddeley (1990) and Gathercole and Adams (1993) in which children were presented with pairs of words that they were required to repeat as quickly as possible. Two measures of articulation rate were taken, for one-syllable word pairs and for three-syllable word pairs. The items were *bedtap*, *egg/lamb*, *cat/pen* and *banana/telephone*, *elephant/ladybird*, *kangaroo/pineapple*, matched across syllable lengths for initial consonant phoneme. The examiner presented each pair of stimuli in the order in which they are listed above. After one correct repetition of a pair of words had been made, the child was asked to repeat the words over and

over again, as quickly as possible, until told to stop. A handheld tape recorder was used to record the child's repetition of the six word pairs, which were later transferred to a digital audiotape recorder. Six scores were recorded for each child: the times taken for the first three complete repetitions of the 3 one-syllable word pairs and for the first three complete repetitions of the 3 three-syllable word pairs. Cronbach's $\alpha = .76$.

Temporal terms. The measure used to assess syntactic comprehension was a test of the comprehension of sentences containing the temporal terms *before* and *after*. Using the approach of Macaruso, Bar-Shalom, Crain, and Shankweiler (1989), we set out to determine whether at-risk and not-at-risk preschoolers showed the same pattern of performance as disabled and nondisabled readers when processing demands were reduced. Processing demands were manipulated in two ways: First, two types of temporal terms sentences were used, ones in which the order in which events were mentioned either conflicted or corresponded with the conceptual order, as in (a) and (b), respectively, below.

(a) *Before you pick up the ball, pick up the car.*

(b) *After you pick up the truck, pick up the car.*

Sentences of Type (a) are assumed to place heavier demands on working memory because one of the clauses must be held in memory while a plan is formulated for acting out the sentence in the correct conceptual order. Second, these sentences were presented in two conditions, one that was designed to satisfy the presupposition of temporal terms sentences and one that was not. The presupposition associated with Sentence (a), for instance, is that the listener intends to pick up the ball. In the *felicity* condition, which is the condition designed to satisfy the presupposition, children were required to establish in advance their intention to perform the action mentioned in the clause introduced by the temporal term. In the *no-felicity* condition, no contextual support was provided; unmet presuppositions had to be accommodated into the listener's mental model of the discourse setting (Lewis, 1979), a procedure that is also presumably costly of processing resources.

The 12 sentences and 4 practice sentences used in the felicity condition were identical to the 12 and 4 used in the no-felicity condition except that in the felicity condition, the noun in the main clause and the noun in the subordinate clause were left unspecified because these were determined by the participant's choice on each trial. Within each of these conditions, half of the items incorporated conflicting word order, and half incorporated corresponding word order. Children were tested individually in two sessions. Half of the children received the no-felicity condition in the first session and the felicity condition in the second session, and the other half received the conditions in the reverse order. There was a minimum of 1 week between testing sessions for each child.

Four objects (*cow, truck, horse, and ball*) were used. For the felicity condition, the child was asked to point to the object he or she would like to pick up this time. The experimenter then incorporated the name of this object into the subordinate clause of the sentence and the name of one of the other objects into the main clause.

Performance was scored in terms of the number of errors made on the two sentence types (conceptual order and order of mention conflict vs. conceptual order and order of mention correspond) on each of the two conditions (felicity vs. no felicity). Maximum errors were six per condition. Cronbach's $\alpha = .60$, possibly reflecting generally low overall error rates (see Results below).

Visual matching. Visual matching ability was assessed using the 15-item Matching of Letter-Like Forms subtest from the British Abilities Scale (BAS; Elliot, Murray, & Pearson, 1983). Published test-retest reliability for children 5 years 0 months to 6 years 3 months = .68.

Procedure

All testing was conducted during the day in a quiet room in the child's preschool or home. Two testers administered the measures given to all children, one tester per child, in two sessions lasting between 20 and 30 min. Tests administered in the first session were, in order, phoneme identity, letter knowledge, rhyme, HPNT, and Concepts About Print. In Session 2, they were Raven's Colored Progressive Matrices, PPVT-R, and Block Design. A separate tester administered the additional measures given to the sample subset in four or five sessions lasting 25 to 35 min. The order was blending, word span, word and nonword repetition, Sentence Memory, temporal terms, visual matching, gating, Goldman-Fristoe and Bishop tests, and articulation rate.

Results

Not every child was able to complete each test, generally because of failures of attention and concentration. In addition, the decision to use the WPPSI-R Block Design test was made only after the first few children had been tested. Thus, the *ns* for each test can be less than the total number of children.

We first present results on the tests administered to all children (maximum 101 in the FAR group and 67 in the nonFAR group). For almost all tests, there was a maximum of 8 missing data points (mostly 0 or 1) except for the parent-completed temperament

rating scale, for which the FAR and nonFAR *ns* were 85 and 64, respectively. The data, along with tests of significance and effect sizes, are given in Table 1. For each of the variables, there was a significant difference in favor of the nonFAR group, with effect sizes ranging from modest (0.41 for Raven's Colored Progressive Matrices) to substantial (0.98 for expressive vocabulary).

Some of the tests (phoneme identity, rhyme, and letter knowledge) had a chance component. On phoneme identity, the FAR group's mean of 10.26 was not significantly above the chance value of 10.00, $t(98) = 0.89$. All other performance levels by both groups were above chance.

There were sex effects on just two of the measures. On Concepts About Print, the mean for girls at 7.15 was higher than the mean for boys at 5.45, $F(1, 164) = 11.88, p < .01$. Sex did not interact with risk status ($F < 1$). On the teacher-completed temperament questionnaire, the mean for girls at 38.52 was lower than the mean for boys at 42.36, $F(1, 160) = 5.86, p = .02$. Again, sex and risk status did not interact. Teachers tended to judge girls as less prone to behavior problems than boys, although the effect size of 0.38 indicates that this was not a strong tendency.

Each of the group differences in Table 1 remained significant after controlling for Block Design except the Raven's means ($F < 1$). We also used highest occupational status as a covariate because the two groups had significantly different means. All differences remained significant, including Raven's, $F(1, 160) = 4.01, p = .05$. The results we report from here on were unaffected by occupational status or Block Design unless noted.

Table 2 provides descriptive statistics for the extra measures from the subsample of 90 children. There was a serious floor effect for phoneme segmentation, with well over half of the children not scoring above zero (69.0% of the FAR group and 49.0% of the nonFAR group). Almost all of the remainder scored close to the ceiling of 20. Because of the highly skewed distributions, this measure was not considered further.

Phoneme blending was managed better by the nonFAR children, consistent with the earlier results for phoneme identity. The FAR children had more limited memory capacity for verbal material, as assessed by word span, WPPSI-R Sentence Memory, and non-

Table 2

Performance on Preintervention Measures (Means and Standard Deviations) Administered to Subsets of Children as a Function of Group

Measure	FAR		nonFAR		<i>t</i>	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Phoneme blending (maximum = 10)	5.00	2.63	7.31	2.28	4.32 ($p < .01$)	1.01
Word-span total (maximum = 18)	7.57	1.51	8.49	1.52	2.86 ($p = .01$)	0.61
WPPSI-R Sentence Memory (scaled score)	11.07	4.78	13.98	4.53	2.89 ($p < .01$)	0.64
Nonword repetition (maximum = 40)	32.42	5.23	35.29	3.99	2.72 ($p < .01$)	0.72
Word repetition (maximum = 40)	38.41	1.68	39.00	1.45	1.72 ($p > .05$)	0.40
Speech discrimination (maximum = 24)	16.79	5.31	18.33	5.14	1.21 ($p > .05$)	0.30
Goldman-Fristoe Test (incorrect articulations, maximum = 44)	4.66	3.47	3.78	4.81	0.95 ($p > .05$)	0.18
Word identification point (mean proportion of gates required for correct identification)	0.71	0.10	0.69	0.09	0.83 ($p > .05$)	0.22
Articulation rate (total time for all items in ms)	25,702	4,282	22,285	3,392	3.92 ($p < .01$)	1.01
Temporal terms (<i>M</i> errors per condition, maximum = 6)	1.99	1.06	1.43	1.13	2.38 ($p < .05$)	0.50
Visual matching	7.50	3.50	7.80	3.34	0.39 ($p > .05$)	0.09

Note. FAR = at-risk group; nonFAR = group not at risk; WPPSI-R = Wechsler Preschool and Primary Scale of Intelligence—Revised.

word repetition. Closer inspection of the data shows that it was only with four-syllable nonwords that the two groups differed, with the Group \times Length interaction significant, $F(3, 237) = 4.55$, $p < .01$, and individual comparisons indicating that only the difference for the longest items was reliable. The difference between the groups in articulation rate, which held for both one- and three-syllable words and with no Group \times Length interaction, furnished the largest effect size of all comparisons. (Note that articulation rate data were collected for only 78 of the 90 children, 41 FAR and 37 nonFAR. Two children per group declined to complete the task, and the remainder could not be tested within the time frame of the project.)

There was no significant difference between the groups on the BAS visual discrimination test, and the group difference on the temporal-terms measure did not survive controlling for Block Design, $F(1, 71) = 1.69$, $p = .20$. Several measures (word span, speech discrimination, word identification point, and temporal terms) contained within-subject manipulations designed to assess the sensitivity of the measures with these groups of children. In each case except word span, there were no significant Group \times Condition interactions, but there were condition effects. In the test of word span, we varied length (one and three syllables) and phonetic similarity within one-syllable items. After controlling for Block Design, there was no group effect for one-syllable words, $F(1, 81) = 2.03$, $p = .16$, but the effect remained significant for three-syllable words, $F(1, 81) = 7.11$, $p = .01$. There was a main effect of phonological similarity for one-syllable items, $F(1, 88) = 138.98$, $p < .001$, with lists of phonologically similar words being more difficult to recall than lists of phonologically dissimilar words. There was also the expected main effect of word length when one-syllable dissimilar words were compared with three-syllable (also dissimilar) words, $F(1, 88) = 147.50$, $p < .001$. Because the word-length effect has often been taken as evidence for the use of speech-based codes in short-term memory (see Snowling, 2000, for a review), we examined the correlations between articulation rate and word span. For one-syllable similar, one-syllable dissimilar, and three-syllable words, the values were $-.40$, $-.45$, and $-.38$, respectively (all $ps < .01$).

For our measure of speech discrimination, it was possible to collect complete data on only 67 of the 90 children (34 FAR and 33 nonFAR); 7 per group declined to complete the test, and the other 9 were unavailable to be tested within the time frame of the project. An effect of consonant type, $F(1, 65) = 21.77$, $p < .001$, favoring the fricative/affricative items was consistent with the results reported by Bishop (1985). This suggested that the measure was sufficiently sensitive to detect group differences, yet none were obtained (see Table 2).

For word identification point, the unit of analysis was the gate at which the word was correctly identified as a proportion of the total number of gates for that word. There was no effect of group (see Table 2) or of frequency, $F(1, 83) = < 1$, but LD words were more easily identified than HD words, $F(1, 83) = 188.62$, $p < .001$, and there was a significant Frequency \times Neighborhood Density interaction, $F(1, 83) = 37.24$, $p < .001$, with HF/LD words more easily identified than any of the other types. This was consistent with prior research using the gating paradigm (Metsala, 1997; Metsala & Walley, 1998), so the failure to find a group difference is unlikely to have been due to insensitivity of our

materials and methods, despite the relatively low Cronbach's alpha of .41.

On interpretation of temporal terms, although the group effect did not survive controlling for Block Design, there were significant effects for condition (felicity vs. no felicity), $F(1, 78) = 11.73$, $p < .01$, and sentence type (order), $F(1, 78) = 20.59$, $p < .01$. Creating testing conditions that honored the pragmatic conditions under which temporal terms were used improved performance for both groups, as did creating a correspondence between order of mention and order of action, consistent with the findings of Macaruso et al. (1989). The mean error scores out of six for the most favorable conditions (felicity with corresponding order) were 1.16 and 0.92 for FAR and nonFAR children, respectively.

Discussion

We discuss the results without distinguishing between those tasks given to all children and those given only to the subsample. The FAR children performed more poorly in many of the tasks, at a similar level in some, and better in none. The pattern of differences is reminiscent of differences between competent and poor readers of older ages, as we discuss below, so perhaps the most striking aspect of the results is that differences were in evidence prior to formal education and largely prior to relative success and failure in reading. The important implication of this observation is that the differences found in older children are unlikely to result from reading-level variation.

Measures of Print-Related Processes

FAR children knew fewer letter names and were less conversant with print conventions than the control children, in both cases with effect sizes over 0.6. Letter knowledge prior to and during the early stages of schooling has been linked to progress in early reading. Candidate sources for the difference include the home and preschool environments, particularly explicit teaching of letters, and child-internal factors such as propensity to profit from experience, explicit or implicit, perhaps under genetic influence (Byrne et al., 2002, 2005).

Phonological Variables Close to Reading

We assessed phonemic awareness in several ways, two of which (recognition of phoneme identity and phoneme blending) gave interpretable data. In both cases, there was a significant group difference in favor of the nonFAR children, with phoneme blending in particular showing a substantial effect size of 1.01 (phoneme-identity effect size was 0.49). The majority of preschool children have negligible levels of phonemic awareness, as evidenced in our data by the near-chance performance even in the nonFAR children on phoneme identity (11.78 out of 20.00, with chance = 10.00) and by the large number of children not scoring at all on phoneme segmentation. It is likely, therefore, that the group differences we did observe were a function of small numbers of children who did have a degree of sensitivity to phonemic structure and that fewer of these children were to be found in the at-risk group.

Sensitivity to rhyme is another measure of phonological awareness that also predicts subsequent reading levels. Here, too, the

at-risk children performed at a lower level than the controls. Means were well above chance, so rhyme may better reflect phonological awareness levels for children of this age.

Other Aspects of Phonological Processing

Some of our other measures (word span, memory for sentences, and nonword repetition) involved the use of short-term memory within the phonological domain. In each case, group differences emerged, particularly where task demands were higher, as with the longer nonword repetition items and the significant three-syllable versus nonsignificant one-syllable word-span effect after controlling for Block Design. Contrary to some suggestions that speech-based memory differences result from reading-level differences (Pennington, Van Orden, Kirson, & Haith, 1991; Wadsworth, DeFries, Fulker, Olson, & Pennington, 1995), our data indicate that they are in place prior to formal reading instruction. Thus, this memory system may be important for developing skill in printed word recognition, as well as contributing to basic language skills that support reading, vocabulary learning, and sentence processing (Baddeley, Gathercole, & Papagno, 1998).

We assessed speech perception and speech production with tests used to detect clinically significant deficiency or delay: Bishop's test for consonant discrimination and the Goldman-Fristoe Test of Articulation. On neither test did the groups differ, indicating that our FAR sample was not, as a group, characterized by readily observable delay in these processes, a finding in general agreement with other research on prediction of reading disability (Scarborough, 1998). This is not to say that more stringent tests of either process would not reveal differences. In speech perception, for instance, inconsistent responses on tests of categorical perception and weak performance on various phonetic discrimination tasks have sometimes been noted in poor readers (see Brady, 1997, and Snowling, 2000, for reviews), and recent data suggest that dyslexic children may be delayed in their perception of amplitude envelope onsets leading to defects in rhythmic timing (Goswami et al., 2002). Furthermore, recent evidence from the Finnish group highlights speech perception anomalies as a function of risk status from a very early age (Leppänen et al., 2003; Richardson, Leppänen, Leiwo, & Lyytinen, 2003).

Articulation rate did show a group difference that, with an effect size of 1.01, was as large as any we detected. The magnitude of this difference can be seen as consistent with Snowling's (2000) hypothesis that children with deficits in output phonology are particularly vulnerable to developing disorders of literacy, both decoding and spelling. As with other aspects of our data, the presence of a possible rate deficit at the preschool level suggests that deficits detectable in older children cannot be attributed to reading difficulties; they predate reading instruction.

There were modest but significant correlations between word span and articulation rate, providing support for the idea that working memory for verbal material involves the speech code. The size of the correlations, around 0.4, indicates that more is involved in short-term memory than whatever articulation rate captures, but the precise relationship among these measures is not further illuminated by our data.

Vocabulary

The two groups were clearly distinguished by vocabulary scores, particularly for expressive vocabulary, with an effect size close to 1. Expressive vocabulary is one of the best predictors of early reading progress when measured prior to or early in schooling (Scarborough, 1998), so finding such a large difference between the risk groups appears to be consistent with the predictive picture.

There are suggestions that a substantive vocabulary supports, or in fact forces, a restructuring of the internal lexicon around phonemic units rather than around larger scale ones such as consonant-vowel or indeed entire lexical entries (Fowler, 1991; Metsala, 1998). By this restructuring mechanism, an appropriate foundation is laid for discovery of the alphabetic principle. It has also been suggested that word identification point can index the degree to which the lexicon is organized phonemically. However, we found no evidence for a group effect in word identification point in the gating task. We did find experimental effects of word frequency and neighborhood density, as have sometimes been reported (Metsala, 1997), confirming the sensitivity of our methods. Thus, insofar as gating is a valid measure of lexical structure, at-risk children had lexical entries no differently organized than control children despite having less comprehensive vocabularies. Our data are consistent with results from Griffiths and Snowling (2001), who failed to detect a difference between dyslexic children and chronological age controls in word identification point. Thus, on the basis of the present data, there is nothing to support the idea that the reading difficulties FAR children are likely to experience can be attributed to less segmentally structured internal lexicons just prior to the start of formal education.

Syntax

Our measure of syntactic processing was limited to one type of structure, the temporal terms *before* and *after*. Although the at-risk children performed worse than the controls, this effect disappeared when Block Design was factored in, suggesting a deficit that is nonspecific to risk status. In any case, error rates for the most favorable conditions were quite low, suggesting that the FAR children were not characterized by slow development of syntactic control. The fact that experimental manipulations, such as failing to honor pragmatic conditions under which temporal terms were used, affected performance indicates that one can appeal to a limitation in processing capacity, rather than a gap in this aspect of syntactic development (Shankweiler & Crain, 1986), to explain differing levels of performance among children. We need to add, however, that our assessment of grammatical functions was very limited, and although other work summarized in Shankweiler and Crain (1986) suggests there is merit in contrasting performance limitation with genuine syntactic lag when considering linguistic deficits in reading-disabled (and at-risk) children, the contribution of the present report to this debate is very restricted.

In summary, at-risk preschool children showed impairments in literacy-related processes such as letter-name knowledge, in phonological processes close to the core of reading such as phoneme-identity detection and phoneme blending, in other phonological processes typically found to be deficient in older children with identified reading difficulties such as nonword repetition and word

span (at least when the task is taxing enough), in output phonology, and in vocabulary. One common feature of the tasks in which at-risk children performed more poorly is that they involved the retrieval and manipulation of phonological codes, often identified as a way to characterize older children with marked reading problems (Snowling, 2000, and many other authors). Lower performance was not across the board, even in phonological processing. For instance, our sample of FAR children was not marked by frank speech discrimination and output difficulties of the kind that characterize children with language delay. Nor did they perform at a lower level in a test of visual discrimination. On the whole, the cognitive profile that typifies poor readers was already present in our preschool at-risk group.

PHASE 2: THE INSTRUCTIONAL PROGRAM

The intervention was based on a prereading program entitled Sound Foundations, designed by Byrne and Fielding-Barnsley (1991b) and evaluated in a longitudinal project following a group of children from preschool to Grade 5 (Byrne & Fielding-Barnsley, 1991a, 1993a, 1995; Byrne et al., 2000). The program had its origins in research conducted by Byrne and Fielding-Barnsley (1989, 1990), and it focused on teaching phoneme identity (e.g., teaching that *sun* and *sail* start the same and that *broom* and *drum* end the same).

Also included in the intervention was a book-reading period for each session modeled on the structured reading protocol developed by Whitehurst and colleagues (Arnold, Lonigan, Whitehurst, & Epstein, 1994; Whitehurst, Arnold, et al., 1994; Whitehurst et al., 1988). Whitehurst, Epstein, et al. (1994) used a similar combination of book reading and the Sound Foundations package with children eligible for Head Start and obtained positive outcomes in literacy. The hallmark of the book-reading protocol is a dialogic exchange between adult and child in which the child is questioned about the story and encouraged to relate it to his or her own experience. The research cited above, as well as other work (e.g., Fielding-Barnsley & Purdie, 2002, 2003; Wasik & Bond, 2001), has demonstrated that this kind of activity enhances various aspects of language (particularly vocabulary and literacy).

Method

Materials

Teaching Materials

In the Sound Foundations package, a limited set of phonemes is used: three continuant consonants, /s/, /f/ (as in *ship*), and /l/; four stops, /m/, /t/, /g/, and /p/; and two vowels, /æ/ (as in *bat*) and /e/ (as in *bed*). For this study, we omitted instruction in /f/, /g/, and /e/ to make it manageable in a reasonable time frame. The program contains colored posters for each phoneme, one in which many of the items depicted begin with the target sound and one in which many end with the target sound, except that for the two vowels, only word-initial items are used. There are also worksheets and card games designed to reinforce the idea of phoneme identity and recorded jingles with alliteration for each sound. For each phoneme, a sandpaper letter on stiff board is present, with the name and typical sound of the letter modeled by the instructor and practiced by the child. For a fuller description, see Byrne and Fielding-Barnsley (1991a).

The book-reading segment was conducted according to the guidelines of Whitehurst's group (e.g., Whitehurst, Epstein, et al., 1994). The acronym

CROWD summarizes the components, each of which takes the form of a question posed to the child during the story reading: completion (e.g., Dad and Amy drove to the . . . ?), recall (e.g., Why were they going to the dump?), open-ended prompts (e.g., You tell me about this page), *wh*-questions (e.g., What are they doing?), and distancing, that is, relating the story to the child's life (e.g., Do you remember when we went to the dump?).

Assessment Materials

Criterion tests. A 12-item criterion test for the day's phoneme in its position (initial or final) consisted of a target picture with an object whose name began or ended with the appropriate phoneme (e.g., *sun* for initial /s/ and *horse* for final /s/) and 12 other pictures. After 4 practice items, the child was asked to sort the 12 test items, 6 with the target phoneme and 6 with a different phoneme, into two piles without feedback. Because there was a chance level of 6 out of 12, a passing performance was set at 10 out of 12 ($p < .02$).

Posttests. The following tests were readministered at the completion of intervention: in the first testing session, phoneme identity, letter knowledge, rhyme recognition, HPNT, and Concepts About Print; in the second session, PPVT-R and two new measures, an extra test of phoneme identity and a structured reading test. For the PPVT-R, we administered the form alternate to the one given to each child at pretest.

The new test of phoneme identity consisted of two parts, one assessing phoneme identity for four of the phonemes that were the subject of the teaching (/s/, /m/, /l/, and /t/) and the other assessing four phonemes that were not part of the teaching (/f/, /n/, /b/, and /k/). For each part, there were 12 items, each consisting of a target picture, for instance, sun, and three choice pictures, for instance, seal, key, and book (for a fuller description, see Byrne & Fielding-Barnsley, 1991a). The purpose of this extra test was to check if phonemic awareness extended to phonemes that were not part of the teaching program. In earlier work, Byrne and Fielding-Barnsley (1991a) found evidence for such transfer in an unselected sample of preschool children, and it was an open question whether we would see this in at-risk children as well.

The structured test of print decoding was adapted from Byrne and Fielding-Barnsley (1991a). From the phonemes we taught, we created 10 words: *mat*, *sat*, *pat*, *tap*, *sam*, *pam*, *lam* (lamb), *lap*, *pal*, and *sap*. We arranged them in pairs and invited the child to point to the word we indicated. For instance, for the pair *sat* and *mat*, we asked, "Which one says *mat*?" There were 12 pairs in total, 6 with initial-letter alternation, as in *sat* and *mat*, and 6 with final-letter alternation, as in *pal* and *pat*. We also tested for knowledge of the six letters themselves by asking the child to point to the correct letter on a card with the six letters plus six other letters. In research described in Byrne and Fielding-Barnsley, preschool children instructed with the Sound Foundations package scored significantly above the chance score of 6 on the decoding test, whereas control children did not. Cronbach's $\alpha = .57$.

Procedure

Each teaching session, which lasted for about 30 min, included the same four main elements: instruction in phoneme identity, visual and tactile exposure to the letter representing the phoneme for that session, the book reading, and a criterion test for the day's phoneme in its position (initial or final). Children were taught individually, at home or at preschool, by one of two teachers. The order of phonemes taught was initial /s/, /m/, /p/, /l/, /t/, and /æ/, followed by final /s/, /m/, /p/, /l/, and /t/. Within the series of initial phonemes and final phonemes, at the start of each lesson other than the first, the previously taught (initial or final) phonemes were reviewed through the colored posters. The criterion test was administered at the end of each session. The schedule of teaching sessions depended primarily on each child's pattern of attendance at preschool, where most teaching took place, but typically, there were two or three sessions per week.

A child's progress through the intervention was partly governed by the criterion tests. Beginning with Session 2, in which initial /m/ was taught, if the child had failed the criterion test for /s/ (Session 1) or failed the test for /m/, an extra lesson was inserted, with both previous phonemes included in the instruction in abbreviated form. Both phonemes were tested at the end of Session 2, but if there was a second failure on either of them, the child was in any case moved on to the next phoneme, /p/. If the criterion test for /p/ was failed, a review lesson using all previous phonemes was inserted, with the child moving on to the next phoneme, /l/, irrespective of the criterion tests at the end of that session. This cycle was repeated throughout the remaining lessons. We elected to follow this modified criterion plan rather than insist on a strict pass for all phonemes in an attempt to maintain the child's interest and in the hope that further phonemic examples might foster phonemic awareness. We also decided to limit the total number of teaching sessions to 17, a decision based only on our intuition that children were aware if they were failing to grasp the ideas we were teaching and on a further intuition that a continued sense of failure might be detrimental. The possible number of sessions ranged from 11 to 17, with the lower number reflecting smooth progress characterized by no repeated lessons. This number served as one index of the child's response to our instruction (see Results and Discussion, below).

In addition to the instructed groups, a wait-list control group of FAR children was recruited as part of the final cohort. This group was pre- and posttested with a gap of around 3 months, the time it took on average to deliver the intervention. Because of time and resource limitations, we posttested these children only on the measures most directly assessing the instructional aims of the program, namely, to teach phonemic awareness and print conventions. Subsequently, the children in the group were given a truncated version of the intervention except for a few cases in which the parents merely wanted testing conducted to help determine the child's suitability for beginning school the following year. Because the wait-list children did not receive the full intervention program, we were unable to use their postintervention scores in combination or comparison with those of the other children.

Assignment of children to this control group was not random. The group was formed from children recruited into the project near the end of Years 2 and 3, mostly the latter, when insufficient time remained prior to the start of the next school year to complete the full intervention schedule. Although there was no reason to believe that timing within the project would bias the sample, on two tests (phoneme identity and letter knowledge), there were significant differences between the two FAR groups, $t(89) = 2.32$ and 1.97, respectively, although, on phoneme identity, the wait-list group's score of 8.88 was below chance, and thus, the group comparison lacks a clear interpretation. Note, however, that the main goal of the project was not to demonstrate the efficacy of the intervention per se; its elements had previously been shown to be effective, as reviewed earlier. Central to the

project were the relative responsiveness of the FAR and nonFAR children to this early intervention and the identification of pretest variables that best predicted responsiveness. Also, follow-up testing was to be based on standardized tests, which afford another kind of controlled comparison based on available norms and other information (see below). Nevertheless, we considered it profitable to add the wait-list group as a further check on the intervention and its effects on the different variables that we included.

Results and Discussion

There was some attrition during the intervention phase, with 69 of the 73 FAR children and 65 of the 68 nonFAR children completing the program. In most cases of attrition, the families moved from the area during the 3 to 4 months the intervention took.

Pre- and Postintervention Scores

We first compared the FAR children with their nonFAR counterparts. We were interested in two questions: whether there was improvement from pre- to posttest and whether the degree of improvement was the same for both groups. In the analyses, we needed to take account of the fact that nonFAR children scored higher on all variables at pretest and that there was thus less room for improvement for them. This may have masked potential differences in responsiveness. Thus, we computed simple change scores for each variable and compared the groups after controlling for pretest scores in analyses of covariance (ANCOVAs).

The pre- and postintervention scores are presented in Table 3, along with adjusted change scores. We can first report that the two intervention groups improved significantly on all variables: All t values for unadjusted change scores comparing the mean with zero were significant for each group on each variable (minimum t value = 5.17 for nonFAR PPVT-R scores). For three variables (phoneme identity, letter knowledge, and Concepts About Print), the ANCOVA analyses showed a greater degree of improvement by the nonFAR children: Respective values were $F(1, 131) = 5.53$, $p = .02$; $F(1, 131) = 8.26$, $p < .01$; and $F(1, 131) = 4.89$, $p = .03$. No other comparison approached significance. Thus, the intervention was effective for both groups, but the at-risk children were somewhat less responsive on the variables that were most directly targeted by the intervention (phoneme identity and print awareness) and likewise improved less on letter knowledge, although

Table 3
Pre- and Postintervention Scores: Means and Standard Deviations

Measure	nonFAR ($n = 65$)				FAR ($n = 69$)				FAR wait-list ($n = 17$)			
	Pre		Post		Pre		Post		Pre		Post	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Phoneme identity	11.83	2.90	16.18	3.14	10.61	2.91	14.43	3.16	8.88	2.84	10.12	3.21
Letter knowledge	13.85	6.05	18.35	6.13	10.87	6.07	14.00	6.15	7.71	5.81	10.12	6.55
Rhyme	7.08	2.66	8.35	7.42	5.13	2.66	7.17	2.41	5.65	2.64	5.82	2.55
PPVT-R	105.12	13.94	111.38	13.62	95.52	13.96	101.77	13.63	93.23	14.76	94.69	15.08
Confrontation naming	82.20	9.59	87.20	7.58	74.71	9.56	81.75	7.56	72.20	11.11	77.73	8.67
Concepts About Print	7.66	3.22	11.49	3.39	5.90	3.24	9.30	3.32	4.57	3.10	5.29	3.37

Note. FAR = at-risk group; nonFAR = group not at risk; Pre = preintervention; Post = postintervention; PPVT-R = Peabody Picture Vocabulary Test-Revised.

this was not as systematically targeted in the intervention (only the letters for the phonemes used were taught). However, one of these analyses, that for phoneme identity, did not survive an ANCOVA with highest occupational status as covariate, $F(1, 131) = 1.31$, *ns*, and another, that for letter knowledge, became borderline, $F(1, 131) = 3.66$, $p = .06$. Concepts About Print remained significant, $F(1, 131) = 4.89$, $p = .03$. These results suggest that the higher responsiveness of the nonFAR children on phoneme-identity training might better be attributed to some function of family occupational status than to risk status per se, although more work is needed to clarify the relations among these variables.

We turn to the comparison of the two FAR groups, the one subject to intervention and the wait-list group. The wait-list group improved on four variables: phoneme identity, $t(16) = 2.76$, $p < .01$, one-tailed; letter knowledge, $t(16) = 2.46$, $p < .05$; Concepts About print, $t(13) = 2.22$, $p < .05$; and confrontation naming, $t(14) = 4.81$, $p < .01$. Note, however, that on phoneme identity, the posttest score of 10.12 was indistinguishable from chance, and hence, the improvement from a below-chance score of 8.88 was more apparent than real. On the other variables, the passage of time appears to have resulted in significant gains.

The ANCOVA analyses testing for the relative degree of change across the two FAR groups showed significant effects in favor of the intervention group on phoneme identity, $F(1, 83) = 18.23$, $p < .01$; rhyme, $F(1, 83) = 9.33$, $p < .01$; PPVT-R, $F(1, 79) = 4.32$, $p = .04$; and Concepts About Print, $F(1, 80) = 6.87$, $p < .01$. The F value of 3.38 for confrontation naming had an associated p value of .07. These two groups did not differ on highest occupational status ($M_s = 5.31$ and 4.82 for intervention and wait-list groups, respectively; $t < 1$), so no adjustment for status was required. Putting these results together with the earlier ones, we can say that the instructed FAR children improved more on phoneme identity (and it is doubtful that the wait-list group improved at all), rhyme, recognition naming, and Concepts About Print but that alphabet knowledge improved in both groups to an equivalent degree. Overall, therefore, the intervention had beneficial effects for the at-risk children compared with natural increase in those variables that we specifically targeted, phoneme and print awareness, and in variables that were more incidental components of instruction, namely, rhyme and vocabulary.

Recall that we also administered another posttest of phoneme identity to check whether both types of preschoolers would generalize their knowledge of phoneme identity to phonemes that were not part of the instruction. The FAR children's means for the instructed and noninstructed phonemes were 8.56 ($SD = 2.72$) and 6.72 ($SD = 2.73$), respectively. The analogous figures for the nonFAR group were 9.11 ($SD = 3.31$) and 7.43 ($SD = 3.28$). The main effect of group was significant, $F(1, 113) = 12.13$, $p < .01$, and so was training status (instructed vs. noninstructed), $F(1, 113) = 5.97$, $p = .02$. However, the interaction term was not ($F < 1$). Thus, although children handled phonemes that had been the subject of explicit teaching better, and nonFAR children generally outperformed FAR children, both groups showed similar amounts of generalization. That is, FAR children were not characterized by a higher level of specificity of response to instruction than their nonFAR counterparts.

On the structured test of word decoding, the FAR children scored an average of 7.94 (maximum = 12, chance = 6; $SD = 2.13$), and the nonFAR children averaged 9.63 ($SD = 1.96$), a

significant group difference, $t(132) = 4.77$, $p < .01$. The FAR children's mean, the lower of the two, was significantly above chance, $t(68) = 7.58$, $p < .01$, an indication that the intervention had successfully transferred to decoding. A noninstructed group of preschoolers in the Byrne and Fielding-Barnsley (1991a) study did not achieve an above-chance mean on this test.

The two groups were also differentiated on the letter-naming test, which assessed just the six letters that were part of the program. The nonFAR children were close to ceiling at 5.35 ($SD = 1.07$), with the FAR children lower at 4.41 ($SD = 1.66$), $t(132) = 3.91$, $p < .01$. Thus, despite being exposed to each letter over at least two sustained sessions, once for the initial phoneme and once for the final phoneme (except for the vowel, in initial position only), the at-risk children did not on average reach as secure a level of letter-name knowledge.

Number of Teaching Sessions Required

There was one further measure of response to the instruction, the number of sessions each child required. Recall that the path through the program was in part governed by a child's success on the criterion test following each lesson, with a higher number of sessions signifying slower progress. The FAR children required a mean of 13.64 ($SD = 2.21$) lessons compared with the nonFAR children's 12.48 ($SD = 1.72$), $t(131) = 3.34$, $p < .01$. This contrast was not affected by covarying on highest occupational status, $F(1, 130) = 10.14$, $p < .01$.

To summarize: The intervention achieved its goals of raising levels of phoneme awareness and familiarity with the conventions of print in both the FAR and nonFAR children, with increases among the FAR children exceeding those from a wait-list control group of similar children. There are indications that the nonFAR group responded better to the central elements of the instruction, although the interpretation of this result is clouded somewhat in the case of phoneme awareness, where the family's occupational status overrode the reading status. The at-risk children required more teaching sessions to reach the levels they did than the not-at-risk children, another indication of a lower degree of responsiveness on the part of the FAR children. However, although rate of response appeared to differentiate the groups, degree of generalization of the products of learning did not—the FAR children transferred their insights into phoneme identity to untaught phonemes as readily as the control children did relative to the overall levels of performance.

Predictors of Response to Instruction

In this section, we examine which of our pretest variables best predicted how well the children responded to the intervention, confining the analyses to those variables given to the subsample of 90 children who received more intensive testing. We used three outcome variables: (a) postintervention phonemic awareness because that was the main focus of the instruction, (b) postintervention print awareness because that was a secondary focus, and (c) the number of sessions children required because that reflected the rate at which they worked through the program. We dubbed this last variable *progress*. In calculating postintervention phonemic awareness, we took advantage of the fact that we administered three measures: the 20-item, two-choice test also given as a pretest

and the two 12-item, three-choice tests given to examine the generalization from instructed to noninstructed phonemes. These three measures correlated in a range of .60–.62, all significant beyond .001. The composite was the summed standard scores of the three components (Cronbach's $\alpha = .84$).

For all analyses, we first entered the control variables identified earlier (Block Design and highest occupational status) together in Step 1, and we included FAR status in Step 2 to determine if classifying children on the basis of parental reading levels accounted for variance independent of the children's own preintervention scores. For analyses involving phoneme awareness and progress, preintervention phoneme identity, rhyme, and letter knowledge were added as predictors at Step 2, along with the factors Vocabulary and Verbal Short-Term Memory (see below). For print awareness, preintervention Concepts About Print was included in Step 2.

To determine which of the remaining preintervention variables should be included as predictors, we conducted a principal-components analysis of the seven language-based measures that clearly differentiated the two groups (except articulation rate, dropped because of the amount of missing data; see above). The variables were the three word-span tests (phonologically similar and dissimilar one-syllable words and three-syllable words), the two vocabulary tests (confrontation and recognition naming), Sentence Memory, and nonword repetition. Using oblimin rotation and Kaiser's rule, we identified two factors accounting for 65.5% of the variance, with the pattern matrix showing that the two vocabulary tests formed a clear factor (loadings of .81 and .95) and that the other five tests also all loaded on a single second factor (loadings from .66 to .84). We dubbed these factors Vocabulary and Verbal Short-Term Memory (Cronbach's $\alpha = .75$ and .82, respectively). For all of the analyses, these variables were entered at Step 2.

The regressions are shown in Table 4. For phonemic awareness, both Block Design and highest occupational status failed to account for a significant percentage of the variance in the first step of the analysis. The predictors added at Step 2 accounted for a total of 51.0% variance, but only preintervention phoneme identity and rhyme made contributions that were independent of the other variables. Thus, children's ability to profit from instruction in phoneme-level identity was particularly dependent on their preexisting phonological awareness at the level of both phoneme and rhyme. The independent contribution of rhyme fits the idea that sensitivity to larger scale phonological units paves the way for sensitivity to smaller scale ones (Goswami & Bryant, 1990).

The prediction of the number of sessions required (progress) looks somewhat different from the prediction of outcome level (postintervention phoneme identity). Block Design uniquely accounted for 14.8% of the variance in progress at Step 1. In Step 2, only Vocabulary and Verbal Short-Term Memory made contributions that were independent of the remaining variables.

For the analysis involving Concepts About Print, Block Design uniquely accounted for a significant 11.8% of the variance in the outcome measure. In Step 2, Block Design remained an independent contributor, joined by preintervention Concepts About Print and phoneme identity to account for 59.7% of variance.

For each of these three analyses, we investigated whether group membership (FAR vs. nonFAR) altered the pattern of results in two ways. By including FAR status among the predictor variables

and finding that it did not contribute to variance explained in an independent way, we showed that the children's own scores overtook parent-derived risk status in predicting response to instruction. We also entered as a third step all of the two-way interactions between group and each of the predictor variables. In no case did that step add significantly to the variance explained. Thus, although FAR children were shown to be somewhat less responsive to the intervention, it appears that the processes contributing to response were similar proportionally in the two groups: The FAR children were just less well endowed with one or more of these component processes.

PHASE 3: KINDERGARTEN FOLLOW-UP

The children were retested in the last 2 months of their first school year. The tests were partly based on those used by Byrne and Fielding-Barnsley (1993a) to assess the kindergarten progress of preschool children instructed with the Sound Foundations kit.

Method

Participants

Again, there was some attrition prior to the follow-up testing, mostly due to families leaving the region or not being able to be contacted. We assessed 57 FAR children and 44 nonFAR children, although, in a few instances, we were unable to administer all three tests.

Tests

Word identification was measured with the Word Identification subtest from the WRMT-R (Woodcock, 1987). Children were required to pronounce single words until they made errors on six successive items. Published Grade 1 split-half reliability with Spearman-Brown correction = .98.

For nonword identification, we created a 20-item list consisting of 5 two-letter, 10 three-letter, and 5 four-letter nonwords. The list was *om, ip, ef, ud, ag, bof, sim, neb, lat, pud, asp, ift, ust, ond, elk, flus, bram, spen, plog, and drib*. To make the task as realistic as possible, we also created drawings of novel animal-like forms and told the children that the words were the names of these "monsters" and that they were to try to read the names. We scored on a letter basis, with one point for each phoneme accurately decoded. Thus, the maximum score was 60. Cronbach's $\alpha = .95$.

For spelling, the children were asked to spell 10 words (*dog, man, one, said, blue, come, plug, went, limp, and tree*) and four nonwords (*ig, sut, frot, and yilt*). The scoring system, adapted from Liberman, Rubin, Duquès, and Carlisle (1985), placed a premium on accurate transcription of the words' phonemes, so that, for instance, the spelling *kum* for *come* earned five points, just one short of the six for *come*. For full details of administration and scoring of the spelling test, see Byrne and Fielding-Barnsley (1993a). Cronbach's $\alpha = .93$.

Results and Discussion

Group Mean Data

The mean scores for the two groups are presented in Table 5 with the WRMT-R Word Identification test in standard score form. The nonFAR group outperformed the FAR group on all three tests: Word Identification, $t(99) = 2.48, p = .01$; nonword decoding, $t(94) = 3.58, p < .01$; and spelling, $t(96) = 3.76, p < .01$.

Table 4

Summary of Hierarchical Regression Analyses for Variables Predicting Three Intervention Outcomes: Phonemic Awareness Composite (A), Progress (B), and Concepts About Print (C)

Outcome	Step	Variable	R^2	B	$SE\ B$	β	Zero-order correlation with dependent variable	S^2_e
A	1	Block Design	.00	.02	.04	.06	.06	.00
		Highest occupational level		.00	.04	.01	.00	.00
	2	Block Design	.52	-.06	.03	-.16	.00	.02
		Highest occupational level		.01	.03	.03	.00	.00
		FAR status		-.10	.08	-.11	-.34**	.01
		Pretest phoneme identity		.07*	.03	.24	.53**	.04
		Pretest letter knowledge		.01	.01	.08	.38**	.00
		Pretest rhyme		.10**	.03	.33	.61**	.06
		Vocabulary		.13	.11	.13	.49**	.01
		Verbal Short-Term Memory		.16	.11	.14	.48**	.01
B	1	Block Design	.15	-.32**	.08	-.39	-.39**	.15
		Highest occupational level		-.06	.08	-.08	-.01	.01
	2	Block Design	.54	-.14	.07	-.17	-.39**	.02
		Highest occupational level		-.07	.06	-.08	-.01	.01
		FAR status		-.06	.18	-.03	.32**	.00
		Pretest phoneme identity		-.08	.06	-.11	-.43**	.01
		Pretest letter knowledge		-.04	.03	-.01	-.37**	.00
		Pretest rhyme		-.13	.08	-.18	-.55**	.02
		Vocabulary		.73**	.25	-.32	-.64	.05
		Verbal Short-Term Memory		-.67**	.25	-.25	-.55**	.04
C	1	Block Design	.12	.46**	.14	.35	.35**	.12
		Highest occupational level		.04	.13	.03	-.03	.00
	2	Block Design	.60	.24*	.11	.18	.35**	.03
		Highest occupational level		.02	.10	.02	-.03	.00
		FAR status		-.25	.27	-.08	-.37**	.00
		Pretest phoneme identity		.34**	.10	.31	.59**	.06
		Pretest letter knowledge		.08	.04	.15	.47**	.02
		Pretest rhyme		.01	.11	.01	.41**	.00
		Pretest Concepts About Print		.37**	.09	.37	.65**	.08
		Vocabulary		-.24	.39	-.07	.50**	.00
		Verbal Short-Term Memory		.37	.38	.09	.46**	.01

Note. FAR status was coded as 0 for nonFAR and 1 for FAR.

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

Table 5

Kindergarten Reading and Spelling Means and Standard Deviations

Test	FAR		nonFAR	
	M	SD	M	SD
WRMT-R Word Identification (standard scores)	109.66	17.42	117.89	15.09
Nonword identification (maximum = 60)	28.33	21.36	42.91	17.94
Spelling (maximum = 84)	54.25	15.35	64.91	11.79

Note. FAR = at-risk group; nonFAR = group not at risk; WRMT-R = Woodcock Reading Mastery Tests—Revised.

Thus, the preschool instruction did not bridge the gap between the two groups.

Despite the group difference, we are able to offer an assessment of the FAR children's performance in terms of norms and of the performance of the control children in the study by Byrne and Fielding-Barnsley (1993a), undertaken with the Sound Foundations kit. The children in that control group afford a check on whether the FAR children reached the levels that Australian children might typically perform at without intervention prior to school. The mean for the 56 kindergarten children from the control group in Byrne and Fielding-Barnsley on the WRMT-R subtest was 108.6 ($SD = 11.1$), similar to the 109.7 ($SD = 17.4$) for the FAR children in the present study. Likewise, the spelling scores for the earlier control group and the current FAR group were close at 53.4 ($SD = 13.6$) and 54.3 ($SD = 15.4$), respectively. Thus, a

case can be made that the intervention raised the FAR children's performance to grade average, although in the absence of a dedicated uninstructed FAR group, this conclusion must remain tentative.

Prediction of Kindergarten Scores From Postintervention Performance

We now consider the prediction of the kindergarten results from the postintervention scores of the children. We modeled this analysis on a similar one in Byrne et al. (2000). In that article, which reported the Grade 5 results of the children in the longitudinal evaluation of Sound Foundations, two outcome measures were contrasted in their ability to predict reading performance 6 years after preschool instruction. These measures were the postintervention score on phoneme identity and the number of sessions before each child was judged secure in his or her understanding of phoneme identity, a measure analogous to the current variable, progress. Following Byrne et al., we conducted regression analyses for each of the three kindergarten variables in which the postintervention phoneme-identity composite score was entered at Step 1 and progress was added at Step 2. For these analyses, we used the entire available sample of children. The results are presented in Table 6.

The phoneme-identity composite accounted for substantial and significant variance in word identification, nonword identification, and spelling but did not contribute independently at Step 2, when progress was added in the cases of nonword reading and spelling. If progress is entered as Step 1, it accounts for 23.3% and 24.6% of variance in nonword identification and spelling, respectively, with, as shown in Table 6, the postintervention phoneme-identity composite adding no significant additional variance. Thus, we have evidence, just as Byrne et al. (2000) had, that rate of response to instruction predicts aspects of literacy growth independent of,

and probably in excess of, the level of phonemic awareness actually achieved during the instruction, although it is not clear why this pattern did not hold for word identification. It is perhaps of interest that nonword identification and spelling each placed special demands on phonological analysis of the target items, and the progress variable was determined by the children's mastery of each step of the phoneme-identity training.

Byrne et al. (2000) proposed that a general learning-rate parameter that determines both response to initial stages of reading instruction and subsequent growth throughout the school years needs to be factored into accounts of reading development and reading disability, counterbalancing the tendency to analyze development largely in terms of static variables, such as phonemic awareness, rapid naming, and working memory, measured on a one-time basis. Other authors, such as Clay (1975) and Vellutino, Scanlon, and Sipay (1997), have made a similar point. This general stance receives support from the results of our study.

To check whether the progress variable represented a specific learning trait or more general IQ, we created a composite from the best proxies for IQ we had available (PPVT-R and Block Design) by adding standard scores. Entering IQ first in regressions ($N = 93$) for the nonword reading and spelling, where progress was a significant predictor, accounted for significant amounts of variance (13.9% and 10.8%, respectively). However, in a model that also included progress, it was only progress that contributed unique variance (to totals of 29.4% and 29.3%, respectively). Thus, progress seems to capture a learning parameter independent of IQ, although more complete IQ assessment would need to be made before this conclusion could be accepted with confidence.

In summary, the follow-up data reveal that the at-risk children progressed less satisfactorily than their not-at-risk controls, but there are indications that the at-risk children were achieving grade-appropriate results of the kind that nondisabled children can be

Table 6
Summary of Hierarchal Regression Analyses for Variables Predicting Three Kindergarten Measures: Word Identification (A), Nonword Identification (B), and Spelling (C)

Measure	Step	Variable	R^2	B	$SE\ B$	β
A	1		.26			
		Phoneme-identity composite		4.08*	0.74	0.51
	2		.29			
		Phoneme-identity composite		3.04*	0.96	0.38
		Progress		-1.32	0.79	-0.20
B	1		.19			
		Phoneme-identity composite		5.48*	1.25	0.44
	2		.25			
		Phoneme-identity composite		2.98	1.57	0.24
		Progress		-3.19*	1.27	-0.31
C	1		.20			
		Phoneme-identity composite		3.90*	0.85	0.45
	2		.28			
		Phoneme-identity composite		1.99	1.05	0.23
		Progress		-2.48*	0.86	-0.35

Note. ΔR^2 for Step 2 is significant ($p < .05$) for B (nonword identification) and C (spelling).

* $p < .05$.

expected to achieve without early intervention. There are also indications that the growth children are likely to show in typical school classrooms was already visible during attempts to foster literacy foundations such as phonemic awareness prior to school entry: Rates of progress, independent of insights achieved, are the signs.

GENERAL DISCUSSION

This study had three goals: (a) characterizing the at-risk FAR population prior to reading instruction, (b) evaluating an intensive intervention program aimed at preventing reading disability, and (c) determining which pretraining measures best predicted responsiveness to the intervention by these and other children. The content of the intervention was based on procedures found to be effective in past research. The program was designed to promote core word-recognition skills by training in phoneme awareness and to promote vocabulary development and other advantages known to accrue from structured, interactive book reading. Trained preschoolers from a sample of FAR families were compared with trained preschoolers lacking the familial risk factor and with untrained FAR preschoolers. Assessment of the comparative effects of the intervention program on each group of children was carried out immediately after completion of the intervention and again a year later, after the children had also received subsequent reading instruction at school.

Preintervention Assessment: Summary and Conclusions

The most striking feature of the results is the degree of conformity to data from older children with emerging or established reading difficulties. Deficits in phonological awareness and print knowledge, in aspects of phonological processing, in working memory for linguistic material, and in vocabulary are characteristic of these older children (see Snowling, 2000, for a summary). The data strongly suggest that this pattern of deficits is in place before reading experience can play a part in determining it, and hence, it offers candidates for accounting for subsequent reading problems.

Demanding operations on phonetic material such as articulating quickly and repeating unfamiliar, lengthy sequences particularly affected the at-risk children. The first of these deficiencies clearly implicates phonological output as relatively impaired, a conclusion consistent with the many reports of deficits in rapid naming in older reading-disabled children. Short-term memory is involved in nonword repetition, and there is other evidence that this was impaired in the FAR group. FAR children appeared to be prone to the familiar *phonological core deficit*, a deficit in the manipulation of phonological codes. According to the data from the gating task, however, the core did not extend to the quality of the representations themselves. Rather, it was their use under demanding conditions that was compromised.

Two other aspects of phonological processing were not compromised in this sample either: speech perception (Bishop's test) and articulation (Goldman-Fristoe test). However, in line with the conclusion that group differences may depend on using specially challenging tests, we cannot say that in the present samples, no differences existed between the two types of children. In speech perception, for instance, others have reported deficits in psycho-

physical functions such as shallower category boundary gradients in reading-disabled children (e.g., Breier et al., 2001; Godfrey, Syrdal-Lasky, Millay, & Knox, 1981), and these may be present in at-risk children of the ages of the present sample. The Finnish group has also found reliable differences in reactions to speech stimuli as early as 6 months of age (H. Lyytinen et al., 2004).

On our syntactic measure, both groups performed well above floor on all conditions, and poorer performance was not uniquely associated with risk status. This provides further evidence that cognitive deficits associated with reading difficulties are primarily phonological in nature.

Yet they are not completely so, as the nonverbal reasoning data indicate. When the lower scores on Block Design, for instance, are considered in conjunction with the vocabulary results, there is evidence of lower overall IQ functioning in the FAR children. This pattern fits with common observation, and the fact that the effect size for Block Design was less than that for expressive vocabulary may also be consistent with the observation that it is lowered verbal IQ, rather than performance IQ, that is most characteristic of dyslexic children (Scanlon & Vellutino, 1996). Finally, the temperament ratings matched observations of an overlap between reading disability and attention disorders, possibly stemming in part from common genes (Willcutt et al., 2002). This aspect of the data is in line with the broad finding that much of what characterizes reading-disabled children is in place prior to formal literacy instruction.

Effects of Intervention: Summary and Conclusions

There is clear evidence that the instructional program was successful in promoting both groups of children's phonological awareness at segment and rhyme levels, their familiarity with books and print, and their receptive vocabularies, although the FAR group was somewhat less responsive than the other in terms of both outcome level and amount of instruction required to reach that level for most of these variables. Both groups showed evidence of generalizing phonemic awareness to phonemes that were not part of the instruction. There was no detectable difference between the groups in the degree of transfer, potentially eliminating generalization problems in the domain of phonological awareness as a source of at-risk children's difficulties in early literacy acquisition. However, although both groups again showed evidence of being able to transfer their newfound phonemic awareness to the task of decoding, as shown by above-chance performance in the structured decoding test, the nonFAR group was more successful at this than the FAR group. Thus, transfer outside the purely phonological domain to print interpretation was not as robust in the at-risk children.

At the end of kindergarten, the FAR group's performances in word identification, nonword decoding, and spelling could be classified as being at grade-appropriate levels even though the nonFAR children outperformed them on all three measures. The lack of a FAR control group that was monitored into kindergarten limits the confidence we are entitled to feel about the preschool intervention's longer term effects. However, the preschool instruction did enhance processes known to underpin literacy development among the FAR children, as attested by comparisons with the wait-list group. Furthermore, much research has shown that early intervention enhances school reading achievement (Byrne et al.,

2000; Torgesen, 2002). These observations combine to suggest that the instructed children performed better in reading and spelling than they would have without the instruction.

Predicting Response to Intervention

Gains in phonemic awareness in the sample as a whole were predicted particularly by preexisting levels of rhyme and phoneme sensitivity. Successful performance on the test of rhyme we used entails the ability to focus on a phonological unit (the rime) and the ability to recognize the identity of two such components. It is perhaps not surprising that children equipped with these abilities are well placed to benefit from instruction in a process that entails focusing on a phonological unit (the segment) and recognizing identity across two of them. Thus, even though evidence exists that instruction in rhyme does not automatically enhance phoneme-level sensitivity in children who lack it (Martin & Byrne, 2002), the current analyses suggest that promoting rhyme helps children profit from phoneme awareness training when it is delivered.

The independent predictors of the rate of response to instruction (progress) were expressive vocabulary and verbal short-term memory. This observation broadly fits with a conclusion reached by Snowling et al. (2003) on the basis of current and retrospective analyses of data from 8-year-old children previously classified as at risk or not. Snowling et al. concluded that a cognitive marker of dyslexia is a deficit in verbal association learning, leading to slow rates of learning of a variety of processes vital for reading, including letters and nursery rhymes. This leads, in turn, to downstream effects on word identification through decoding and encoding processes. The fact, too, that rate of response to early stages of instruction continues to predict reading growth for up to 6 years as further processes of word identification are added to the reading repertoire (Byrne et al., 2000) is consistent with this analysis.

Effects of Instruction 1 Year Later

In kindergarten, the nonFAR children were more able readers and spellers than the FAR children. Nevertheless, there were signs that the at-risk sample was performing on average at grade-appropriate level, as if it were a nonFAR sample with no preschool intervention. This is encouraging. It is also helpful to know that responsiveness to early instruction can be taken as a sign of future prospects, especially responsiveness measured by the rate variable we have called progress. If a standardized test of this variable can be created, researchers and educators would be well placed to identify those children needing special attention of an early and sustained kind.

References

- Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology*, 86, 235–243.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158–173.
- Bishop, D. V. (1985). Spelling ability in congenital dysarthria: Evidence against articulatory coding in translating between phonemes and graphemes. *Cognitive Neuropsychology*, 3, 229–251.
- Brady, S. A. (1997). Ability to encode phonological representations: An underlying difficulty of poor readers. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 21–47). Hillsdale, NJ: Erlbaum.
- Breier, J. I., Gray, L., Fletcher, J. M., Diehl, R. L., Klaas, P., Foorman, B. R., & Molis, M. R. (2001). Perception of voice and tone onset time continua in children with dyslexia with and without attention deficit/hyperactivity disorder. *Journal of Experimental Child Psychology*, 80, 245–270.
- Bruck, M. (1990). Word recognition skills of adults with childhood diagnoses of dyslexia. *Developmental Psychology*, 26, 439–454.
- Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91, 403–414.
- Byrne, B., Delaland, C., Fielding-Barnsley, R., Quain, P., Samuelsson, S., Hoien, T., et al. (2002). Longitudinal twin study of early reading development in three countries: Preliminary results. *Annals of Dyslexia*, 52, 49–73.
- Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, 81, 313–321.
- Byrne, B., & Fielding-Barnsley, R. (1990). Acquiring the alphabetic principle: A case for teaching recognition of phoneme identity. *Journal of Educational Psychology*, 82, 805–812.
- Byrne, B., & Fielding-Barnsley, R. (1991a). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology*, 83, 451–455.
- Byrne, B., & Fielding-Barnsley, R. (1991b). *Sound Foundations: An introduction to prereading skills*. Sydney, Australia: Peter Leyden Educational.
- Byrne, B., & Fielding-Barnsley, R. (1993a). Evaluation of a program to teach phonemic awareness to young children: A 1-year follow-up. *Journal of Educational Psychology*, 85, 104–111.
- Byrne, B., & Fielding-Barnsley, R. (1993b). Recognition of phoneme invariance by beginning readers: Confounding effects of global similarity. *Reading and Writing: An Interdisciplinary Journal*, 5, 315–324.
- Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology*, 87, 488–503.
- Byrne, B., Fielding-Barnsley, R., & Ashley, L. (2000). Effects of phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology*, 92, 659–667.
- Byrne, B., Fielding-Barnsley, R., Ashley, L., & Larson, K. (1997). Assessing the child's and the environment's contribution to reading acquisition: What we know and what we don't know. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 265–285). Hillsdale, NJ: Erlbaum.
- Byrne, B., Wadsworth, S., Corley, R., Samuelsson, S., Quain, P., DeFries, J. C., et al. (2005). Longitudinal twin study of early literacy development: Preschool and kindergarten phases. *Scientific Studies of Reading*, 9, 219–235.
- Carver, R. P. (1997). Reading for one second, one minute, or one year from the perspective of reading theory. *Scientific Studies of Reading*, 1, 3–43.
- Castles, I. (1989). *Australian Standard Classification of Occupations*. Canberra: Australian Bureau of Statistics.
- Castles, I. (1994). *Australian Standard Classification of Occupations* (2nd ed.). Canberra: Australian Bureau of Statistics.
- Clay, M. (1975). *The early detection of reading difficulties: A diagnostic survey*. Auckland, New Zealand: Heinemann.
- Cope, N., Harold, D., Hill, G., Moskvina, V., Stevenson, J., Holmans, P., et al. (2005). Strong evidence that KIAA0319 on Chromosome 6p is a susceptibility gene for developmental dyslexia. *American Journal of Human Genetics*, 76, 581–591.

- Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test—Revised* (3rd ed.). Circle Pines, MN: American Guidance Service.
- via shared book reading: The assessment of a family intervention program for preschool children at risk of becoming reading disabled. *Australian Journal of Learning Disabilities*, 7, 13–19.
- Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure. *Support for Learning: A Journal of the National Association for Special Education Needs*, 18, 73–78.
- Fisher, J. P., & Glenister, J. M. (1992). *The Hundred Pictures Naming Test*. Hawthorn: Australian Council for Education Research.
- Fisher, S. E., & DeFries, J. C. (2002). Developmental dyslexia: Genetic dissection of a complex trait. *Nature Reviews Neuroscience*, 3, 767–782.
- Fowler, A. (1991). How early phonological development might set the stage for phonological awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 97–117). Hillsdale, NJ: Erlbaum.
- Fox, B., & Routh, D. K. (1975). Analysing spoken language into words, syllables, and phonemes: A developmental survey. *Journal of Psycholinguistic Research*, 4, 331–342.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88, 3–17.
- Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk for dyslexia. *Journal of Child Psychology and Psychiatry*, 41, 203–213.
- Gathercole, S. E., & Adams, A. (1993). Phonological working memory in very young children. *Developmental Psychology*, 29, 770–778.
- Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, 29, 336–360.
- Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, 28, 887–898.
- Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K. K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology*, 32, 401–424.
- Goldman, R., & Fristoe, M. (1986). *Goldman-Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Service.
- Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read*. Hove, England: Erlbaum.
- Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences, USA*, 99, 10911–10916.
- Griffiths, Y. M., & Snowling, M. J. (2001). Auditory word identification and phonological skills in dyslexia and average readers. *Applied Psycholinguistics*, 22, 419–439.
- Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Educational Psychology*, 38, 241–253.
- Karlsen, B., & Gardner, E. F. (1984). *Stanford Diagnostic Reading Test* (3rd ed.). New York: Psychological Corporation.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Lefly, D. L., & Pennington, B. F. (1996). Longitudinal study of children at high family risk for dyslexia: The first two years. In M. L. Rice (Ed.), *Toward a genetics of language* (pp. 49–75). Hillsdale, NJ: Erlbaum.
- Leppänen, P. H. T., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M., & Lyytinen, H. (2003). Brain responses to changes in speech. In D. B. Gray & J. L. Kavanagh (Eds.), *Neurophysiological measures of dyslexia*. Parkton, MD: York Press.
- Locke, J. L., Hodgson, J., Macaruso, P., Roberts, J., Lambrecht-Smith, S., & Guttentag, C. (1997). The development of developmental dyslexia. In C. Hulme & M. Snowling (Eds.), *Dyslexia: Biology, cognition and intervention* (pp. 72–96). London: Whurr Publishers.
- Luce, P. (1986). *Neighborhoods of words in the mental lexicon* (Research on Speech Perception Tech. Rep. No. 6). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.
- Lyytinen, H., Aro, M., Eklund, K., Erskine, J., Guttorm, T., Laakso, M.-L., et al. (2004). The development of children at familial risk for dyslexia: Birth to early school age. *Annals of Dyslexia*, 54, 184–220.
- Lyytinen, P., Poikkeus, A.-M., Laakso, M.-L., Eklund, K., & Lyytinen, H. (2001). Language development and symbolic play in children with and without familial risk for dyslexia. *Journal of Speech, Language, and Hearing Research*, 44, 873–885.
- Macaruso, P., Bar-Shalom, E., Crain, S., & Shankweiler, D. (1989). Comprehension of temporal terms by good and poor readers. *Language and Speech*, 32, 45–67.
- Martin, M., & Byrne, B. (2002). Teaching children to recognise rhyme does not directly promote phonemic awareness. *British Journal of Educational Psychology*, 72, 561–572.
- Metsala, J. L. (1993). *An examination of speech perception in reading disabled and normally achieving children using the gating paradigm*. Unpublished doctoral dissertation, University of Toronto, Toronto, Ontario, Canada.
- Metsala, J. L. (1997). Spoken word recognition in reading disabled children. *Journal of Educational Psychology*, 89, 159–169.
- 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 and beginning literacy* (pp. 89–120). Hillsdale, NJ: Erlbaum.
- Olson, R. K., & Byrne, B. (2005). Genetic and environmental influences on reading and language ability and disability. In H. Catts (Ed.), *Connections between language and reading disorders* (pp. 173–200). Hillsdale, NJ: Erlbaum.
- Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, 72, 816–833.
- Pennington, B. F., & Olson, R. K. (2005). Genetics of dyslexia. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 453–472). Oxford, England: Blackwell.
- Pennington, B. F., Van Orden, G., Kirsan, D., & Haith, M. (1991). What is the causal relation between verbal STM problems and dyslexia? In S. A. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 173–186). Hillsdale, NJ: Erlbaum.
- Prior, M., Sanson, A., & Oberklaid, F. (1989). The Australian Temperament Project. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 537–556). New York: Wiley.
- Raven, J. C. (1976). *Colored progressive matrices*. London: H. K. Lewis.
- Richardson, U., Leppänen, P. H. T., Leiwo, M., & Lyytinen, H. (2003). Speech perception of infants with high familial risk for dyslexia differs at the age of 6 months. *Developmental Neuropsychology*, 23, 385–397.
- Roodenrys, S., Hulme, C., & Brown, G. (1993). The development of

- short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology*, 56, 431–442.
- Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behavior*. London: Longmans.
- Sanson, A. V., Prior, M., & Oberklaid, F. (1985). Normative data on temperament in Australian infants. *Australian Journal of Psychology*, 37, 185–195.
- Satz, P., Taylor, H. G., Friel, J., & Fletcher, J. M. (1978). Some developmental and predictive precursors of reading disability: A six year follow-up. In A. L. Benton & D. Pearl (Eds.), *Dyslexia: An appraisal of current knowledge* (pp. 313–347). New York: Oxford University Press.
- Scanlon, D. M., & Vellutino, F. R. (1996). Prerequisite skills, early instruction, and success in first-grade reading. *Mental Retardation and Developmental Research Reviews*, 2, 54–63.
- Scarborough, H. S. (1989). Prediction of reading disability from familial and individual differences. *Journal of Educational Psychology*, 81, 101–108.
- Scarborough, H. S. (1990). *A comparison of methods for identifying reading disabilities in adults*. Unpublished manuscript.
- Scarborough, H. S. (1998). Early detection of children at risk for reading disabilities: Phonological awareness and other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York Press.
- Shankweiler, D., & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition*, 24, 139–168.
- Shankweiler, D., Crain, S., Brady, S., & Macaruso, P. (1992). Identifying the causes of reading disability. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 275–305). Hillsdale, NJ: Erlbaum.
- Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., et al. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading*, 3, 71–96.
- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Schneider, A. E., Marchione, K. E., Stuebing, K. K., et al. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study of Adolescence. *Pediatrics*, 104, 1–13.
- Snowling, M. J. (2000). *Dyslexia*. Oxford, England: Blackwell.
- Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development*, 74, 358–373.
- SoundEdit Pro 1.0. (1992). San Francisco, CA: Macromedia.
- Thomas, C. (1905). Congenital word-blindness and its treatment. *Ophthalmoscope*, 3, 380–385.
- Torgesen, J. K. (2002). Lessons learned from intervention research in reading: A way to go before we rest. In R. Stainthorpe & P. Tomlinson (Eds.), *Learning and teaching reading* (pp. 89–104). Leicester, England: British Psychological Society.
- Van der Leij, Lyytinen, H., & Zwarts, F. (2001). The study of infant cognitive processes in dyslexia. In A. J. Fawcett (Ed.), *Dyslexia: Theory and good practice* (pp. 160–181). London: Whurr Publishers.
- Vellutino, F. R., Scanlon, D. M., & Sipay, E. R. (1997). Towards distinguishing between cognitive and experiential deficits as primary sources of difficulty in learning to read: The importance of early intervention in diagnosing specific reading disability. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 347–379). Hillsdale, NJ: Erlbaum.
- Wadsworth, S. J., DeFries, J. C., Fulker, D. W., Olson, R. K., & Pennington, B. F. (1995). Reading performance and verbal short-term memory: A twin study of reciprocal causation. *Intelligence*, 20, 145–167.
- Walley, A. C. (1988). Spoken word recognition by young children and adults. *Cognitive Development*, 3, 137–165.
- Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, 13, 286–350.
- Waring, S., Prior, M., Sanson, A., & Smart, D. (1996). Predictors of "recovery" from reading disability. *Australian Journal of Psychology*, 48, 160–166.
- Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology*, 93, 243–250.
- Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scale of Intelligence for Children—Revised*. San Antonio, TX: Psychological Corporation.
- Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in daycare and home for children from low-income families. *Developmental Psychology*, 30, 679–689.
- Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology*, 86, 542–555.
- Whitehurst, G. J., Falco, F., Lonigan, C. J., Fischel, J. E., Valdez-Manchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture-book reading. *Developmental Psychology*, 24, 552–558.
- Willcutt, E. G., Pennington, B. F., Smith, S. D., Cardon, L. R., Gayán, J., Knopik, V. S., et al. (2002). Quantitative trait locus for reading disability on Chromosome 6p is pleiotropic for attention deficit/hyperactivity disorder. *American Journal of Medical Genetics*, 114, 260–268.
- Wimmer, H., Mayringer, H., & Landerl, K. (1998). Poor reading: A deficit in skill automatization or a phonological deficit? *Scientific Studies of Reading*, 2, 321–340.
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5, 211–238.
- Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests—Revised*. Circle Pines, MN: American Guidance Service.

Received February 26, 2004

Revision received April 20, 2005

Accepted June 6, 2005 ■