Who Are the Young Children for Whom Best Practices in Reading Are Ineffective? : An Experimental and Longitudinal Study

Stephanie Al Otaiba and Douglas Fuchs

J Learn Disabil 2006 39: 414

DOI: 10.1177/00222194060390050401

Who Are the Young Children for Whom Best Practices in Reading Are Ineffective?

An Experimental and Longitudinal Study

Stephanie Al Otaiba and Douglas Fuchs

Abstract

The primary purpose of this study was to identify student characteristics that reliably predict responsiveness and nonresponsiveness to generally effective early literacy interventions. Participants were 104 children, including 7 with special needs and Individualized Education Programs (IEPs), who were tested in kindergarten and first grade. Responsiveness/nonresponsiveness status was determined after 2 years during which children participated in best practice instruction (a) in kindergarten and first grade, (b) in kindergarten only, (c) in first grade only, or (d) in neither year. This facilitated the study of three groups. *Always responsive* students met responsiveness criteria in both years. *Sometimes responsive* students met the criteria in only one year. *Nonresponsive students* did not meet the criteria in either year. Multivariate analysis of variance and discriminant function analysis indicated that the three groups were reliably different from one another on measures of problem behavior, verbal memory, sentence imitation, syntactic awareness, vocabulary, naming speed, and segmentation. A combination of naming speed, vocabulary, sentence imitation, problem behavior, and amount of intervention correctly predicted 82.1% of nonresponsive students, 30.0% of sometimes responsive students, and 84.1% of always responsive students. Fifty students from kindergarten and first grade were tested again at the end of what should have been their third-grade year. All but 1 of the nonresponsive students who received intervention had been identified as requiring special education and had an IEP with reading goals.

The gap between proficient and less proficient readers widens over the elementary years (Stanovich, 1986), and remediation of reading problems becomes increasingly difficult after third grade (Fletcher & Foorman, 1994; Kennedy, Birman, & Demaline, 1986; Lyon, 1985). Moreover, the long-term negative effects of illiteracy have been well documented (e.g., Adams, 1990). Thus, it is only fitting that the prevention of reading difficulties has become a national priority. Two recent federally sponsored reviews of the literature, *Preventing Reading Difficulties* (Snow, Burns, & Griffin, 1998) and *The National Reading Panel Report* (NRP; 2000), have documented the effectiveness of explicit and systematic early literacy interventions and have recommended providing more intensive early interventions to help all children read by the end of third grade. The findings and recommendations of these two reports, known collectively as *scientifically based reading research* (SBRR), have influenced recent educational policy, as reflected by the frequent references to it in the Reading First Initiative of the No Child Left Behind Act (P.L. 107-110, H.R. 1).

No Child Left Behind and Nonresponders

In the spirit of this bold new law, we join a growing number of researchers and educators who have expressed concern that as many as 30% of children at risk for reading difficulties (e.g., Blachman, 1994, 1997; Brown & Felton, 1990; Juel, 1994; Mathes, Howard, Allen, & Fuchs, 1998; Shanahan & Barr, 1995; Smith-Burke & Jaggar, 1994; Torgesen, Morgan, & Davis, 1992) may not benefit from generally effective early literacy interventions of the kind suggested by SBRR. These students have been called "treatment resisters" or "nonresponders" (e.g., Blachman, 1994; Torgesen, 2000). Our work and investigations by others (O'Connor, Jenkins, Leicester, & Slocum, 1993; O'Connor, Jenkins, & Slocum, 1995) have suggested that the percentage of nonresponders among children with learning disabilities may be as high as 50%.

Such unsettling information led us to ask, What do we know about nonresponders that might lead to their early identification and more effective remediation? To answer this question, we recently examined 23 studies in which (a) participants ranged from preschoolers to third graders, (b) participants were at risk for reading disabil-

ities, (c) interventions targeted early literacy, (d) outcomes reflected reading development, and (e) descriptions of students unresponsive to the interventions were provided. Our findings indicated that early literacy interventions often helped students, including many students with disabilities. However, depending on the individual study and outcome measure, between 8% and 80% of students showed little or no improvement (see Al Otaiba & Fuchs, 2002).

More important, perhaps, we found seven categories of child characteristics associated with nonresponsiveness: (a) phonological awareness; (b) verbal memory; (c) rapid naming; (d) vocabulary, verbal ability, and IQ; (e) attention or behavior problems; (f) orthographic awareness; and (g) home background (including socioeconomic status). Phonological awareness was explored in 21 of the 23 studies; in 70% of them, it correlated significantly with nonresponsiveness to intervention. Vocabulary, verbal ability, or IQ was investigated in 15 studies. Five research groups reported that children with low vocabulary, low verbal ability, or low IQ were more likely to be nonresponders. Seven research teams reported no statistically significant relationship between these child characteristics and nonresponsiveness. Three investigations produced mixed results.

The remaining child characteristics were explored much less often: 61% of research teams did not address the importance of verbal memory, and 70%, 61%, 70%, and 80% of the studies did not explore rapid naming, attention or behavior, orthographic processing, or demographics, respectively. To date, only a few investigators have assessed the importance of several of these child characteristics within the same study (see Torgesen & Davis, 1996; Torgesen et al., 1999; Uhry & Shepherd, 1997; Vellutino et al., 1996).

Methodological issues made it difficult to compare percentages and characteristics of nonresponders across the 23 studies in our review. These issues included different definitions of nonresponders, contrasting cutoff scores, and a general absence of information on the fidelity of treatment implementation. Furthermore, teachers conducted few interventions: Classroom teachers conducted three; resource room teachers, two. In fact, there was no direct test of the relationship between teachers' demonstrably accurate implementation of scientifically based reading interventions and nonresponsiveness. This is important because without fidelity data, one does not know whether nonresponsiveness is a consequence of improper treatment implementation or of child characteristics. We suggested in our review that the proper study of nonresponsiveness to scientifically based reading practice requires fidelity of treatment data and, perhaps, the tracking of nonresponders over an extended time. This second recommendation would be particularly important with respect to young children, given the relative instability of their school performance from one year to the next. Such research might, incidentally but not unimportantly, provide support for nonresponsiveness to intervention as a valid means of identifying children with disabilities (see Donovan & Cross, 2002; L. S. Fuchs & Fuchs, 1998; Vaughn & Fuchs, 2003).

Recently, Nelson, Benner, and Gonzalez (2003) extended our review in two ways. First, they found seven additional studies of nonresponders. Second, they used meta-analytic techniques to explore the relative importance of individual learner characteristics to treatment effectiveness. Noteworthy predictors were rapid naming (Z_r = 0.51), problem behavior (Z_r = 0.46), phonological awareness (Z_r = 0.42), alphabetic understanding (Z_r = 0.35), memory (Z_r = 0.31), IQ (Z_r = 0.26), and demographics (Z_r = 0.07).

Nonresponders and Multilevel Instruction

The recognition that generally effective early literacy programs do not accommodate the learning needs of all students has led to a strong interest in "multilevel" models of instruction and assessment (President's Commission on Excellence in Special Education, 2002). Multilevel models appear necessary because many nonresponders need more intensive instruction than is delivered in general education classrooms (e.g., Case, Speece, & Molloy, 2003; Dickman, Hennessy, Moats, Rooney, & Tomey, 2002; McMaster, Fuchs, Fuchs, & Compton, 2005; National Association of School Psychologists, 2002; O'Connor, 2000; Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino et al., 1996). Multilevel models are also preferable to traditional service delivery models because they are seen as providing intensive services sooner than special education does (e.g., Grimes, 2002; President's Commission on Excellence in Special Education, 2002). Currently, special education services are not available for most children with learning problems until third or fourth grade, when some children fall far enough below the level of reading achievement that is expected of them and they are given a learning disability (LD) label (e.g., Grimes; Lyon, Fletcher, Shaywitz, Shaywitz et al., 2001; President's Commission on Excellence in Special Education, 2002).

The first level of a multilevel approach typically calls for the classroom teacher to faithfully implement scientifically based instruction, with the expectation that the teacher will accelerate most children's learning. At the second level, instruction is more intensive. It may involve more instructional components, and it may be delivered more frequently and with greater duration. Because of its comparative complexity and intensity, this so-called *secondary intervention* is typically conducted by a reading teacher or paraprofessional—that is, someone other than the classroom teacher—and it is implemented in small student groups or individual tutorials (see McMaster et al., 2005 ; O'Connor, 2000; Vaughn et al., 2003; Vellutino et al., 1996).

Multilevel models or not, the question remains, Who are the children who do not respond to generally effective instruction? This is a truly pivotal issue because understanding the characteristics of chronically nonresponsive students may guide researchers and school personnel in developing (a) more sensitive screening measures that might better target children who require intensive services and (b) more effective interventions informed by knowledge of students' strengths and weaknesses.

Study Purpose

Our primary purpose was to contribute to the extant literature on the characteristics of children who are nonresponsive to early literacy interventions conducted by their classroom teachers in kindergarten and first grade. To extend what is known from prior investigations, we deliberately incorporated five features into our study design. First, we employed long-term (i.e., 2-year), potentially more stable definitions of responsiveness and nonresponsiveness. Relatively stable definitions could provide more valid information about the percentages and characteristics of persistent nonresponders and could lead to more accurate identification of students with LD.

Second, we evaluated responsiveness to intervention in "normal" school contexts. Kindergarten and first-grade teachers (rather than graduate students) were trained to conduct scientifically validated early literacy interventions to a heterogeneous mix of boys and girls, African American and European American students, children from middle class and low-income homes, and students with and without disabilities. Third, we carefully monitored the fidelity with which teachers and students conducted the interventions. Fourth, we selected a relatively large set of child measures that reliably differentiated nonresponders according to Al Otaiba and Fuchs' (2002) literature review. Our aim here was to begin with a sufficiently large number of relatively sensitive measures in hopes of developing an efficient and useful set of markers, or predictors, of nonresponsiveness to intervention. Finally, we followed our study participants over time to determine whether nonresponders in kindergarten and first grade would be identified as requiring special education services in reading by third grade—the grade in which all children should be reading at grade level according to the No Child Left Behind Act.

Method

Earlier Study

This study was part of an earlier and larger investigation designed to examine the 2-year effects of early literacy intervention on children's reading acquisition (see D. Fuchs, Thompson, Al Otaiba, et al., 2000). In the earlier study, teachers in kindergarten and first grade were assigned randomly to intervention conditions. As a result, the 312 participating students received either (a) 2 years of intervention, (b) only 1 year of intervention, in kindergarten, (c) only 1 year of intervention, in first grade, or (d) typical classroom instruction in both years (control condition). All participating teachers used the same district-adopted basal reading program (Harcourt Brace; Farr & Strickland, 1995), which represented an implicit approach to the teaching of phonological awareness and phonics instruction (Stein, Johnson, & Gutlohn, 1999).

The kindergarten and first-grade interventions, in contrast, were explicit and systematic, and they supplemented typical classroom reading instruction. At kindergarten, the intervention took one of two forms. One version consisted of teacher-directed phonological awareness activities selected from *Ladders to Literacy* (O'Connor, Notari-Syverson, & Vadasy, 1998), which were conducted in 5- to 15-min sessions three times per week for 20 weeks. The second variant was a combination of these teacher-directed activities and *Kindergarten Peer-Assisted Learning Strategies* (K-PALS; D. Fuchs, Fuchs, Thompson, Al Otaiba, Yen, McMaster, et al., 2001), which is peer mediated and provides phonological awareness and phonemic decoding practice. K-PALS sessions were conducted three times per week for 16 weeks. Each session lasted approximately 20 min. Because of concerns about small sample sizes, we combined *Ladders* and K-PALS plus *Ladders* study groups into a single "kindergarten intervention" group.

The first-grade intervention consisted of first-grade PALS (e.g., D. Fuchs, Fuchs, Svenson, et al., 2001), which teachers conducted in 20-min sessions three times a week for 20 weeks. First-grade PALS lessons included phonological awareness, decoding, and sight word training as well as reading in connected text (see Note 1).

Participants

Defining Nonresponsiveness and Responsiveness. At the end of kindergarten, nonresponsiveness to intervention was defined as performing in the lowest 30th percentile of intervention students on the amount of pre-to-posttreatment growth on letter-sound and segmentation fluency measures. In other words, nonresponsive kindergarteners could not segment more than 12 phonemes in 1 min or identify more than 11 letter sounds per minute. Good, Simmons, and Kame'enui (2001) reported that end-of-year kindergarteners who named fewer than 10 phonemes per minute on the *Dynamic Indicators of Beginning Literacy Skills* (DIBELS; Kaminski & Good, 1996) phoneme segmentation fluency measure were at risk for poor reading outcomes in first grade.

Nonresponsiveness in first grade was defined in terms of oral reading fluency, or the number of words read

aloud and correctly per minute. Evaluating student growth on this measure was not possible because it was too difficult for many first-grade students in the fall. Therefore, a cutoff criterion, or benchmark, of 40 words or more per minute read correctly from unfamiliar grade-level text at the end of first grade was selected, because it seemed to predict adequate or better future reading (Good, Simmons, & Smith, 1999).

Our decision to use fluency measures to identify nonresponders in kindergarten and first grade was informed by a meta-analysis conducted by D. Fuchs, Fuchs, Mathes, and Lipsey (2002), who found that timed measures were better than untimed measures at differentiating poor readers with an LD label from poor readers without an LD label. Also see Jenkins, Fuchs, van den Broek, Espin, and Deno's (2003) data-based discussion of the validity of fluency measures as indicators of reading competence and the *Final Report on Reading First Reading Assessment Analysis* (Kame'enui, 2002), which recommended fluency measures similar to those used in the present study for screening and progress monitoring.

Students were identified as *responsive* to kindergarten intervention if their pre-to-posttreatment growth on the segmentation and rapid letter-sound fluency measures was at or above the intervention group's mean. First-grade students were considered responsive if their end-of-year oral reading fluency was at or above the intervention group's mean.

Identifying Students Who Are Nonresponsive, Sometimes Responsive, and Always Responsive. We eliminated 14 of the 312 children in the earlier study because they repeated kindergarten, thereby making it impossible for us to follow them into first grade. Of the remaining 298 students, 146 met the aforementioned nonresponsive and responsive criteria. From this group, an additional 42 students were lost when they moved out of the school district during first grade, leaving a total of 104 participants. Of these 104 students, 79 participated in the interventions (40 in kindergarten but not in first grade; 11 in first grade but not in kindergarten; 28 in both kindergarten and first grade) and 25 received only typical classroom instruction.

At the end of first grade, all 104 students were classified into three groups in terms of their responsiveness to kindergarten and first-grade instruction: nonresponsive (n = 34), sometimes responsive (n = 26), and always responsive (n = 44). *Nonresponsive* indicated students who did not meet responsiveness criteria in kindergarten and first grade. *Sometimes responsive* denoted those who met the criteria in kindergarten or in first grade, but not in both. *Always responsive* referred to children who met the criteria in both years. Controls, who received typical classroom instruction, were also categorized as nonresponsive, sometimes responsive, or always responsive in order to compare their progress to that of the students who participated in the interventions and to describe nonresponsive and responsive children to typical classroom instruction. In Table 1, the numbers of students who were nonresponsive, sometimes responsive, and always responsive are displayed by whether they participated in (a) kindergarten and first-grade interventions, (b) kindergarten-only intervention, (c) first-grade–only intervention, and (d) the control condition (i.e., typical classroom instruction).

Table 2 shows student demographic data for the nonresponsive, sometimes responsive, and always responsive children. A one-way ANOVA indicated no statistically significant differences between the groups on chronological age, $F(2, 101) = 0.11$, *ns*. Chi-square analyses also revealed no statistically significant between-group differences on the type of school attended (Title I vs. non–Title I), $\chi^2(2, N = 104) = 0.82$, *ns*; gender, $\chi^2(2, N = 104) = 1.05$, *ns*; and racial identity, $\chi^2(4, N = 104) = 4.81$,*ns*. However, a chi-square test on the number of children with special needs who had an IEP was statistically significant for group, $\chi^2(2, N = 104) = 9.97$, $p < .01$. More nonresponsive students had IEPs, all but one of which called for speech services.

Measures

As indicated earlier, measures were selected in accordance with the child characteristics that were reliably associated with nonresponsive students in prior investigations (see Al Otaiba & Fuchs, 2002). These were rapid naming (Berninger et al., 1999; Torgesen & Davis, 1996; Torgesen et al., 1999; Uhry & Shepherd, 1997; Vellutino et al., 1996), verbal memory (Schneider Ennemoser, Roth, Kuspert, 1999; Uhry & Shepherd, 1997; Vellutino et al., 1996), phonological discrimination (Hurford,

TABLE 1
Numbers and Percentages of Students by Responsiveness Status Across Kindergarten and First Grade

Intervention	*n*	Nonresponsive	Sometimes responsive	Always responsive
Kindergarten and first grade	28	7 (25.00%)	5 (17.86%)	16 (57.14%)
Kindergarten only	40	5 (13.00%)	14 (35.00%)	21 (52.50%)
First grade only	11	4 (36.36%)	4 (36.36%)	3 (27.27%)
Controls	25	18 (72.00)%	3 (12.00%)	4 (16.00%)
Total	104	34	26	44

TABLE 2
Student Demographics by Responsiveness Status and Intervention

	Kindergarten and first grade[a]				First grade[b]				Kindergarten[c]				Controls[d]			
	M	*SD*	*n*	(%)	*M*	*SD*	*n*	(%)	*M*	*SD*	*n*	(%)	*M*	*SD*	*n*	(%)
Nonresponsive students (*n* = 34)																
Total			7				4				5				18	
Age	6.65	(0.46)			6.40	(0.25)			6.80	(0.46)			6.55	(0.46)		
ESL			0	0.00			0	0.00			0	0.00			0	0.00
IEP in kindergarten			1	14.30			1	25.00			0	0.00			3	16.70
IEP in first grade			3	42.90			1	25.00			0	0.00			3	16.70
Race																
African American			4	57.10			1	25.00			4	80.00			7	34.00
Caucasian			3	42.90			3	75.00			1	20.00			10	55.60
Other			0	0.00			0	0.00			0				1	5.60
Sex																
Female			2	28.60			2	50.00			2	40.00			10	55.60
Title 1			7	100.00			4	100.00			1	20.00			10	55.60
Sometimes responsive students (*n* = 26)																
Total			5				4				14				3	
Age	6.37	(.70)			6.68	(0.29)			6.59	(0.45)			6.70	(.56)		
ESL			0	0.00			0	0.00			0	0.00			0	0.00
IEP in kindergarten			0	0.00			0	0.00			1	7.10			0	0.00
IEP in first grade			0	0.00			0	0.00			1	7.10			0	0.00
Race																
African American			3	60.00			1	25.00			5	35.70			0	0.00
Caucasian			1	20.00			3	75.00			7	50.00			3	100.00
Other			1	20.00			0	0.00			2	14.30			0	0.00
Sex																
Female			3	60.00			2	50.00			6	42.90			1	33.33
Title I			4	80.00			1	25.00			9	64.30			0	0.00
Always responsive students (*n* = 44)																
Total			16				3				21				4	
Age	6.52	(0.32)			6.57	(0.47)			6.79	(0.17)			6.54	(0.41)		
ESL			0	0.00			0	0.00			0	0.00			0	0.00
IEP in kindergarten			0	0.00			0	0.00			0	0.00			0	0.00
IEP in first grade			0	0.00			0	0.00			0	0.00			0	0.00
Race																
African American			5	31.30			1	33.30			5	23.80			2	50.00
Caucasian			10	62.50			2	66.70			11	52.40			1	25.00
Other			1	6.30			0	0.00			5	23.80			1	25.00
Sex																
Female			11	68.80			0	0.00			10	47.60			4	100.00
Title I			8	50.00			2	66.70			14	66.70			1	25.00

Note. ESL = English as a second language; IEP = Individualized Education Program.
[a]*n* = 28. [b]*n* = 11. [c]*n* = 40. [d]*n* = 25.

1990), syntactic knowledge (Vellutino et al., 1996), verbal ability (Berninger et al., 1999; Foorman et al., 1998; O'Connor, Notari-Syverson, & Vadasy, 1996; Torgesen & Davis, 1996), and attention or behavior (O'Shaugnessy & Swanson, 2000; Snider, 1997; Torgesen et al., 1999; Uhry & Shepherd, 1997; Vellutino et al., 1996). Orthography was not included in the final set of measures because most kindergarten students could not write and did not know how to spell.

Measures Used to Define Responsiveness to Intervention. Three 1-min fluency measures administered during the earlier study (i.e., D. Fuchs, Fuchs, Thompson, Al Otaiba, Yen, Yang, et al., 2001) were used to determine children's responsiveness to intervention in this nonresponder study: the *Rapid Letter Sound* (RLS) test, the Yopp-Singer Segmentation Test, and oral reading fluency. At kindergarten, the RLS test (see Levy & Lysynchuk, 1997) was adapted to assess the number of letter sounds that a student named correctly. Students are shown a sheet with lowercase letters and asked to say each letter sound. The Yopp-Singer Segmentation Test (Yopp, 1995) requires children to say the sounds in words (e.g., "Say the sounds in *cat*"). This test correlates well with the DIBELS phoneme segmentation fluency measure (r = .77; Good, Simmons, & Kame'enui, 2001). At first grade, we used oral reading fluency. Its test–retest reliability ranges from .93 to .96 (L. S. Fuchs, Deno, & Marston, 1983).

Measures Used to Describe Nonresponders. To evaluate naming speed, we used student scores on the *Rapid Letter Naming Test* (RLN; D. Fuchs, Fuchs, Thompson, Al Otaiba, Yen, Yang, et al., 2001). Students are shown upper- and lowercase letters and are asked to name them as quickly as they can in 1 min. The RLN is a strong predictor of future reading ability (Adams, 1990; Juel, 1988).

We selected two subtests of the *Detroit Tests of Learning Aptitude*, third edition (DTLA-3; Hammill, 1991) to assess two types of verbal memory or encoding. The first subtest, Word Sequences, determines students' ability to encode and repeat a series of unrelated words (e.g., *cold*, *late*, *full*). The second subtest, Sentence Imitation, was used to assess whether the ability to encode sequential verbal information is enhanced by the syntactic structure of the sentence or whether the students still rely on word-by-word memory strategies (e.g., "My dog chases the white cat"). Internal consistency of Word Sequences is .88; for Sentence Imitation it is .84. Test–retest reliability coefficients for the two subtests are .89 and .92, respectively.

Syntactic knowledge was measured using the Grammatical Closure subtest of the *Test of Language Development–2–Primary* (TOLD-P2; Hammill & Newcomer, 1988). Students are asked to supply the missing word in a sentence (e.g., "John likes to cook every day. Yesterday, he _____" [cooked]). Internal consistency for this subtest is .92. Test–retest reliability is .91.

The Word Discrimination subtest of the TOLD-P2 was employed to assess phonological discrimination. Students are asked to say whether pairs of words are the same or different (e.g., *chop*, *shop*). Test–retest reliability is .84.

Because the literature seems equivocal on the importance of verbal ability, we added a vocabulary measure—specifically, the *Peabody Picture Vocabulary Test–Revised* (PPVT-R; Dunn & Dunn, 1981). Children are shown four pictures and asked to point to the one described by the examiner. Split-half reliability for children ages 6 to 8 ranges from .77 to .99; test–retest reliability ranges from .67 to .82.

Teachers administered the *Achenbach Child Behavior Checklist–Teacher Report Form* (Achenbach, 1994). The checklist consists of 118 items (e.g., "can't sit still, restless, or hyperactive"; "disturbs other children"; "disrupts class discipline"). Teachers assign a 0, 1, or 2 to each item. Test–retest reliability of the composite score is .92; internal consistency is .97 (downloaded March 22, 2004, from *Buros Mental Measurement Yearbook*, http://buros.unl.edu/buros/jsp/search.jsp). Whereas syndrome scores (for attention, conduct, etc.) can also be derived, they are less reliable. Hence, only a composite score was used (hereafter referred to as "Achenbach").

Fidelity of Intervention. Fidelity of intervention implementation was evaluated during the earlier study and again during the present study. To determine the quality of implementation of the *Ladders* activities in kindergarten, staff in the earlier study observed teachers' lessons and gave them a weekly global rating ranging from 1 (*poor*) to 3 (*excellent*). The ratings were designed to address (a) lesson clarity; (b) how well the teacher's instruction fit the intent of the lesson; and (c) the degree to which all students were engaged, including low-achieving and special education students. Prior to conducting the ratings, all staff scored videotaped instruction and obtained 100% agreement with a designated expert rater.

Staff in the earlier study used checklists to evaluate the accuracy of teachers' and students' implementation of K-PALS and first-grade PALS. On two occasions in both kindergarten and first grade, staff observed three randomly chosen student pairs in each PALS classroom. On the checklists, behavior was scored as *demonstrated*, *not demonstrated*, or *not applicable*. Each observation yielded a teacher score and an averaged student score. To create an overall classroom score, teacher and student fidelity scores were combined. To establish interrater agreement, two staff members simultaneously observed students and teachers in four classrooms in each of the 2 study years. Interrater agreement was consistently higher than 98%. Finally, staff associated with the present nonresponder study observed each student who participated in first-grade PALS to evaluate his or her individual fidelity of implementation (see later for details).

Procedure

Training. Research staff included five graduate students with teaching and assessment experience who were unaware of treatment conditions and responsiveness/nonresponsiveness status. Staff was trained to administer and score all measures according to instructions in the user manuals. Training occurred in two sessions, totaling 4 hours. Staff was also trained to establish rapport with children in about 10 min prior to administering the measures because examiner unfamiliarity can depress the test performance of students in high-poverty schools (e.g., D. Fuchs & Fuchs, 1986) and of young children with disabilities (e.g., D. Fuchs, Fuchs, Power, & Dailey, 1985).

Testing and Scoring. Measures used to identify responsiveness to intervention were administered after intervention: Kindergarten responsiveness was based on RLS and phoneme segmentation measures that were administered in the spring of kindergarten; first-grade responsiveness was based on the oral reading fluency measure that was administered in the spring of first grade. All but one measure used to predict responsiveness to intervention was administered after intervention in the spring of first grade. The exception, the RLN, was administered in the fall of kindergarten prior to intervention.

Staff tested the students individually in one session outside the students' classrooms in a quiet hallway. Measures were administered in random order and were scored by the examiner and rescored by a second staff member. Students' first-grade teachers completed the Achenbach checklist.

Third-Grade Follow-Up. Among the 104 students tested in kindergarten and first grade, 50 were located and tested again in May of what should have been their third-grade year. We found 13 of 34 nonresponsive students (19 students moved, 1 was in a private school for students with LD, and the family of 1 child did not give consent for testing); 14 of 26 sometimes responsive students (11 students moved and 1 was home schooled); and 23 of 44 always responsive students (16 students moved, 3 were home schooled, and 2 students' families withheld permission). We found these students in second- and third-grade classrooms. Examiner training, measures employed, and testing procedures were virtually identical to the procedures followed in kindergarten and first grade.

Data Analysis

First, to determine whether responsiveness to intervention was related to the fidelity with which the interventions were implemented, ANOVAs and post hoc tests were conducted. Second, the percentages of nonresponsive, sometimes responsive, and always responsive students were calculated. Third, the characteristics of nonresponsive students were compared with those of sometimes responsive and always responsive students using MANOVA and discriminant function analysis. The discriminant function analysis was used to determine how accurately children's responsiveness status could be predicted based on their characteristics and on whether they received treatment (Stevens, 1986). Due to this study's exploratory nature, we used a conservative approach to discriminant function analysis (Huberty, 1994): We screened the variables; cross-validated the classifications; and used kappa, which provides a chance-corrected index of agreement between actual and predicted group membership (Landis & Koch, 1977). Furthermore, we examined the standardized and structural coefficients to determine the relative importance of the measures. Finally, we examined the long-term relationship between responsiveness to intervention and reading difficulties at the end of third grade using a chi-square test. Statistical significance was determined as $p < .05$, using a two-tailed statistical test.

Results

Fidelity of Intervention Implementation

Information on the number of *Ladders* activities conducted was taken from teacher calendars—that is, the records they kept on treatment implementation during the study. Data on the quality of *Ladders* implementation came from direct observations. Teachers reported spending between 5 and 15 min per *Ladders* lesson three or more times per week, as recommended in the earlier study. The mean quality of implementation ratings for kindergarten teachers ranged from 1.8 to 2.8 on the 3-point scale. Table 3 displays *Ladders* fidelity scores by responsiveness/nonresponsiveness status.

We evaluated the accuracy of PALS implementation five times across kindergarten and first grade (see Table 3). During the earlier study, the mean combined teacher and student K-PALS fidelity scores ranged from 71% to 88% in the fall and from 59% to 79% in the spring (see Note 2). First-grade PALS fidelity scores ranged from 83.82% to 91.21% in the fall and from 84.55% to 86.37% in the spring. As mentioned earlier, an additional fidelity check during the present study was performed for each of the first-grade PALS intervention students. Their scores ranged from 81.91% to 88.05% (see Table 3).

To determine whether responsiveness to intervention was associated with the fidelity with which the intervention was implemented, univariate ANOVAs were conducted on quality of *Ladders* implementation and PALS fidelity scores. Responsiveness status (nonresponsive vs. sometimes responsive vs. always responsive) was a between-group factor. There were statistically significant differences in average quality ratings for *Ladders*, $F(2, 67) = 7.42$, $p < .001$; and for combined teacher–student fidelity of first-grade PALS in the fall, $F(2, 36) = 12.70$, $p < 001$. There were no statistically significant differences for fall K-PALS, $F(2,$

TABLE 3
Combined Teacher and Student Fidelity of Ladders, K-PALS, and First-Grade PALS

		Nonresponsive			Sometimes responsive			Always responsive			Total	
Intervention	**Time**	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***M***	***SD***
Ladders	Yearly average[a]	8	2.10	(0.29)	12	2.39	(0.21)	17	2.39	(0.22)	2.34	(0.25)
K-PALS	Fall	1	71.00	(0.00)	7	86.33	(7.94)	10	88.00	(8.81)	86.78	(8.74)
K-PALS	Spring	1	59.00	(0.00)	7	77.67	(10.79)	8	79.21	(11.72)	78.00	(11.56)
First-grade PALS	Fall	5	83.82	(4.26)	4	86.22	(6.07)	8	91.21	(2.51)	87.97	(5.14)
First-grade PALS	Spring	5	84.55	(5.43)	4	86.22	(5.54)	8	86.37	(5.61)	85.82	(5.46)
First-grade PALS	Spring[b]	5	81.91	(8.92)	4	84.56	(9.42)	8	88.05	(11.33)	85.51	(10.37)

Note. "n" signifies the number of classrooms for which nonresponsive, sometimes responsive, and always responsive students came. Looking at the second row from the top (K-PALS, fall implementation), the combined teacher and student implementation score for the one class from which nonresponders came was 71%; the mean combined implementation score for the seven classes with which sometimes responsive students were associated was 86.33%; and the mean combined implementation score from which came always responsive students was 88%.
[a]"Yearly average" is the mean score for all observations across the year. Possible *Ladders* ratings ranged from 0 to 3, with 0 as the lowest rating and 3 as the highest rating. [b]An additional fidelity check of participants in the present study.

29) = 1.91, *ns*; for spring K-PALS, $F(2, 29) = 1.51$, *ns*; for spring first-grade PALS, $F(2, 36) = .41$, *ns*; or for the individual first-grade PALS fidelity score of the participants in the present study, $F(2, 36) = 1.29$, *ns*.

Post hoc pairwise comparisons using the Tukey HSD indicated that nonresponsive students were in classrooms in which *Ladders* activities across the kindergarten year were implemented with significantly lower quality than in the classrooms of sometimes ($p < .01$) and always ($p < .01$) responsive students. Similarly, post hoc comparisons indicated that nonresponsive students were in classrooms with significantly lower fidelity for first-grade PALS implementation in the fall than were always responsive students ($p < .01$). Sometimes responsive students were also in first-grade PALS classrooms with lower implementation fidelity in the fall than always responsive students ($p < .01$).

End of First Grade Findings

Percentages of Nonresponders. We calculated the percentages of nonresponders (i.e., students who did not respond to intervention during kindergarten or in first grade) in relation to the number of students who participated in the earlier study. In the earlier study, 227 students received intervention for 1 or 2 years, and 7.05% ($n = 16$) were nonresponders. In contrast, of 71 control students, 25.35% ($n = 18$) were nonresponders. There was only 1 nonresponder who received K-PALS plus *Ladders*, and she had special needs and an IEP. Of all the students with IEPs in the earlier study, 26.67% of intervention children were nonresponsive, compared to 66.76% of control students.

Characteristics of Nonresponders. To understand the characteristics of the nonresponders, we first conducted a MANOVA that included responsiveness status (always responsive vs. sometimes responsive vs. nonresponsive) and the amount and timing of the intervention (kindergarten and first grade vs. first grade only vs. kindergarten only vs. control). The nine dependent variables were students' performance on three pretreatment measures (RLN, RLS, and Segmentation) and six posttreatment measures (Achenbach, Grammatical Closure, PPVT-R, Sentence Imitation, Word Discrimination, and Word Sequences). Two students did not complete the Word Discrimination test, and the analysis was conducted using only complete cases (Tabachnick & Fidell, 1996). The Wilks' Λ test statistic indicated a statistically significant multivariate main effect for responsiveness status, $F(9, 164) = 5.29$ $p < .001$; absence of a main effect for intervention, $F(27, 240.13) = .76$, *ns;* and absence of an interaction between student responsiveness status and intervention, $F(54, 422.71) = 1.00$, *ns*. In Table 4, means and standard deviations for these nine dependent variables are displayed separately for nonresponsive, sometimes responsive, and always responsive students.

There were statistically significant effects for responsiveness status on Achenbach, $F(2, 90) = 8.04$, $p < .001$; Grammatical Closure, $F(2, 90) = 24.68$, $p < .001$; PPVT-R, $F(2, 90) = 18.19$, $p < .001$; pretreatment RLN, $F(2, 90) = 12.06$, $p < .001$; pretreatment RLS, $F(2, 90) = 9.26$, $p < .001$; pretreatment Segmentation, $F(2, 90) = 3.98$ $p < .01$; Sentence Imitation, $F(2, 90) = 6.73$, $p < .01$; and Word Discrimination, $F(2, 90) = 9.28$, $p < .001$. Word Sequences was not significant, $F(2, 90) = 1.45$, and was not

TABLE 4
Characteristics by Responsiveness Status and Intervention

Variable	Kindergarten and first grade		First grade		Kindergarten		Controls		Overall	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Nonresponsive students (*n* = 34)										
Achenbach[a]	48.43	(36.77)	50.75	(35.01)	19.00	(10.42)	37.89	(22.98)	38.79	(27.19)
Grammatical Closure[b]	10.43	(7.30)	6.50	(7.85)	8.40	(7.37)	13.87	(6.69)	10.76	(7.43)
PPVT-R[c]	62.43	(16.10)	54.75	(5.91)	62.00	(9.77)	64.31	(13.81)	62.34	(12.94)
Sentence Imitation[d]	6.00	(3.74)	7.00	(2.45)	6.00	(3.08)	6.75	(2.67)	6.50	(2.85)
Word Discrimination[b]	13.43	(5.38)	13.75	(7.18)	11.00	(4.12)	13.38	(4.83)	13.06	(4.99)
Word Sequences[d]	7.14	(4.53)	7.00	(1.63)	5.20	(3.27)	6.43	(2.89)	6.47	(3.17)
Rapid Letter Naming[e]	10.14	(7.86)	6.75	(4.79)	14.80	(9.42)	7.94	(9.13)	9.26	(8.58)
Rapid Letter Sound[f]	1.29	(1.60)	0.00	(0.00)	1.00	(1.73)	0.72	(1.32)	0.79	(1.37)
Segmentation[g]	0.57	(0.79)	0.00	(0.00)	2.00	(3.46)	1.72	(3.25)	1.43	(2.72)
Sometimes responsive students (*n* = 26)										
Achenbach[a]	20.00	(24.29)	10.50	(9.04)	37.07	(29.26)	6.00	(2.00)	26.12	(26.68)
Grammatical Closure[b]	11.60	(10.16)	20.75	(6.08)	15.07	(5.27)	17.33	(6.66)	15.54	(6.85)
PPVT-R[c]	66.20	(19.97)	91.00	(11.80)	77.71	(17.72)	92.33	(11.85)	79.23	(18.25)
Sentence Imitation[d]	7.40	(2.88)	9.25	(3.10)	9.29	(3.54)	10.33	(1.53)	9.04	(3.16)
Word Discrimination[b]	14.40	(4.16)	17.75	(0.50)	16.86	(1.83)	17.67	(1.53)	16.62	(2.47)
Word Sequences[d]	5.20	(2.59)	6.50	(2.65)	7.00	(3.64)	7.33	(3.06)	6.62	(3.18)
Rapid Letter Naming[e]	10.00	(9.75)	17.75	(10.81)	13.43	10.01	21.00	(14.80)	14.31	(10.50)
Rapid Letter Sound[f]	4.20	(4.92)	5.00	(4.83)	1.57	(1.60)	10.67	(4.93)	3.65	(4.32)
Segmentation[g]	1.60	(2.51)	3.00	(1.83)	1.21	(1.81)	3.67	(2.31)	1.85	(2.09)
Always responsive students (*n* = 44)										
Achenbach[a]	11.63	(12.37)	25.67	(32.35)	12.67	(18.58)	6.50	(4.51)	12.61	(16.73)
Grammatical Closure[b]	20.13	(4.18)	24.33	(2.08)	20.57	(4.17)	21.50	(4.20)	20.75	(4.09)
PPVT-R[c]	79.19	(11.34)	81.67	(13.87)	85.05	(14.32)	86.75	(9.74)	82.84	(12.83)
Sentence Imitation[d]	8.88	(3.90)	9.33	(2.52)	10.76	(3.51)	10.25	(2.87)	9.93	(3.56)
Word Discrimination[b]	16.56	(2.28)	16.33	(2.52)	16.33	(2.58)	17.75	(1.26)	16.55	(2.34)
Word Sequences[d]	7.38	(3.10)	7.67	(4.51)	9.62	(4.47)	7.50	(3.11)	8.48	(3.93)
Rapid Letter Naming[e]	26.38	(10.47)	13.33	(11.93)	27.67	(14.43)	32.25	(7.18)	26.64	(12.70)
Rapid Letter Sound[f]	7.00	(6.34)	3.67	(4.04)	8.67	(8.59)	10.50	(9.57)	7.89	7.61
Segmentation[g]	4.13	(3.70)	4.33	(4.51)	3.52	(4.59)	4.50	(3.70)	3.89	4.59

[a]*Achenbach Child Behavior Checklist* (Achenbach, 1994). [b]*Test of Language Development 2–Primary* (Hammill & Newcomer, 1988). [c]*Peabody Picture Vocabulary Test–Revised* (Dunn & Dunn, 1981). [d]*Detroit Tests of Learning Aptitude* (Hammill, 1991). [e]*Rapid Letter Naming Test* (D. Fuchs, Fuchs, Thompson, Al Otaiba, Yen, Yang et al., 2001). [f]*Rapid Letter Sound Test* (Levy & Lysynchuk, 1997). [g]*Yopp-Singer Segmentation Test* (Yopp, 1995).

included in the subsequent discriminant analysis reported later. Post hoc pairwise comparisons were conducted using the Tukey HSD, which revealed that nonresponsive students had significantly lower scores than always responsive students on Grammatical Closure ($p < .001$), PPVT-R ($p < .001$), pretreatment RLN ($p < .001$), pretreatment RLS ($p < .001$), pretreatment Segmentation ($p < .05$), Sentence Imitation ($p < .001$), and Word Discrimination ($p < .001$). Nonresponsive students had significantly higher (i.e., more negative) scores on the Achenbach ($p < .001$).

The follow-up multiple comparisons revealed a similar pattern of differences involving nonresponsive and sometimes responsive students. Nonresponsive students scored lower on Grammatical Closure ($p < .05$), PPVT-R ($p < .001$), pretreatment RLN ($p < .05$), pretreatment RLS ($p < .01$), Sentence Imitation, ($p < .01$), and Word Discrimination, ($p < .001$). There was no reliable difference between the two groups on the Achenbach. The multiple comparisons also revealed that sometimes responsive students scored lower than always responsive students on Grammatical Closure ($p < .001$), pretreatment RLN ($p < .001$), pretreatment RLS ($p < .01$), and pretreatment Segmentation ($p < .05$), and scored higher (i.e., more negatively) than al-

ways responsive students on the Achenbach ($p < .05$). However, there were no reliable differences between these groups on PPVT-R, Sentence Imitation, or Word Discrimination.

We computed *ES*s as the difference between the means divided by the pooled standard deviation (Hedges & Olkin, 1985). As shown in Table 5, the largest *ES*s in absolute value ranged from 0.58 to 1.72 for the always responsive versus the nonresponsive group. For the sometimes responsive versus the nonresponsive group, *ES*s ranged from 0.08 to 1.10; for the sometimes responsive versus the always responsive students, they ranged from 0.03 to 1.03.

Discriminant Function Analysis. A discriminant function analysis was used to determine which variables reliably predicted nonresponsiveness. The variables were screened to eliminate those that (a) were statistically nonsignificant in the overall MANOVA (i.e., Word Sequences), (b) correlated highly with measures used to define responsiveness (e.g., RLS correlated highly with RLN), (c) correlated highly with other child characteristic variables (e.g., Grammatical Closure was associated strongly with both PPVT-R and Sentence Imitation), and (d) had non-normal distributions (pretreatment Segmentation and Word Discrimination). Thus, assumptions of normality and homogeneity of variance were met. Table 6 shows the correlations among pretreatment measures and measures used to define responsiveness. Table 7 displays the correlations among the child characteristic measures.

Following this screening process, the four remaining child characteristic variables (i.e., pretreatment RLN, Achenbach, PPVT-R, and Sentence Imitation) and the amount and timing of intervention were entered into a discriminant function analysis to determine which combinations of variables best predicted responsiveness status (Huberty, 1994; Tabachnick & Fidell, 1996). The analysis had an adequate sample size to variable ratio. The Box M Test for equivalence of covariance matrices between groups showed no statistically significant difference ($p = .10$), which indicates that the distribution of scores on the measures met the assumption of multivariate normality suggested for discriminant function analysis.

Given our three groups (nonresponsive, sometimes responsive, and always responsive), two discriminant functions were possible, because the number of functions equals the number of groups minus one. With regard to classification into groups, although unequal sample sizes are not a problem in discriminant function analysis, a conservative approach is recommended to adjust for them (e.g., Tabachnick & Fidell). This is important because in discriminant function analysis, classification into groups is based on two probabilities: *prior* likelihood, or the prior probability of group membership (e.g., with three equal groups, a student would have a 33% chance of being in any given group), and *posterior* likelihood (i.e., the functions representing performance on the predictor variables). Prior values, therefore, were adjusted in proportion to the groups' sample sizes (nonresponsive, $n = 34$; sometimes responsive, $n = 26$; always responsive, $n = 44$).

The discriminant function analysis yielded high classification rates in predicting responsiveness to intervention. These rates were cross-validated using a leave-one-out (i.e., jackknife) method (Huberty, 1994). Table 8 shows both the original and cross-validated classification rates. Using the more conservative cross-validation method,

TABLE 5
Effect Sizes for Child Characteristic Measures by Responsiveness Status

	Always responsive vs. nonresponsive	Sometimes responsive vs. nonresponsive	Always responsive vs. sometimes responsive
Achenbach[a]	–1.20 ***	–0.47	–0.65*
Grammatical Closure[b]	1.72***	0.67*	0.99***
PPVT-R[c]	1.60***	1.10***	0.24
Sentence Imitation[d]	1.06***	0.86***	0.26
Word Discrimination[b]	0.94***	0.87***	0.03
Word Sequences[d]	0.58	0.08	0.51
Rapid Letter Naming[e]	1.57***	0.53*	1.03***
Rapid Letter Sound[f]	1.23***	0.95***	0.64**
Segmentation[g]	0.66**	0.21	0.53*

[a]*Achenbach Child Behavior Checklist* (Achenbach, 1994). [b]*Test of Language Development 2–Primary* (Hammill & Newcomer, 1988). [c]*Peabody Picture Vocabulary Test–Revised* (Dunn & Dunn, 1981). [d]*Detroit Tests of Learning Aptitude* (Hammill, 1991). [e]*Rapid Letter Naming Test* (D. Fuchs, Fuchs, Thompson, Al Otaiba, Yen, Yang, et al., 2001). [f]*Rapid Letter Sound Test* (Levy & Lysynchuk, 1997). [g]*Yopp-Singer Segmentation Test* (Yopp, 1995).
*$p < .05$. **$p < .01$. *** $p < .001$.

group membership was correctly predicted for 82.40% of nonresponsive, 30.80% of sometimes responsive, and 84.10% of always responsive students. Overall, 70.20% of the students were correctly classified through the cross-validated method. The kappa approach (Cohen, 1960) was used to further test whether the relationship between the actual group membership and the predicted group membership was due to chance. The value of kappa was .64, which represents respectable agreement beyond chance (Landis & Koch, 1977).

Table 9 shows the standardized and structural coefficients. Standardized coefficients are products of the raw correlation coefficient of each variable and the standard deviation and are analogous to the standardized coefficient in multiple regression. Therefore, in common with beta weights, standardized coefficients provided a standardized metric that allowed a direct comparison of the relative importance or degree to which each variable contributed to the function (Huberty, 1994; Pedhazur, 1982). Standardized coefficients also showed whether variables were redundant across the two functions.

An examination of the structural coefficients, which are analogous to factor loadings in factor analysis, revealed the correlation between each variable and the discriminant function. The first discriminant function was significant ($p < .001$) and accounted for 49% of the variance between groups. The second function was not significant but explained an additional 7% of the variance. Inspection of the structural coefficients indicated that pretreatment RLN (i.e., naming speed), PPVT-R (i.e., vocabulary), Achenbach

TABLE 6
Correlations among Pretreatment Measures and Measures Used to Define Responsiveness Status

Measures	1	2	3	4	5
Pretreatment					
1. Segmentation[a]	—				
2. Rapid Letter Sound[b]	0.61*	—			
3. Rapid Letter Naming[c]	0.36*	0.60*	—		
Defining nonresponsiveness					
4. Growth in Segmentation	0.10	0.29*	0.41*	—	
5. Growth in Rapid Letter Sound	0.20*	0.33*	0.37*	0.82*	—

[a]*Yopp-Singer Segmentation Test* (Yopp, 1995). [b]*Rapid Letter Sound Test* (Levy & Lysynchuk, 1997). [c]*Rapid Letter Naming Test* (D. Fuchs, Fuchs, Thomson, Al Otaiba, Yen, Yang, et al., 2001).
*$p < .01$.

TABLE 7
Correlations Among Child Characteristics Measures

Measures	1	2	3	4	5
1. PPVT-R[a]	—				
2. Word Discrimination[b]	0.35*	—			
3. Sentence Imitation[c]	0.62*	0.35*	—		
4. Grammatical Closure[b]	0.72*	0.37*	0.51*	—	
5. Achenbach[d]	–0.32*	–0.27	–0.10	–0.34*	—

[a]*Peabody Picture Vocabulary Test–Revised* (Dunn & Dunn, 1981). [b]*Test of Language Development 2–Primary* (Hammill & Newcomer, 1988). [c]*Detroit Tests of Learning Aptitude* (Hammill, 1991). [d]*Achenbach Child Behavior Checklist* (Achenbach, 1994).
* $p < .01$.

TABLE 8
Original and Cross-Validated Classification Rates by Responsiveness Status

	Predicted group membership					
	Nonresponsive		Sometimes responsive		Always responsive	Total
Student status	*n*	%	*n*	%	*n* %	*n* %
Original Classification						
Nonresponsive	29	(85.30%)	4	(11.80%)	1 (2.90%)	34 (100%)
Sometimes responsive	7	(26.90%)	9	(34.60%)	10 (38.50%)	26 (100%)
Always responsive	4	(9.10%)	3	(6.80%)	37 (84.10%)	44 (100%)
Cross-Validated Classification						
Nonresponsive	28	(82.40%)	4	(11.80%)	2 (5.90%)	34 (100%)
Sometimes responsive	8	(30.80%)	8	(30.80%)	10 (38.50%)	26 (100%)
Always responsive	4	(9.10%)	3	(6.80%)	37 (84.10%)	44 (100%)

(i.e., problem behavior), and Sentence Imitation (i.e., phonological encoding with syntactic awareness) correlated with Function 1; and RLN and PPVT-R correlated with Function 2.

As indicated earlier, we tested 50 of the original 104 students at the end of what should have been third grade for these children. Table 10 indicates whether they had reading difficulties at that time. There was a statistically significant difference among nonresponsive, sometimes responsive, and always responsive students in terms of school-identified reading problems 2 years beyond first grade, $\chi^2(6, n = 50) = 40.60$, $p < .001$.

Discussion

The purpose of this study was to provide a relatively comprehensive and accurate description of nonresponders to generally effective beginning reading interventions. We took steps to extend prior work, both methodologically and conceptually. First, we used a long-term definition of responsiveness/nonresponsiveness to try to produce more stable information about the number and characteristics of nonresponders. Second, we trained teachers to implement the reading interventions in their classrooms to create a more realistic context for the study. Third, we monitored their fidelity of treatment implementation and that of their students to eliminate improper implementation as a contributing and complicating cause of nonresponsiveness. Fourth, we selected a broad range of child measures associated with nonresponsiveness in previous research. Fifth, we conducted a 2-year follow-up of our sample to determine how many nonresponders would eventually be identified as requiring special education because of a reading disability.

Number of Nonresponders Over Time

Of the 227 students who received intervention (in kindergarten-only, first grade-only, or both kindergarten and first grade) as part of the earlier study, only 7.05% were nonresponsive. A much higher proportion (25.35%) of the 71 control students was nonresponsive. This finding corroborates prior investigations showing that K-PALS plus *Ladders to Literacy* is a generally effective beginning reading program (e.g., Fuchs, Fuchs, Thomson, Al

TABLE 9
Standardized Canonical and Structural Coefficients for the Measures

	Function 1		Function 2	
	All predictors		All predictors	
Measure	Standardized	Structural	Standardized	Structural
Achenbach[a]	–0.47	–0.44	0.10	0.08
Treatment	0.45	0.25	0.19	0.08
PPVT-R[b]	0.39	0.54	0.76	0.69
Rapid Letter Naming[c]	0.60	0.62	–0.69	–0.60
Sentence Imitation[d]	0.25	0.41	0.10	0.37

[a]*Achenbach Child Behavior Checklist* (Achenbach, 1994). [b]*Peabody Picture Vocabulary Test–Revised* (Dunn & Dunn, 1981). [c]*Rapid Letter Naming* (D. Fuchs, Fuchs, Thompson, Al Otaiba, Yen, Yang, et al., 2001). [d]*Detroit Tests of Learning Aptitude* (Hammill, 1991).

TABLE 10
Third-Grade Educational Placement by Responsiveness Status

	Nonresponsive		Sometimes responsive		Always responsive		Total	
Educational placement	Intervention	Controls	Intervention	Controls	Intervention	Controls	Intervention	Controls
Third grade, no reading difficulties	0	2	21	2	11	2	32	6
Third grade, reading difficulties (receiving tutoring and/or student study team in progress)	1	2	0	0	0	0	1	2
Third grade; IEP has reading goals	4	1	0	0	0	0	4	1
Second grade; retained due to reading difficulties	0	4	0	0	0	0	0	4
Total	5	9	21	2	11	2	37	13

Note. *Intervention* refers to intervention during kindergarten, first grade, or kindergarten and first grade.

Otaiba, Yen, Yang, et al., 2001; Fuchs, Fuchs, Thompson, Al Otaiba, Yen, Yang, et al., 2002); it underscores the importance of such programs in mainstream classrooms; and it supports Vellutino et al.'s (1996) warning that many children with apparent reading disabilities are really students without reading disabilities who are in need of competent early literacy instruction.

Among the children with reading disabilities, the proportion of nonresponders was higher. Of those with IEPs in the earlier study, 26.67% of intervention students were nonresponsive, as compared to 66.76% of controls. Many children with reading disabilities require an intensity and systematicity of instruction that exceeds what PALS, *Ladders*, or most other best practices developed for mainstream classrooms can offer. This fact was substantiated not long ago by a best-evidence synthesis of research on the efficacy of peer tutoring for students with high-incidence disabilities (Mathes & Fuchs, 1994). Peer tutoring was more effective than typical reading instruction, but it was less effective than one-to-one teacher tutoring or teacher-led small-group instruction.

Virtually our entire sample (91.89%)—and literally all children with IEPs—who were nonresponsive to kindergarten intervention remained nonresponsive in first grade. At third grade, four out of five children who had been nonresponsive to kindergarten intervention had IEP goals in reading. The remaining student was receiving tutoring, and his case was before a school study team. Of nine control students who were nonresponsive to classroom instruction in kindergarten, two had no reading difficulties, one had IEP goals in reading, four had been retained, and two were receiving tutoring (and one of these two had a school study team meeting pending). To our knowledge, this 2-year follow-up is unique in the sense that it may be the first study to indicate that nonresponsiveness to generally effective interventions is a valid indicator of long-term reading disability—a proposition expressed 10 years ago by Vellutino et al. (1996). Responsiveness to generally effective intervention seems like it would be an important component in a more comprehensive approach to identify children who require more intensive instruction and possibly special education.

Characteristics of Nonresponders

To better understand our nonresponders, we looked for statistically reliable and educationally meaningful differences between the characteristics of nonresponsive and always responsive students. Nonresponsive students scored about 1.5 *SD* lower than always responsive students on measures of vocabulary, rapid naming, and problem behavior; they were at least 1.0 *SD* lower on measures of verbal memory. Performance on all these measures, in addition to the amount of intervention that children had received (none vs. one year vs. two years), correctly distinguished a large majority of children who were nonresponsive (82.4%) and always responsive (84.1%).

These measures provide a preliminary but promising set of markers to help practitioners to identify students who are likely to be chronically nonresponsive to generally effective interventions and to target them for most intensive preventive instruction. Generally, these markers are consistent with findings from a small but growing number of investigations exploring the relations between child characteristics and responsiveness to intervention. A recent and pertinent meta-analysis (Nelson et al., 2003) reported moderately large *ES*s for rapid naming (0.51) and problem behavior (0.46) and smaller *ES*s for memory (0.31), IQ (0.26), and demographic factors (0.07).

Vocabulary. Our nonresponders performed poorly on the PPVT-R, a measure of receptive vocabulary. Poor vocabulary is strongly associated with other weak verbal skills (e.g., Lonigan, Burgess, & Anthony, 2000; Scarborough, 2002) and predictive of poor reading ability (Catts, Fey, Zhang, & Tomblin, 1999; Tabors, Snow, & Dickinson, 2001). The apparent importance of weak performance on the PPVT-R deserves further comment, especially in light of the No Child Left Behind Act of 2001. Under this act, states are required to ensure adequate reading progress for all students. Accordingly, schools are strongly encouraged to offer scientifically based (or evidence-based) instruction. Yet, the body of scientifically based reading research (National Reading Panel, 2000; Snow et al., 1998) has often excluded children with disabilities, such as those with a verbal IQ below 70. Relatively little is known about what constitutes effective reading practices and adequate progress for this subgroup.

Some have argued that verbal ability, or IQ in general, is not related to responsiveness to intervention. Fletcher, Lyon, Barnes, et al. (2002), for example, wrote, "The concept of IQ as it is applied to LD is outmoded and reflects an obsolete practice. . . . IQ tests do not measure aptitude for learning or provide an index of response to intervention" (p. 234; see also Fletcher, Morris, & Lyon, 2003). To the extent to which we may legitimately jump from our nonresponders' poor PPVT-R performance to verbal ability to IQ (and at least some would counsel that these jumps are too many and too large), our findings suggest that it is premature to conclude that verbal ability and IQ are *not* related to responsiveness. In others' work, the importance of verbal ability and IQ to responsiveness has often been explored for a rather restricted range of children. Vellutino et al. (1996), for example, studied only those children with IQ scores greater than 90. Restrictions in range, of course, diminish the likelihood of obtaining statistically significant findings. In contrast to much of the prior research, we included students with disabilities and, partly for this reason, we had a more heterogeneous sample with a broader range of verbal skills.

The presence or absence of a relationship between verbal ability or IQ and nonresponsiveness to generally effective intervention may also depend on the researchers' definitions of nonresponsiveness. In a recent review of 13 studies, Fuchs and Young (2006) found that IQ was a better predictor when the dependent measure was defined as "reading comprehension" than when it represented "word reading" or "phonological awareness." In this study, we used fluency measures rather than word-level reading measures to define responsiveness/nonresponsiveness. That vocabulary was related to fluency measures is consistent with a view of reading development whereby verbal skills provide support for fluent reading and comprehension (Wolf & Bowers, 1999). Scarborough (2002), for example, has argued that "'reading comprehension' deficits are essentially oral language limitations" (p. 98). In short, whether vocabulary, verbal ability, or IQ distinguishes responders from nonresponders may be affected by the range of performance of the students studied and by how their responsiveness/nonresponsiveness is defined. Their importance is no doubt further influenced by the nature of the intervention and the fidelity with which it is conducted. At present, the claim that verbal ability or IQ is unrelated to responsiveness/nonresponsiveness may be insufficiently nuanced.

Problem Behavior. The relationship between problem behavior (as rated by teachers on the Achenbach) and nonresponsiveness corroborates findings from prior studies (O'Shaugnessy & Swanson, 2000; Snider, 1997; Uhry & Shepherd, 1997; Vellutino et al., 2000; Vellutino et al., 1996). Torgesen et al. (1999) and Vadasy et al. (1997), for example, reported that many nonresponders demonstrated inattention and conduct problems even when instruction was delivered in one-to-one tutorials. We, too, have observed this in recent efforts to provide nonresponders with more intensive reading instruction in groups of two and three children. The distractibility or conduct problems of some are so severe that much of the tutoring time is lost to behavior management. It should come as no surprise, therefore, that levels of reading achievement among special education students with severe behavior disorders in elementary school are often 1.5 to 2 grade levels lower than those of their peers; they are lower still when these students reach high school (e.g., Kauffman, Cullinan, & Epstein, 1987).

This prompts at least two thoughts. First, nonresponsiveness to generally effective early reading instruction may well reflect the co-occurrence of two or more disabilities for many children. Pennington's (2004) work has suggested that reading disabilities and attention-deficit/hyperactivity disorder (ADHD) are often comorbid. Additional research like Pennington's should focus on the nature of comorbidity to better understand nonresponsiveness. Second, such research may lead to more comprehensive treatments that address multiple dimensions of disability (e.g., reading disability and ADHD) rather than a single dimension. Reading interventions, for example, may be combined with cognitive behavior modification or self-regulation strategies (e.g., Graham & Harris, 2005). If such comprehensive and complex treatments prove effective, a future question will be, How can they be made sufficiently practical for use in schools?

Race, Age, Socioeconomic Status, and Memory. Finally, there were no differences between responders and nonresponders on race, age, or socioeconomic status. In this respect, our study differs from that of Foorman et al. (1997), who found that even in the most effective reading treatment, Hispanic children achieved less growth than native English speakers. None of the children in our sample were English language learners. We also failed to find reliable differences between our groups on verbal memory for abstract words or on phonological discrimination. Vellutino et al. (1996), in contrast, reported differences between responsive and nonresponsive students on a measure assessing immediate and delayed recall of abstract words. These differences might be explained by the fact that our measure, the Word Sequences subtest of the DTLA-3, did not require delayed recall.

Study Limitations

We wish to emphasize several important study limitations. First, our interpretation of intervention responsiveness may well be confounded by the significantly lower fidelity of intervention implementation for nonresponsive students than for always responsive students. This was true during two of six fidelity checks—the first in kindergarten for *Ladders* activities, and the second in the fall of first grade for PALS. Thus, it is unclear whether the lower fidelity of intervention implementation disadvantaged students from the start, and whether differences in implementation rather than differences in child characteristics were associated with student responsiveness status.

The second limitation relates to our sample size. To maximize the number of students for our various analyses, we combined two kindergarten interventions (PALS plus *Ladders* and *Ladders* only) despite the fact that one (PALS plus *Ladders*) appeared to be more effective than the other (*Ladders* only; see D. Fuchs, Fuchs, Thompson, Al Otaiba, Yen, Yang, et al., 2001). Thus, it is unclear whether nonresponsiveness to intervention was due to differences in the two interventions rather than to characteristics of the participants. Recall that at follow-up in third grade, only one student who participated in PALS plus *Ladders* was a nonresponder, whereas 11 students who received *Ladders* only were nonresponders. Unrelated but important, attrition during our 2-year follow up was considerable.

Third, several child characteristic measures (RLN, RLS, and Segmentation) were administered prior to treatment in the earlier study, whereas other measures (Achenbach, Grammatical Closure, PPVT-R, Sentence Imitation, Word Discrimination, and Word Sequences) were administered following the treatment implementation in the present nonresponder study. A given child's relative performance across these measures may thus have been influenced by historical, maturational, or other factors, possibly threatening the validity of our findings.

Fourth, general classroom instruction was not systematically observed. Fidelity of treatment data were obtained in the intervention classes, and intervention teachers' instruction was informally noted. In control classes, there was no formal and little informal observation. Hence, besides the type of basal reading curriculum, little is known about the quality of instruction in our study classes. Similarly, we cannot provide information about reading instruction for any of the children in second and third grades beyond reporting the type of basal reading curriculum. Finally, our results are tied to our definition of nonresponsiveness as well as to our specific intervention; different definitions of nonresponsiveness and different interventions may well yield different numbers and different types of nonresponders (see Al Otaiba & Fuchs, 2002).

Conclusions

These limitations notwithstanding, the primary contribution of our study is in extending the knowledge base on nonresponders to more typical school contexts in which classroom teachers provided early literacy intervention to heterogeneously grouped students. Such research is critically important in light of the No Child Left Behind requirements to help all children learn to read at grade level. We have shown that a generally well-implemented, systematic, explicit, peer-mediated intervention, targeting phonological and alphabetic awareness and supplemented by teacher-directed phonological awareness training, can substantially reduce the number of students at risk for reading problems. By following our children through third grade, we confirmed that none of the children who responded initially to this classroom instruction later developed reading difficulties.

In light of our findings and study limitations, there is a need for continued research to address the needs of children who do not respond to scientifically based classroom (or primary-level) instruction. Our nonresponders clearly required a secondary level of intervention, using a different method or combination of methods delivered with greater intensity than was available to our study participants. Given the heterogeneity of the nonresponders, we believe that this secondary level of instruction must not only be more intensive, but must be tailored to children's individual strengths and weaknesses. Perhaps the most humbling implication of our study is that although we and others have learned much about what is necessary to reduce the number of nonresponders and to understand some of the characteristics that make it difficult for them to learn to read in mainstream classrooms, we do not yet know nearly enough about which specific approaches are key to ensure that no child is left behind.

ABOUT THE AUTHORS

Stephanie Al Otaiba, *PhD, is an assistant professor of special education at Florida State University and is on the research faculty at the Florida Center for Reading Research. Her research interests include early intervention to prevent reading difficulties, teacher preparation and professional development, and cultural and linguistic diversity.* ***Douglas Fuchs,*** *PhD, holds the Nicholas Hobbs Chair in Special Education and Human Development at Peabody College of Vanderbilt University, where he is co-director of the National Research Center on Learning Disabilities, funded by the Office of Special Education Programs in Washington, DC. Address: Stephanie Al Otaiba, City Centre Building 227 N. Bronough St., Suite 7250, Tallahassee, FL 32301; e-mail salotaiba@fcrr.org.*

AUTHORS' NOTES

1. *The research described in this article was funded by Grants H324D000033 and H324B0049 from the Office of Special Education Programs in the U.S. Department of Education to Vanderbilt University. Points of view expressed in the article do not necessarily reflect the policies of the funding agency, and no official endorsement by it should be inferred.*
2. *Parts of this work were presented at meetings of the American Educational Research Association in New Orleans, LA (April, 2000), the Council for Exceptional Children in Kansas City, KA (April, 2001), the International Reading Association in New Orleans, LA (April 2001), and the Society for the Scientific Study of Reading in Chicago, IL (July 2002)*
3. *We thank Ebba Svenson, Kristen McMaster, Anneke Thompson, Loulee Yen, Nancy Yang, and Mary Braun for their invaluable assistance during study implementation, and Lynn Fuchs for her collegiality.*

NOTES

1. *K-PALS and first-grade PALS teacher manuals are available from the second author or at http://www.peerassistedlearningstrategies.net*
2. *The combined teacher–student fidelity score of 59% was for one dysfunctional class.*

REFERENCES

Achenbach, T. M. (1994). *Child behavior checklist.* Burlington, VT: University Medical Education Associates.

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

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education, 23,* 300–316.

Berninger, V. W., Abbott, R. D., Zook, D., Ogier, S., Lemos-Britton, Z., & Brooksher, R. (1999). Early intervention for reading disabilities: Teaching the alphabetic principle in a connectionist framework. *Journal of Learning Disabilities, 32,* 491–503.

Blachman, B. A. (1994). What we have learned from longitudinal studies of

phonological processing and reading, and some unanswered questions: A response to Torgesen, Wagner, and Rashotte. *Journal of Learning Disabilities, 27,* 287–291.

Blachman, B. (1997). Early intervention and phonological awareness: A cautionary tale. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 408–430). Mahwah, NJ: Erlbaum.

Brown, I. S., & Felton, R. H. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. *Reading and Writing: An Interdisciplinary Journal, 2,* 223–241.

Case, L. P., Speece, D. L., & Molloy, D. E. (2003). The validity of a response-to-instruction paradigm to identify reading disabilities: A longitudinal analysis of individual differences and contextual factors. *School Psychology Review 32(4), 447–582.*

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3,* 331–361.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Dickman, E. G., Hennessy, N. L., Moats, L. C., Rooney, K. J., & Tomey, H. A. (2002). *The International Dyslexia Association response to OSEP Summit on Learning Disabilities.* Unpublished paper.

Donovan, M. S., & Cross, C. T. (2002). *Minority students in special and gifted education.* Washington, DC: National Academy Press.

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

Farr, R. C., & Strickland, D. S. (1995). *Treasury of literature: First street.* Orlando, FL: Harcourt Brace.

Fletcher, J. M., & Foorman, B. R. (1994). Issues in definition and measurement of learning disabilities: The need for early intervention. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 185–200). Baltimore: Brookes.

Fletcher, J. M., Lyon, G. R., Barnes, M., Stuebing, K., Francis, D. J., Olson, R. K., et al. (2002). Classification of learning disabilities: An evidence-based evaluation. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities* (pp. 185–250). Mahwah, NJ: Erlbaum.

Fletcher, J. M., Morris, R. D., & Lyon, G. R. (2003). Classification and definition of learning disabilities: An integrative perspective. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 30–56). New York: Guilford Press.

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

Fuchs, D., & Fuchs, L. S. (1986). Test procedure bias: A meta-analysis of examiner familiarity effects. *Review of Educational Research, 56,* 243–262.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Lipsey, M. L. (2002). Is "learning disabilities" just a fancy term for low achievement? A meta-analysis of reading differences between low achievers with and without the label. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities* (pp. 737–762). Mahwah, NJ: Erlbaum.

Fuchs, D., Fuchs, L. S., Power, M. H., & Dailey, A. M. (1985). Bias in the assessment of handicapped children. *American Educational Research Journal, 22,* 185–198.

Fuchs, D., Fuchs, L. S., Svenson, E., Yen, L., Thompson, A., McMaster, K., et al. (2001). *Peer Assisted Learning Strategies: First grade reading.* Nashville, TN: Vanderbilt University.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., McMaster, K., et al. (2001). *Peer Assisted Learning Strategies: Kindergarten reading.* Nashville, TN: Vanderbilt University.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N., et al. (2001). Is reading important in reading-readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology, 93,* 251–267.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N., et al. (2002). Exploring the importance of reading programs for kindergartners with disabilities in mainstream classrooms. *Exceptional Children, 68,* 295–311.

Fuchs, D., Thompson, A., Al Otaiba, S., Yen, L., Svenson, E., Braun, M., & Yang, N. (2000, February). *Exploring effective and efficient reading readiness programs in Title I and non-Title I schools.* Paper presented at the annual Pacific Coast Research Conference, La Jolla, CA.

Fuchs, D., & Young, C. (2006). On the irrelevance of intelligence in predicting responsiveness to reading instruction. *Exceptional Children, 73,* 8–30.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice, 13,* 204–219.

Fuchs, L. S., Deno, S. L., & Marston, D. (1983). Improving the reliability of curriculum-based measures of academic skills for psychoeducational decision making. *Diagnostique, 6,* 135–149.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5,* 257–288.

Good, R. H., Simmons, D. C., & Smith, A. (1999). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *Educational and Child Psychology, 15,* 56–70.

Graham, S., & Harris, K. R. (2005). Improving the writing performance of young struggling writers: Theoretical and programmatic research from the Center on Accelerating Student Learning. *Journal of Special Education, 39*(1), 19–33.

Grimes, J. (2002, August). *Responsiveness to interventions: The next step in special education identification, service, and exiting decision making.* Paper presented at OSEP's Learning Disabilities Summit conference, Washington, DC.

Hammill, D. D. (1991). *Detroit tests of learning aptitude* (3rd ed.). Austin, TX: PRO-ED.

Hammill, D. D., & Newcomer, P. L. (1988). *Test of language development: 2–Primary.* Austin, TX: PRO-ED.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego: Academic Press.

Huberty, C. J. (1994). *Applied discriminant analysis.* New York: Wiley.

Hurford, D. P. (1990). Training phonemic segmentation ability with a phonemic discrimination intervention in second- and third-grade children with reading disabilities. *Journal of Learning Disabilities, 23,* 564–569.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research & Practice, 18,* 237–245.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437–447.

Juel, C. (1994). At-risk university students tutoring at-risk elementary school children: What factors make it effective? In E. H. Hiebert & B. M. Taylor (Eds.), *Getting reading right from the start: Effective early interventions* (pp. 39–62). Boston: Allyn & Bacon.

Kame'enui, E. (2002). *Final report: An analysis of reading assessment instruments for K–3.* Retrieved December 5, 2003, from the Institute for the Development of Educational Achievement, University of Oregon, at http://idea.uoregon.edu/assessment/index.html

Kaminski, R. R., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25,* 215–227.

Kauffman, J. E., Cullinan, D., & Epstein, M. H. (1987). Characteristics of students placed in special programs for the seriously emotionally disturbed. *Behavioral Disorders, 12,* 175–184.

Kennedy, M. M., Birman, M., & Demaline, B. (1986). *The effectiveness of Chapter 1 services. Second interim report from the National Assessment of Chapter 1.* Washington, DC: Office of Educational Research and Improvement.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. *Developmental Psychology, 36,* 596–613.

Levy, B. A., & Lysynchuk, L. (1997). Beginning word recognition: Benefits of training by segmentation and whole word methods. *Scientific Studies of Reading, 1,* 359–387.

Lyon, G. R. (1985). Identification and remediation of learning disability subtypes: Preliminary findings. *Learning Disability Focus, 1,* 21–35.

Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., et al. (2001). *Rethinking learning disabilities.* Washington, DC: Hudson Institute.

Mathes, P. G., & Fuchs, L. (1994). The efficacy of peer tutoring in reading for students with mild disabilities: A best-evidence synthesis. *School Psychology Review, 23,* 59–80.

Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly, 33,* 62–94.

McMaster, K. L., Fuchs, D., Fuchs, L. S. & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71,* 445–463.

National Association of School Psychologists. (2002). *Draft: NASP recommendations on LD eligibility and identification for IDEA reauthorization.* Unpublished paper.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: National Institutes of Child Health and Human Development.

Nelson, R. J., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research & Practice, 18,* 255–267.

No Child Left Behind Act of 2001, P.L. 107-110, H.R. 1.

O'Connor, R. E. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilities Research & Practice, 15,* 43–54.

O'Connor, R. E., Jenkins, J., Leicester, N., & Slocum, T. (1993). Teaching phonological awareness to young children with learning disabilities. *Exceptional Children, 59,* 532–546.

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

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

O'Connor, R. E., Notari-Syverson, A., & Vadasy, P. F. (1998). *Ladders to Literacy: A kindergarten activity book.* Baltimore: Brookes.

O'Shaugnessy, T. E., & Swanson, H. L. (2000). A comparison of two reading interventions for children with reading disabilities. *Journal of Learning Disabilities, 33,* 257–277.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanations and prediction.* New York: Holt, Rinehart & Winston.

Pennington, B. (2004, April 8). *Comorbidities of dyslexia, ADHD, and speech–language disorders.* Paper presented at the John F. Kennedy Center for Research on Human Development, Vanderbilt University, Nashville, TN.

President's Commission on Excellence in Special Education. (2002). *A new era: Revitalizing special education for children and their families.* Washington, DC: Author.

Scarborough, H. S. (2002). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York: Guilford Press.

Schneider, W., Ennemoser, M., Roth, E., & Kuspert, P. (1999). Kindergarten prevention of dyslexia: Does training in phonological awareness work for everybody? *Journal of Learning Disabilities, 32,* 429–436.

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

Smith-Burke, M. T., & Jaggar, A. M. (1994). Implementing Reading Recovery in New York: Insights from the first two years. In E. H. Hiebert & B. M. Taylor (Eds.), *Getting reading right from the start: Effective early interventions* (pp. 63–84). Boston: Allyn & Bacon.

Snider, V. E. (1997). Transfer of decoding skills to a literature basal. *Learning Disabilities Research & Practice, 1,* 54–62.

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

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

Stein, M., Johnson, B., & Gutlohn, L. (1999). Analyzing beginning reading programs: The relationship between decoding instruction and text. *Remedial and Special Education, 20,* 275–287.

Stevens, J. (1986). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics.* New York: HarperCollins.

Tabors, P. O., Snow, C. E., & Dickinson, D. K. (2001). Homes and schools together

supporting language and literacy development. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language* (pp. 289–339). Baltimore: Brookes.

Torgesen, J. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice, 15,* 55–64.

Torgesen, J. K., & Davis, C. (1996). Individual difference variables that predict response to training in phonological awareness. *Journal of Experimental Child Psychology, 63,* 1–21.

Torgesen, J. K., Morgan, S., & Davis, C. (1992). The effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology, 84,* 364–370.

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

Uhry, J. K., & Shepherd, M. (1997). Teaching phonological recoding to young children with phonological processing deficits: The effect on sight word acquisition. *Learning Disability Quarterly, 20,* 104–125.

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

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*, 137–146.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children, 69,* 391–409.

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers. *Journal of Learning Disabilities, 33,* 223–238.

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

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91,* 415–438.

Yopp, H. K. (1995). A test for assessing phonemic awareness in young children. *Reading Teacher, 49*(1), 20–29.