Responsiveness-to-Intervention Symposium

December 4-5, 2003 • Kansas City, Missouri

The National Research Center on Learning Disabilities, a collaborative project of staff at Vanderbilt University and the University of Kansas, sponsored this two-day symposium focusing on responsiveness-to-intervention (RTI) issues.

The symposium was made possible by the support of the U.S. Department of Education Office of Special Education Programs. Renee Bradley, Project Officer. Opinions expressed herein are those of the authors and do not necessarily represent the position of the U.S. Department of Education.

Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods

Kristen L. McMaster

University of Minnesota

Douglas Fuchs, Lynn S. Fuchs, and Donald L. Compton

Peabody College of Vanderbilt University

When citing materials presented during the symposium, please use the following: "McMaster, K L., Fuchs, D., Fuchs, L. S., and Compton, D. L. (2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO."

Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods

Abstract

First graders (N = 323) participated in an evidence-based peer-mediated classwide reading program (Peer-Assisted Learning Strategies; PALS). A "dual-discrepancy" approach was used to identify 66 children unresponsive to PALS. Unresponsiveness was defined as performance levels and growth rates substantially below those of average readers based on Curriculum-Based Measures. An exploration of this approach revealed that the dual-discrepancy approach reliably distinguished among unresponsive at-risk, responsive at-risk, and average-performing readers. Unresponsive students were assigned randomly to one of three increasingly individualized treatments: PALS, Modified PALS, or one-to-one tutoring by an adult. The relative effectiveness of the three treatments was evaluated by comparing the three groups' performance on phonological awareness and reading-related measures. No statistically significant between-group differences were found. Effect sizes comparing the treatments and proportions of nonresponders following treatment suggest that one-to-one tutoring was the most promising for reducing unresponsiveness. Implications for further research and service delivery are discussed.

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods

Reading research over the past 20 years has greatly advanced our understanding of reading problems. We know that children who experience difficulty learning to read have phonological processing weaknesses (e.g., Liberman, Shankweiler, & Liberman; 1989) and poor word recognition skills (e.g., Ehri, 1998; Siegel, 1989; Share & Stanovich, 1995). Moreover, a number of researchers (e.g., Blachman, Ball, Black, & Tangel, 1994; Byrne & Fielding-Barnsley, 1991; D. Fuchs, Fuchs, Thompson, Al Otaiba et al., 2001; Torgesen, Wagner, & Rashotte, 1997) have demonstrated that programs emphasizing phonological awareness and decoding can significantly improve young students' reading achievement, at least in the short term.

Yet, there is a persisting problem: Not all children respond even to the most effective interventions. Researchers have reported that 20% to 30% of children at risk for reading difficulties (e.g., Blachman et al., 1994; Brown & Felton, 1990; Mathes, Howard, Allen, & Fuchs, 1998; Torgesen et al., 1999), and more than 50% of children with disabilities (e.g., D. Fuchs, Fuchs, Thompson, Al Otaiba et al., 2001; Torgesen et al., 2001), do not appear to benefit from generally effective early reading intervention. Such students have been dubbed "nonresponders." Recently, unresponsiveness-to-intervention has been suggested as an alternative criterion to the current IQ-achievement discrepancy approach for identifying students with learning disabilities (e.g., President's Commission on Excellence in Special Education, 2002). This has contributed to researchers' interest in finding the best ways to identify nonresponders and develop effective interventions to reduce unresponsiveness (e.g., Case, Speece, & Molloy, in press; O'Connor, 2000; Speece & Case, 2001; Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino et al., 1996).

Identifying Nonresponders.

It is generally agreed that nonresponders to reading intervention are students who do not make adequate reading progress despite their participation in evidence-based instruction. However, there is little agreement about what constitutes "adequate reading progress" (Torgesen, 2000). In an extensive review of research addressing unresponsiveness to reading instruction, Al Otaiba and Fuchs (2002) showed that researchers have used one of two basic indicators of reading progress: performance level or growth rate. In terms of performance level, researchers have identified unresponsiveness as performance below the 10^{th} percentile to below the 50^{th} percentile on a given measure (e.g., Foorman et al., 1998; Torgesen et al., 1999; Vellutino et al., 1996). With respect to growth, nonresponders have been identified on the basis of no growth (e.g., Berninger et al., 1999; Torgesen & Davis, 1996, Uhry & Shepherd, 1997) or limited growth (e.g., Vadasy, Jenkins, Antil,

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Wayne, & O'Connor, 1997; Vellutino et al., 1996).

There are serious conceptual problems related to performance-level-only and growth-rate-only approaches. For example, a child's performance level may be low but, without considering her growth rate, it is difficult to determine whether she is responsive to intervention. She may be making important growth. Likewise, using only growth to determine unresponsiveness ignores information about a child's performance relative to meaningful educational benchmarks. A child may be making steady progress, but may still be performing at such a low level that she cannot be expected to reach an adequate level of competency in a timely manner.

Developing valid methods of identifying nonresponders is a recognized goal of early reading intervention research. One alternative to the performance-level-only and growth-rate-only methods is a "dual-discrepancy" approach (L. Fuchs & Fuchs, 1998), whereby students must be discrepant from their peers in both performance level *and* growth rate to be considered unresponsive. Researchers are determining whether this discriminates among average readers and poor readers who do and do not respond to instruction (e.g., Speece & Case, 2001; Speece, Case, & Molloy, 2003). Other researchers are testing its utility by comparing it to alternative procedures like "median split," "normalized," and "benchmark" posttreatment scores (see L. Fuchs, 2003). The dual discrepancy approach seems to hold promise, but more work remains.

Treating Nonresponders

A second question important to the study of nonresponders is how to address such children's failure to learn to read. Many who have implemented interventions for struggling readers have reported the proportions of nonresponders (e.g., Berninger et al., 1999; Foorman et al., 1998; Torgesen et al., 1999). However, only a few have attempted multi-phased interventions in an attempt to decrease rates of unresponsiveness (e.g., Case et al., in press; O'Connor, 2000; Speece et al., 2002; Vaughn et al., 2003; Vellutino et al., 1996). In this subgroup of ambitious studies, the first phase has consisted of instruction for students identified as at risk for early reading failure, and subsequent phases have consisted of continued instruction directed at nonresponders in the previous phase. Some of these studies have defined unresponsiveness in the context of general education instruction; others in the context of more intensive, small-group instruction.

Special-education-like approaches.

Some researchers have conceptualized unresponsiveness as a failure to respond to instruction resembling traditional special education service delivery (e.g., Berninger et al., 1999; Foorman et al., 1998; Torgesen et al., 1999, Vellutino et al., 1996). That is, students were temporarily removed from the classroom to receive focused supplemental reading instruction from well-trained teachers. Vellutino et al., for example, implemented an intensive, one-to-one tutoring intervention for 15 weeks with at-risk first-graders. Students who did not make substantial progress during tutoring were designated as "difficult-to-remediate" and received further intervention. Vaughn et al. (2003) implemented daily small-group instruction with at-risk second-graders. Nonresponders (i.e., students who did not meet exit criteria after 10 weeks) received a second round of tutoring. Students who did not

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

meet exit criteria after another 10 weeks received a third phase. After 30 weeks of intervention, less than 25% of the original at-risk sample had not met the exit criteria.

Whereas researchers working within this "special-education-like" framework have demonstrated that many poor readers improve when intensive intervention is in place, they have not examined whether or how general educators might modify their instruction, curricula, or materials to accommodate the learning needs of at-risk students. This, of course, was not the purpose of their research, and we mean no criticism of it. Yet, the role of the general education teacher and the nature of mainstream instruction seem pivotal in identifying the most difficult-to-teach students and best ways to meet their needs.

We offer two reasons for this view. First, as Vellutino et al. (1996) have suggested, some students' reading difficulties are no doubt due to inadequate instruction rather than a true reading disability. Improving general education instruction may be sufficient to help many struggling readers and to identify those in need of more intensive instruction at a lower cost than providing intensive instruction to all at-risk students. Second, current education reforms emphasize evidence-based, general classroom interventions and modifications as a first step in addressing students' academic difficulties (e.g., President's Commission on Excellence in Special Education, 2002). Thus, for both pragmatic and policy-related reasons, the quality and effectiveness of classroom instruction seem important. More comprehensive examinations of unresponsiveness should begin by determining whether modifying instruction in the general education classroom is effective.

General education approaches.

O'Connor (2000) and Case et al. (in press) have done precisely this. O'Connor (2000) implemented four increasingly intensive levels of beginning reading interventions to kindergartners. Intervention at the first level was an evidence-based, whole-class, phonological awareness program conducted by general education teachers. Unresponsive students then received one-to-one tutoring from teaching assistants. Children who remained unresponsive received small-group instruction from their teachers at the beginning of first grade. Finally, still unresponsive first-graders received one-to-one tutoring from a researcher. The proportion of nonresponders decreased with each level of intervention. O'Connor's findings suggest that some poor readers benefit from evidence-based classroom instruction, whereas others require more intensive, individualized instruction.

Unlike O'Connor (2000), Case et al. (in press) did not implement an evidence-based classroom intervention before identifying nonresponders; rather, they defined unresponsiveness as a dual discrepancy between poor readers and their peers participating in regular classroom instruction. Case et al. then worked with the classroom teachers of the unresponsive students to design interventions supported by research, such as phonological awareness and phonics instruction, partner reading, or computer programs to address behavior and motivational needs. Most interventions were delivered only to the nonresponders. Speece et al. (2002) found that students who received the classroom interventions in the Case et al. study made greater academic gains than those who did not. Like O'Connor's findings, these results support the idea that evidence-based intervention within the general

education classroom may be beneficial for at least some nonresponders.

Comparing levels of service delivery.

The "special-education-like" approach described above is important to research examining unresponsiveness, in part because it identifies the most-difficult-to-teach students in relatively few steps: Students unresponsive to intensive, individualized instruction are likely to be among those most in need of ongoing, specialized services (e.g., Vellutino et al., 1996). But given the current emphasis on providing all students access to the general education curriculum, it is likely that most schools will encourage teachers to implement evidence-based classroom instruction and modifications before removing lagging students for individualized or small-group instruction (Vaughn, Gersten, & Chard, 2000). Moreover, schools typically do not have the resources for intensive, individualized interventions for all at-risk children. Thus, we believe that students' responsiveness to general education instruction (modified or unmodified) should be included in research focusing on how to identify and help nonresponders. As we have indicated, researchers working within a general education framework have begun to do this. What is still needed is a better understanding of approaches that are both effective for many nonresponders and practical to implement. To examine these issues, we directly compared the effectiveness of (a) evidence-based classwide intervention delivered by the general education teacher, (b) individualized modifications to generally effective classroom instruction, and (c) more intensive one-to-one intervention in strengthening struggling students' reading achievement.

Purpose of the Study.

The purpose of this study was twofold. First, we wished to further explore the validity of the dual-discrepancy approach (L. Fuchs & Fuchs, 1998); specifically, to refine a process that (a) distinguishes a "risk pool" of lowest-performing readers, (b) monitors their progress using valid measures, and (c) identifies nonresponders to intervention in a way that minimizes false positives and false negatives. In this study, we identified nonresponders to a generally effective classwide intervention, Peer-Assisted Learning Strategies (PALS; e.g., D. Fuchs, Fuchs, Thompson, Svenson et al., 2001). Both performance level and growth rate were measured using Curriculum Based Measurement (CBM; see Deno, 1985). Students dually discrepant from average readers with respect to performance level and growth rate were identified as nonresponders to PALS.

Our second purpose was to compare the achievement of nonresponders who continued receiving PALS to nonresponders who received one of two treatments designed to represent parts of the continuum of services available in schools for most students who are referred to or identified for special education: modified classroom intervention (Modified PALS), or individualized pull-out instruction (Tutoring). In this study, PALS served as a control, enabling us to compare the reading achievement of students who received increasingly individualized interventions to students who did not. PALS and Modified PALS may be considered best practices to meet the needs of most students in the general education classroom through evidence-based classwide interventions or through modifications (e.g., D. Fuchs & Fuchs, 1994). Tutoring reflects a more intensive, costly, and, according to some, "restrictive" level of intervention because the

student is temporarily removed from his or her peers. Whereas this more intensive level, with its "pull-out" dimension, may not currently be desired by all, support for it is emerging from reading research (e.g., Ehri & Robbins, 1992; Torgesen et al., 1999; Uhry & Shepherd, 1997; Vadasy et al., 1997; Vellutino et al., 1996).

Method

Participants

This study took place in eight Metropolitan Nashville schools participating in a large-scale investigation of the First-Grade PALS reading program from October, 2000, to April, 2001, (Fuchs, Fuchs, Yen et al., 2001). Four of the schools were high-poverty "Title I" schools, and four were middle-class "Non-Title I" schools. Thirty-three first-grade teachers volunteered to participate. The teachers were stratified by school type (Title I or Non-Title I) and assigned randomly within school to one of three conditions. In the large-scale study, 11 classrooms used a standard version of the PALS program (Standard PALS), 11 classrooms used a fluency-building version (PALS + Fluency), and 11 classrooms served as "no-treatment" controls. Pre- and posttreatment data were available for 496 students (168 in Standard PALS classes, 155 in PALS + Fluency classes, and 173 in control classes). The control classrooms did not participate in the nonresponder portion of the study. A three-step process guided selection of the nonresponders from the 22 Standard PALS and PALS + Fluency classrooms: (a) selecting students at risk for unresponsiveness to PALS, (b) monitoring the at-risk students' progress, and (c) identifying nonresponders among the at-risk group.

Selecting the risk pool. In October, 2000, written parental consent was obtained for students to participate in the large-scale study. These students were given a Rapid Letter Naming test, an effective predictor of future reading achievement (Torgesen et al., 1997). Within each classroom, students' were rank-ordered from highest- to lowest-performing based on their Rapid Letter Naming scores. Adjustments to these rankings were made based on teacher judgment. For example, if a student's Rapid Letter Naming score was low, but the teacher believed the student to be an average reader, the teacher's judgment overrode the Rapid Letter Naming ranking. Based on the adjusted rankings, 4 average-performing and 8 lowest-performing students were identified in each classroom. The 8 lowest-performing students from each of the 22 Standard PALS and PALS + Fluency classrooms were considered "at risk" to be unresponsive to the PALS program (n = 176). The 4 average-performing students per class served as a comparison group (n = 88).

Monitoring progress. From October to December, the at-risk and average-performing students' reading progress was monitored weekly. Monitoring measures included "chapter tests," which were criterion-referenced measures of students' progress in PALS, and two word-level reading CBM measures. These were Nonword Fluency probes from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2001), and Dolch word probes consisting of high-frequency pre-primer, primer, and first-grade-level words.

Identifying nonresponders. After seven weeks of PALS participation, complete monitoring data were available for 166 at-

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

risk and 87 average students (representing an attrition rate of 6% and 1% for at-risk and average students, respectively). At this point, nonresponders were identified in accordance with a five-step process. First, the CBM levels and growth rates (i.e., slopes) of each of the at-risk and average students were calculated. Level indicated the number of correct words per min the student read at the end of the monitoring period. It was calculated as the mean of each student's last two scores. Slope, by contrast, indicated how many more correct words per min students read each time they were monitored. For example, a slope of 1 indicated a growth rate of one word read correctly per min for each monitoring session. It was calculated using a least squares regression between monitoring scores and calendar days.

Next, means and standard deviations (*SD*) of the average-performing students' CBM levels and slopes were calculated. Third, all students from each classroom were identified who had scored 90% or less on the last chapter test, or who were the lowest-scoring in their class on this measure (n = 97). Fourth, z-scores were calculated separately on the CBM levels and slopes of these 97 students, using the means and *SD*s of the average-performing students. Finally, we examined the z-scores of each and every low-performing student. Students were identified as nonresponders if they scored more than 0.5 *SD* below the average readers in terms of both level and slope on the CBM measures (n= 66).

Assigning nonresponders to groups. Within the Standard PALS and PALS + Fluency classes, the 66 nonresponders were stratified in terms of "low" (n = 28) vs. "very low" (n = 38) status based on CBM levels and slopes. Then, they were assigned randomly to one of three treatments: PALS, Modified PALS, or Tutoring. Each of the three treatments was nested within the PALS and PALS + Fluency treatments, such that Modified PALS and Tutoring students received either a Standard or Fluency version of these programs, depending on their initial PALS condition.

Twenty-two students were assigned randomly to each of the three treatments. Due to attrition, there were 21, 15, and 20 students in the PALS, Modified PALS, and Tutoring treatments, respectively, for a total of 56 nonresponders, at the end of the study. Analyses of variance (ANOVAs) showed no statistically significant differences across the three groups in terms of Dolch level (F [2, 53] = 0.37, p = 0.70), Dolch slope (F [2, 53] = 1.09, p = 0.34), Nonword Fluency level (F [2, 53] = 0.24, p = .79), and Nonword Fluency slope (F [2, 53] = 0.94, p = 0.40) before Modified PALS and Tutoring began. Students in the three conditions were also compared on several demographic variables, including sex, ethnicity, English Language Learner (ELL) status, socioeconomic status, and special education status. Chi-square tests indicated no statistically significant differences across conditions on these variables.

Treatments

PALS. First-Grade PALS was developed by researchers at Vanderbilt University (e.g., D. Fuchs, Fuchs, Thompson, Svenson et al., 2001; Mathes et al., 1998). PALS is a structured, peer-mediated program that emphasizes phonological awareness, beginning decoding, word recognition, and fluency, all skills that researchers have demonstrated to be important for successful beginning reading programs (e.g., Blachman et al., 1999; Hatcher, Hulme, & Ellis, 1994; Juel, 1996;

Torgesen et al., 1997; Vellutino et al., 1996). Results of several large-scale experimental studies indicate that PALS helps develop beginning reading skills for a majority of low-, average-, and high-achieving students, and for many children with disabilities (e.g., D. Fuchs et al., 1997; D. Fuchs, Fuchs, Thompson, Al Otaiba et al., 2001; Mathes et al., 1998). PALS has proven to be effective in schools with many minority children and children living in poverty as well as in schools with predominantly white, middle-class student populations. However, an estimated 10% to 20% of low-achieving non-disabled students (Mathes et al.), and more than 50% of students with disabilities (D. Fuchs, Fuchs, Thompson, Al Otaiba et al.) have not responded to PALS.

In the large-scale investigation that provided the context for this study, Standard PALS was compared to PALS + Fluency, which was designed to promote reading fluency and comprehension. Standard PALS and PALS + Fluency were implemented three times per week for approximately 35 min per session. Teachers paired higher performing readers with lower performing readers. Each lesson began with a brief teacher-led introduction of new sounds and words. Then, the students conducted the PALS activities in pairs.

The higher performing student was always the tutor or "Coach" first, and the lower-performing student was the "Reader" first. For each activity, the Coach provided prompts, praise, and corrective feedback to the Reader. After completing each activity, the students switched roles. PALS activities included letter-sound recognition, decoding, sight word recognition, and reading short stories. Students also conducted Partner Reading in books that corresponded to the reading level of the lower-performing student in each pair. The Coach read a page, then the Reader read the same page. When the partners finished a book, they switched roles and read it at least three more times. Students in PALS + Fluency conducted the same activities, with two modifications. The sight words were presented in phrases rather than in isolation, and the short stories were read in a repeated reading, "Speed Game" format in which a student read the story in a fixed time, then had two chances to read more words than the first time.

Modified PALS. Modified PALS lessons were conducted in the classroom during the scheduled PALS time. The teachers selected Coaches who were capable of reading the PALS lessons independently and demonstrated the ability to work well with lower-performing students. Modified PALS activities were similar to PALS; however, they incorporated three important modifications. First, fewer sounds and words were introduced at one time, and the students worked on lessons that matched their skill level. Second, the Coach modeled the sounds and words for the Reader. Opportunities for reading without a model were also built into the activities. Third, more emphasis was placed on phonological awareness and decoding skill.

Tutoring. Tutoring took place three times per week, for 35 min per session, and substituted for PALS. Students received tutoring from a trained adult and the tutoring roles were not reciprocal. Tutoring was more individualized than PALS and Modified PALS in several ways. First, tutors were trained to teach students to mastery. Sounds and words used in tutoring were grouped into sets. Students did not progress to a new set until they had mastered the sounds and

words in the first set. Second, the tutors spent more time on activities that were especially difficult for the students. Third, an additional motivational component was built into the tutoring lessons. During each lesson, the student determined how many sounds and words were needed to master the set, and marked this "goal" on a bar graph. For example, if there were 15 sounds in a set, the student would draw a line at the number 15 on the y-axis of the graph. At the end of the activity, the student counted the number of sounds mastered, and graphed this number in relation to the goal.

Pre- and Posttest Measures

A set of measures was individually administered to all study participants prior to and immediately following the treatment period. The set included tests of rapid naming, phonological awareness, reading words, and spelling.

Rapid naming. There were two rapid naming measures, the first of which was Rapid Letter Naming, given to assess letter recognition. It consisted of upper and lower case letters presented randomly in black type on a single sheet of paper. Students were instructed to name the letters as quickly as they could. The score was recorded as the number of letters named correctly in 1min. The Rapid Letter Sound test is based on a measure used by Levy and Lysunchuk (1997) and was developed for use in a previous PALS study (D. Fuchs, Fuchs, Thompson, Al Otaiba et al., 2001). All 26 letters of the alphabet were presented randomly in black type on a sheet of paper. Students were instructed to say the sounds as quickly as they could. The score was recorded as the number of sounds produced correctly in 1min.

Phonological awareness. The ability to segment words into phonemes correlates highly with future reading ability (Torgesen et al., 1997). A segmentation test based on the Yopp-Singer test (Yopp, 1988) and developed for use in previous PALS studies (e.g., D. Fuchs, Fuchs, Thompson, Al Otaiba et al., 2001) was administered. Students were instructed to say the sounds in each word provided. One point was recorded for each correct phoneme. The score was recorded as the number of phonemes expressed correctly in 1min.

A blending task, again developed previously (D. Fuchs, Fuchs, Thompson, Al Otaiba et al., 2001), measured students' ability to blend phonemes into words. One point was recorded for each correctly blended word. For example, if the examiner said "s-oa-p," the student earned 1 point for saying "soap." Two scores were recorded: a timed score (the number of words blended correctly in 1min) and an untimed score (the total number of words blended correctly).

Reading words. The Word Identification and Word Attack subtests of the Woodcock Reading Mastery Test-Revised (WRMT-R; Woodcock, 1987) were given to measure word recognition and decoding skills. Scores on the Word Identification and Word Attack subtests correlate highly with other tests of reading, and internal consistency exceeds .80.

Spelling. The Wechsler Individual Achievement Test (WIAT; Psychological Corporation, 1992) spelling subtest was administered. Students were instructed to write letters and words on a sheet of paper. Two scores were recorded: a standard score based on words written correctly, and a developmental score that allowed partial credit for

including key spelling patterns. The WIAT correlates well (.70s to .80s) with other individually administered achievement tests, and has a test-retest reliability coefficient of .94.

Posttest Only Measures

Near-transfer reading passages. Two "near-transfer" passages were administered to all PALS participants at posttest. The stories are similar to PALS stories in terms of words used, style, and format. Students were instructed to read the stories quickly and correctly. The score was recorded as the number of words read correctly in 1min.

Far-transfer reading passages and comprehension. Two "far-transfer" reading passages, taken from the Comprehensive Reading Assessment Battery (CRAB; L. Fuchs, Fuchs, & Hamlett, 1989), were administered to all PALS students. The passages are traditional folktales. Students were instructed to read the stories quickly and correctly. The score is recorded as the number of words read correctly in 1min. Test-retest reliability for the fluency measure ranges from .93 to .96, and concurrent validity with the comprehension subtest of the Stanford Achievement Test (SAT) is .91 (L. Fuchs, Fuchs, & Maxwell, 1988). The passages were followed by 10 open-ended comprehension questions. For the number of questions answered correctly, test-retest reliability is .92; the correlation with the SAT is .82 (L. S. Fuchs et al.).

Monitoring Measures

Chapter tests. Research staff developed seven chapter tests. These tests covered material presented in PALS lessons. The tests were cumulative and untimed, and included sounds and words that had been introduced before the test was administered. The score was recorded as the percentage of sounds and words read correctly.

Dolch probes. Research staff developed the Dolch probes for this study. These probes are equivalent forms of 100 sight words selected randomly from a pool of 126 high frequency words. Students were instructed to read the words quickly and correctly. The score was recorded as the number of words read correctly in 1min.

Nonword Fluency Probes. The Nonword Fluency probes (Good & Kaminski, 2001) consist of consonant-vowel-consonant and vowel-consonant nonwords. Students were instructed to say the individual sounds of the letters or read the whole word. The score is recorded as the total number of phonemes pronounced correctly in 1 min. Alternate-form reliability is .83, concurrent validity with the Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R) Readiness Cluster is .36, and predictive validity with the WJ-R Total Reading Cluster is .66 (Good et al., 2002).

Procedures

Test-administration training. Research staff were taught to administer the pretreatment, posttreatment, and monitoring measures in several training sessions. Interrater agreement sessions followed, during which the staff observed simulated testing sessions and scored protocols. Interrater agreement was calculated using a point-by-point method comparing each staff member's scored protocols with protocols scored by an experienced tester. Staff members who failed to reach 90% agreement on one or more of the measures received additional training and practice until they achieved criterion.

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Pre- and posttreatment testing. In October, staff members obtained signed parental consent forms from students participating in PALS. These students were tested individually in a quiet location (such as a library or conference room). Tests were administered in two sessions. During the first session, the examiner spent several minutes establishing rapport with the student. The Rapid Letter Naming test was always administered first; the Rapid Letter Sound, segmentation, and WRMT-R subtests were then given in random order. During the second session, the blending test was administered, followed by the spelling test. For a given student, the staff member who administered measures in session one also worked with the child in session two. In April, research staff again tested students in two sessions. To avoid examiner bias, staff did not test students whom they had tutored. Again, the examiner spent several minutes establishing rapport with the student. The near-transfer passages were added to the end of the first session, and the CRAB was added to the end of the second session. Again, testers worked with the same child across the two sessions.

Monitoring progress. For the first seven weeks of PALS, research staff administered the monitoring measures individually to the at-risk and average-performing students. After the nonresponders were identified and assigned randomly to treatment conditions, the CBM measures were administered biweekly for the remaining 13 weeks of treatment.

Training, technical assistance, and implementation. In October, all teachers from the Standard PALS and PALS + Fluency classrooms attended a full-day workshop in which the PALS procedures were described, demonstrated, and practiced. Teachers then returned to their classrooms and trained their students to conduct PALS. A staff member visited each classroom twice weekly. All participants implemented PALS for 7 weeks (from October to December).

In January, the research staff attended a full-day workshop to learn the Modified PALS and Tutoring procedures. Each staff member was assigned tutoring and/or classroom support roles. Those designated to support PALS classrooms were responsible for training nonresponders and their Coaches to conduct Modified PALS, and for providing weekly technical assistance. Tutors were responsible for working one-to-one with their assigned students three times per week, 35 minutes per session. The PALS, Modified PALS, and Tutoring treatments were implemented for an additional 13 weeks (from January to April).

Treatment fidelity. For PALS, Modified PALS, and Tutoring, a checklist of teacher and student behaviors was developed to assess fidelity of implementation. In December, and again in March, fidelity checks of the 22 PALS classrooms were conducted. On average, PALS was implemented with 92% fidelity. In April, the first author conducted a fidelity check of Modified PALS. Each of 15 pairs participating in Modified PALS was observed. On average, it was implemented with 86% fidelity, with a range of 49% to 100%.

Because of scheduling problems, it was not practical to do on-site observations of the Tutoring activities. Instead, each tutor participated in a simulated tutoring session. The tutor conducted the activities with another staff member serving as the tutee. Across the 8 tutors, the simulations were implemented with 97% fidelity, as scored by

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

the first author. In addition, one session conducted with each nonresponder was audiotaped and reviewed by a staff member. These sessions were conducted, on average, with 96% fidelity. (For more detailed information on the fidelity of implementation of PALS, Modified PALS, and Tutoring, contact the first author.)

Results

Identifying Nonresponders

One purpose of this study was to explore a dual-discrepancy approach to identifying nonresponders to generally effective intervention. A series of analyses was conducted to determine (a) the success of the screening process for identifying a very low-performing risk pool, (b) the reliability and validity of the progress monitoring measures, and (c) the sensitivity of the dual-discrepancy approach in identifying nonresponders to PALS while minimizing the numbers of false positives and false negatives.

As indicated, we first identified average readers and a risk pool of low-performing readers, using Rapid Letter Naming scores and teacher judgment. To determine whether this screening process accurately identified a group of very poor readers, one-way ANOVAs comparing at-risk and average readers were run on all of the pretreatment measures. The average readers statistically significantly, and dramatically, outperformed the risk pool on all measures, indicating that this screening process successfully earmarked a group of students who were performing reliably and substantially below their average peers (see Table 1).

The second step was to monitor the progress of the risk pool during the first seven weeks of PALS. To determine criterion validity and test-retest reliability, Pearson *r*s were calculated among the monitoring levels and pre- and posttreatment scores. Dolch levels correlated strongly with the Word Identification and Word Attack subtests, the near- and far-transfer fluency measures, and the comprehension measure, with coefficients of .82, .52, .92, .93, and .73, respectively. Nonword Fluency levels correlated well with the same measures, with coefficients of .65, .51, .78, .80, and .54, respectively. Each of these correlations was significant at the $p < .01$ level. In addition, test-retest correlations for the Dolch and Nonword Fluency measures were calculated. These correlations were .88 for Dolch and .87 for Nonword Fluency. Results suggest the monitoring measures were valid indicators of reading-related skills and were stable over time.

Finally, to determine whether the dual discrepancy approach successfully differentiated unresponsive at-risk students from responsive at-risk students and average readers, the CBM levels and slopes of these three groups were compared using ANOVAs. Not surprisingly, there were statistically significant differences among the three groups (Dolch level: F [2, 250] = 47.73; $p < .001$; Dolch slope: F [2, 250] = 19.13; $p < .001$; Nonword Fluency level: F [2, 250] = 64.97; $p < .001$; Nonword Fluency slope: F [2, 250] = 3.99; $p < .05$). As shown in Table 2, follow-up comparisons indicated that both average-performing and responsive at-risk readers had statistically significantly higher CBM levels and slopes than nonresponders. Effect size contrasts revealed that these differences were not only reliable but, for the most part, large in magnitude, indicating the dual-discrepancy criterion successfully identified a very low-performing group of readers.

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Table 1

Comparisons of At-Risk and Average-Performing Students on Pretreatment Measures

	At-Risk[a] (*n*= 174)		Average[a] (*n*= 87)			
Measure	*M*	*(SD)*	*M*	*(SD)*	*F*	(*df*)
Rapid Letter Naming	34.02	(13.40)	48.75	(10.37)	80.74**	(259)
Rapid Letter Sound	20.60	(13.96)	31.75	(11.51)	41.36**	(259)
Segmentation	20.40	(14.68)	26.14	(12.61)	9.70**	(259)
Word ID	9.20	(9.54)	17.63	(9.75)	44.63**	(259)
Word Attack	3.45	(5.27)	6.92	(4.74)	26.87**	(259)
Timed Blending	12.44	(7.47)	16.52	(6.66)	18.58**	(259)
Untimed Blending	14.78	(7.74)	18.91	(6.06)	18.94**	(259)
Spelling	8.83	(3.48)	11.63	(3.02)	40.49**	(258)

**$p < .01$

[a]Pretreatment data were not available for 2 at-risk students and 1 average-performing student.

Table 2

Comparisons of At-Risk Nonresponders, At-Risk Responders, and Average-Performers in December

Comparisons

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Measures	Responders vs. Nonresponders			Average Performers vs. Nonresponders			Average Performers vs. Responders		
	Mean Difference	*SE*	*ES*[a]	Mean Difference	*SE*	*ES*[b]	Mean Difference	*SE*	*ES*[b]
Dolch									
Level	8.41***	1.86	0.85	18.91***	1.96	1.09	10.51***	1.74	0.61
Slope	0.57***	0.15	0.57	1.00***	0.16	0.82	0.43**	0.14	0.35
NWF									
Level	17.34***	2.17	1.35	25.72***	2.28	1.51	8.38***	2.04	0.49
Slope	0.97**	0.34	0.41	0.60	0.36	0.26	-0.36	0.32	-0.15

Note. *SE* = Standard error; *ES* = Effect size; NWF = Nonword Fluency.

[a]Effect sizes were calculated by dividing mean difference by the *SD* of the responders.

[b]Effect sizes were calculated by dividing mean difference by the *SD* of the average performers.

**$p < .01$, ***$p < .001$

Comparing Treatment Effectiveness

Significance tests. A second purpose of this study was to compare the effects of three increasingly individualized treatments on nonresponders' reading performance. A three-factor nested design was used to analyze pre- and posttreatment group differences after 13 weeks of additional treatment. Treatment (PALS vs. Modified PALS vs. Tutoring) was nested within PALS program (Standard PALS vs. PALS + Fluency). Nonresponder status (low vs. very low) was nested within treatment. The pretreatment data were analyzed with 2 x 3 x 2 (PALS program x treatment x nonresponder status) ANOVAs. There were no statistically significant pretreatment main effects or interactions.

Table 3

Reading Performance by Nonresponder Treatment

Measures	PALS (*n*= 21)		Modified PALS (*n*= 15)		Tutoring (*n*= 20)	
	M	*(SD)*	*M*	*(SD)*	*M*	*(SD)*

Pre- and Post-test

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

	Rapid Letter Naming						
	Pre	30.14	(7.64)	27.60	(12.84)	27.30	(12.83)
	Post	47.43	(17.11)	44.47	(19.23)	45.80	(19.19)
	Growth	17.28	(13.96)	16.87	(13.50)	18.50	(15.83)
	Rapid Letter Sound						
	Pre	13.29	(8.44)	18.20	(15.62)	12.90	(9.42)
	Post	41.71	(11.15)	48.27	(22.75)	45.50	(15.54)
	Growth	28.42	(9.11)	30.07	(16.44)	32.60	(13.59)
	Segmentation						
	Pre	13.62	(13.00)	18.53	(13.19)	15.05	(13.68)
	Post	35.52	(13.39)	33.87	(14.91)	35.25	(14.04)
	Growth	21.90	(12.40)	15.33	(12.61)	20.20	(13.03)
	Word Identification						
	Pre	2.86	(2.85)	5.40	(5.53)	4.65	(6.05)
	Post	20.62	(9.56)	20.47	(10.13)	25.60	(9.51)
	Growth	17.76	(9.74)	15.07	(8.33)	20.95	(9.78)
	Word Attack						
	Pre	1.10	(1.87)	2.27	(3.08)	1.40	(2.46)
	Post	6.71	(5.26)	8.67	(5.55)	8.95	(5.36)
	Growth	5.62	(5.10)	6.40	(4.45)	7.55	(4.99)
	Timed Blending						
	Pre	9.67	(7.07)	10.00	(6.01)	10.75	(7.01)
	Post	19.38	(8.54)	19.00	(5.94)	22.70	(7.14)
	Growth	9.71	(6.70)	9.00	(6.23)	11.95	(7.56)
	Untimed Blending						
	Pre	12.67	(7.64)	13.67	(7.79)	13.00	(7.68)
	Post	20.81	(7.69)	20.13	(5.71)	23.20	(6.87)
	Growth	8.14	(6.30)	6.47	(5.48)	10.20	(7.73)
	Standard Spelling						
	Pre	6.86	(2.01)	7.47	(2.59)	6.95	(3.33)
	Post	12.67	(3.37)	11.27	(3.61)	12.45	(2.86)
	Growth	5.81	(3.28)	3.80	(2.75)	5.50	(3.75)
	Developmental Spelling						
	Pre	28.43	(16.56)	32.87	(18.92)	30.65	(26.18)
	Post	74.29	(27.25)	67.60	(23.18)	75.05	(22.18)
	Growth	45.86	(24.87)	34.73	(19.88)	44.40	(30.02)
Post-test Only							
	Near-Transfer Fluency	18.26	(12.55)	19.07	(12.71)	22.20	(10.68)
	Far-Transfer Fluency	19.29	(12.22)	19.90	(13.69)	22.95	(10.08)
	Comprehension	0.38	(0.55)	0.70	(0.92)	0.68	(0.82)

The posttreatment data were analyzed with 2 x 3 x 2 (PALS program x treatment x nonresponder status) analyses of covariance (ANCOVAs). ANCOVAs were used because, although there were no statistically reliable pretreatment group differ-

ences, we suspected that possible between-group differences may not have been detected due to low statistical power. Thus, we used ANCOVA with December Dolch level as the covariate because it was a more proximate pretreatment measure than those administered in October. Because there were no statistically significant main effects or interactions of interest, we simplified the design of our analysis to one-between group ANCOVAs. Table 3 shows means and standard deviations of the pre- and posttreatment and growth scores of the three groups on all measures. No statistically significant between-group differences were found on any of these measures. Respective *F*s (*df* 2, 52) are as follows: Rapid Letter Naming = .26, p = .77; Rapid Letter Sound = 0.59, p = .56; segmentation = 0.07, p = .93; Word Identification = 1.47, p = .24; Word Attack = 0.79, p = .46; timed blending = 1.48, p = .24; untimed blending = 1.03, p = .36; standard spelling = 1.13, p = .33; developmental spelling = 0.53, p = .59; near transfer = 0.35, p = .70; far transfer = 0.30, p = .74; or comprehension = 0.78, p = .46.

In addition to comparing pre- and posttreatment performance, repeated measures ANOVAs were conducted on the CBM levels and slopes to compare nonresponders' performance during the first seven weeks of PALS to their performance during the additional 13 weeks of treatment. Students' CBM levels and slopes in April were significantly greater following the PALS, Modified PALS, or Tutoring treatments than their levels and slopes in December (Dolch level: F [2, 53] = 133.45, $p < .001$; slope: F [2, 53] = 60.38, $p < .001$; Nonword Fluency level: F [2, 53] = 313.25, $p < .001$; slope: F [2, 53] = 65.59, $p < .001$). This indicates that students made greater gains during spring implementation of the three treatments than during the first seven weeks of PALS implementation. However, there were no statistically significant interactions between treatment and time; that is, the higher levels and slopes in April cannot easily be attributed to the treatments.

Power analysis and effect sizes. A power analysis was conducted to determine whether the treatment groups were sufficiently large to yield statistically significant between-group differences that may have existed. Assuming that the difference between groups would be small to moderate (i.e., an effect size of .30), the sample size needed for a power level of .70 is about 103 students per group. Posttreatment data were available for 21, 15, and 20 students in PALS, Modified PALS, and Tutoring, respectively. Thus, the statistical tests used were low powered for detecting relatively moderate differences. Whereas one might reasonably suggest we could have been more planful, it should be noted that a study of this kind cannot guarantee a specific number of participants. We attempted to maximize the number of nonresponders in our selection process by starting with a relatively large sample (N = 323); nevertheless, it was impossible to ensure an appropriately large sample of nonresponders.

Because statistical analyses were relatively low powered, we calculated effect sizes as another means to explore the importance of the study. We used effect sizes for ANCOVA, again relying on December Dolch level as the covariate. Moderate effects, reported in Table 4, were found (a) favoring Tutoring vs. PALS on Word Identification, Word Attack, blending, and comprehension, (b) favoring Tutoring vs. Modified PALS on Word Identification, blending, and spelling; and (c) favoring Modified

PALS vs. PALS on Rapid Letter Sound, | Word Attack, and comprehension.

Table 4

Posttreatment Effect Sizes by Nonresponder Treatment

	Comparisons		
Measures	Tutoring vs. PALS	Tutoring vs. Modified PALS	Modified PALS vs. PALS
Rapid Letter Naming	-0.09	0.07	-0.16
Rapid Letter Sound	0.23	-0.17	0.40
Segmentation	-0.02	0.09	-0.11
Word ID	0.43	0.44	-0.01
Word Attack	0.38	0.05	0.33
Blending	0.44	0.49	0.05
Spelling	-0.06	0.31	-0.37
Near Transfer Fluency	0.22	0.18	0.05
Far Transfer Fluency	0.21	0.17	0.03
Comprehension	0.32	-0.02	0.34

Note. Effect sizes were calculate using effect size for ANCOVA (Wilson, 1996).

Proportions of Nonresponders Following Treatment

Dual-discrepancy approach. We also calculated proportions of students who continued to be unresponsive following implementation of the three treatments. The dual-discrepancy approach used to identify initial nonresponders in December was used again to identify nonresponders following the 13 weeks of additional treatment. As in December, students were identified as nonresponders if they scored more than 0.5 *SD* below average-performers' levels and slopes

on the CBM measures. Accordingly, 81% of PALS students, 80% of Modified PALS students, and only 50% of Tutoring students remained unresponsive to treatment at the end of the study. Overall, 70% of the initial nonresponders were still unresponsive. This translates to 22% of the 174 at-risk students who were monitored during the first half of the study, or 12% of the total number of students participating in Standard PALS or PALS + Fluency (n = 323).

Performance-level-only and growth-rate-only approaches. To determine whether the procedures we used to identify nonresponders yielded similar proportions as alternative procedures used by other researchers, two additional approaches to identifying nonresponders (i.e., performance-level-only and growth-rate-only) were explored. Performance levels and growth rates on the Word Identification and Word Attack subtests and the near- and far-transfer fluency measures were examined (see Berninger et al., 1999; Brown & Felton, 1990; D. Fuchs, Fuchs, Thompson, Al Otaiba et al., 2001; Torgesen et al., 1999; Vellutino, 1996). Table 5 presents proportions of nonresponders identified using the dual discrepancy approach as well as the approaches used in previous studies.

In terms of performance-level-only approaches, we first determined how many of the initial 56 nonresponders for whom we have complete data performed below the 30th percentile (as recommended by Torgesen, 2000) at the end of the study. Next, we calculated how many of the initial 56 nonresponders read less than 40 words correctly in 1 min (as suggested by Good, Kaminski, & Shinn, 1999) at study's end. In terms of growth rate, we figured how many students made no growth on the Word Identification and Word Attack subtests or limited growth (less than 10 words gained on the Word Identification subtest; less than 5 words gained on the Word Attack subtest).

As shown in Table 5, each approach to identifying nonresponders yielded a different proportion of students, sometimes strikingly so. Using percentile as a performance-level-only criterion (as per Torgesen, 2000) resulted in fewer nonresponders than the dual-discrepancy approach. However, students who scored above the 30th percentile on the WRMT-R subtests, but still met the dual-discrepancy criterion, had slopes that were .80 *SD* below the average-performers on the monitoring measures. This finding suggests that performance above the 30th percentile may mask very poor progress toward higher levels of reading. On the other hand, using a specific criterion level as a performance-level-only criterion (as per Good et al., 1999) resulted in many more nonresponders. Some of the students who had not yet reached 40 correct words per min were making mean gains of .08 *SD above average*, indicating that they were making similar word-reading gains to their average-performing peers. Thus, these students were probably not true "nonresponders." Similarly, the "no growth" (e.g., Berninger et al., 1999) and "limited growth" (e.g., Vellutino, 1996) criteria resulted in fewer nonresponders than the dual-discrepancy criteria, most likely overlooking many students who made some, but not sufficient growth.

Discussion

Identifying Nonresponders

The first purpose of this study was to explore the usefulness of a dual-discrepancy

approach to identifying nonresponders (see L. Fuchs & Fuchs, 1998). Findings show that our CBM measures were reliable and valid indicators of students' reading skill, and that the dual-discrepancy approach identified students performing statistically significantly below their peers. This result is consistent with recent evidence that dual-discrepancy can successfully distinguish nonresponders from responsive at-risk and average-achieving children on reading-related measures (e.g., Speece & Case, 2001; Speece et al., 2003).

Table 5
Proportions of Nonresponders at the End of the Study Identified by Dual-Discrepancy, Performance-Level-Only, and Growth-Rate Only Criteria

			Performance Level Only				Growth Rate Only			
	Dual Discrepancy[a]		Percentile Rank[b]		Criterion Level[c]		No Growth[d]		Limited Growth[e]	
Treatment	*n*	(%)	*n*	(%)	*n*	(%)	*n*	(%)	*n*	(%)
PALS (*n*= 21)	17	(81)	8	(38)	21	(100)	2	(10)	10	(48)
Modified PALS (*n*= 15)	12	(80)	8	(53)	15	(100)	1	(7)	7	(47)
Tutoring (*n*= 20)	10	(50)	9	(45)	20	(100)	1	(5)	5	(25)
Total nonresponders (*N*= 56)	39	(70)	25	(45)	56	(100)	4	(7)	22	(39)
Total risk pool *(N* = 174)	39	(22)	25	(14)	56	(32)	4	(2)	22	(13)
Total PALS participants (*N* = 323)	39	(12)	25	(8)	56	(17)	4	(1)	22	(7)

[a]Dual discrepancy was determined by slopes and levels > .5 *SD* below average on the Dolch and/or Nonword Fluency measures. [b]Percentile rank was determined by scores below the 30th percentile on the WRMT-R Word Identification and/or Word Attack subtests. [c]Criterion level was determined by < 40 correct words per min on the Near-Transfer and/or Far-Transfer Fluency measures. [d]No growth was determined by a gain of 0 words (or less) on Word Identification and/or Word Attack. [e]Limited growth was determined by a gain of 10 words or less on Word Identification and/or a gain of 5 words or less on Word Attack.

Findings also suggest that dual-discrepancy holds promise as a better method of identification than performance-level-only and growth-rate-only approaches. In this study, the performance-level-only approach would identify a student as a non-responder who scored below the 30th percentile on the WRMT-R Word Identification or Word Attack subtests. The dual-discrepancy approach would not identify such a student as a nonresponder if she were making growth similar to average-performing readers. Likewise, using growth-rate only criteria, a student making limited growth may be identified as a nonresponder. However, if this student were reading at a level commensurate with his average-performing peers, he probably would not need additional instruction. Using the dual-discrepancy approach, both his growth rate and performance level would need to be below those of average performers for him to be considered unresponsive. Our dual-discrepancy approach provided performance-level criteria (based on average students' performance) needed to make such a decision.

Nevertheless, we offer two important caveats about our favored approach. First, just as performance-level and growth-rate criteria used by previous researchers are arbitrary, our discrepancy criterion of .5 *SD* below average is arbitrary, too. Currently, there is no consensus regarding how far below average a student must perform to warrant a change in intervention. Second, non-responders were identified using word-level CBM measures rather than other indicators of reading skill. These measures were selected because they were likely to be sensitive to weekly growth in low-performing first graders. Yet, others have used measures of phonological awareness, spelling, fluency, or comprehension to identify younger and older nonresponders. Continued research is needed to determine which measures are most appropriate to identify nonresponders at different points of reading development.

Responding to Nonresponders

Our second purpose was to compare the effects of PALS, Modified PALS, and Tutoring to determine which was most effective in improving the reading performance of nonresponders to PALS and in reducing rates of unresponsiveness. There were no statistically significant differences across the treatment approaches on the monitoring measures and on any of the other reading-related measures.

There are several possible explanations for this. First, the treatments' relatively low intensity and short duration may simply

have been insufficient to yield dramatic gains in students with severe reading difficulties. A number of researchers (e.g., Berninger et al., 2002; Blachman et al., 1999; Foorman et al., 1998; Torgesen et al., 2001; Vellutino et al., 1996) have conducted treatments for similar students that have totaled 65 hours to more than 300 hours, some spanning 2 years or more. Moreover, these interventions have often been implemented by highly trained teachers or reading specialists. Even such intensive interventions have not succeeded in improving the reading achievement of a subset of at-risk readers, with proportions of nonresponders similar to those reported in this study. Another explanation is that PALS, Modified PALS, and Tutoring all incorporated comparable activities. Whereas the treatment activities and formats were deliberately designed to be similar so we might explore the importance of varying levels of individualization, it is possible that they were not sufficiently different from each other to produce different results.

A third explanation, already discussed, may be the low statistical power of the study. In recognition of this fact, we calculated effect sizes and proportions of nonresponders at the end of the study as additional ways to compare the three treatments. Effect sizes favored Tutoring over PALS and Modified PALS on a number of reading-related measures, and favored Modified PALS over PALS on Rapid Letter Sound naming, Word Attack, and comprehension. A similar pattern is seen in the proportions of students who continued to be unresponsive at the end of treatment. Tutoring reduced unresponsiveness by 50%, whereas the Modified PALS and PALS treatments lessened unresponsiveness by 25% and 19%, respectively.

The Tutoring program deserves some additional comment. The addition of one-to-one instruction provided by a trained research assistant may explain why half of the tutored students were responsive. The research assistant was trained to ensure that immediate, correct feedback was provided, and that students mastered the content of the tutoring lessons before moving to new content. Although peers have been demonstrated to be effective tutors for many students, perhaps the most difficult-to-teach students require the more individualized support that trained adult tutors are able to provide. Also, tutored students had more opportunities to respond than PALS and Modified PALS students, since tutored students were always the "Readers" and there were fewer transitions.

A final feature that may have enhanced the Tutoring treatment is difficult to quantify, but was mentioned by many of the tutored students' teachers. The teachers often spoke of the special relationship shared by the students and their tutors, reporting that many of the students looked forward to the tutoring. Also, many of the tutored students enjoyed sharing their progress with their teachers, a behavior not observed among students in the other two treatments. Perhaps there is an important motivational component associated with one-to-one instruction from an adult that leads to a stronger desire to engage in reading activities—leading, perhaps, to increased learning. Future researchers may wish to investigate this aspect of individualized tutoring through more systematic observation and interviewing.

We hoped that Modified PALS would also benefit nonresponders. Modified PALS effects would indicate that the needs of struggling readers might be met using less

costly resources more readily available to teachers than individualized tutoring. However, Modified PALS was only half as successful as Tutoring in reducing rates of unresponsiveness. Specific implementation features may have contributed to this outcome. First, teachers were asked to supervise the students' implementation of the Modified PALS activities, but, given that they had an entire class to oversee at the same time, it is not clear that they were able to ensure that Modified PALS was conducted correctly. The wide range of Modified PALS fidelity (49% to 100%) indicates that it was conducted incorrectly at least some of the time. In addition, making modifications to a PALS program that was not benefiting the nonresponders was likely an inadequate response to their reading problems. Unfortunately, and importantly, this reflects what can happen when the general curriculum is modified for struggling students: classroom teachers are not always able to monitor them closely, and instruction is basically "watered down" rather than individualized (D. Fuchs & Fuchs, 1994; Vaughn & Schumm, 1995).

Findings indicate that classroom instruction—even generally effective classroom instruction—can be inappropriate for struggling readers, with or without modifications. For those students for whom modifications are ineffective, it is important that options such as one-to-one or small group tutorials are available. Whereas some may see such intervention as “restrictive,” especially if located outside the general education classroom, it is important to remember why federal law compels educators to provide such an option—to ensure that education is not only provided in the least restrictive environment, but that it is also most appropriate for meeting students’ unique learning needs.

Study Limitations

There are at least several study limitations, the first of which is that the sample size was small, reducing statistical power. It should be noted, however, that the participants were drawn from a pool of 323 students, which represents a large field-based study. A much larger pool of students would be needed to generate sufficiently large numbers of nonresponders to power the necessary inferential statistical analyses. Second, 15 of the 56 nonresponders (27%) were ELL students. Although these students spoke English well enough to interact with their peer or adult tutors, language differences may have complicated treatment effects. For example, several ELL students were observed to make very rapid gains, suggesting that, as they learned more English, they began to overcome their reading difficulties. Conversely, some ELL students made little growth; it is difficult to determine whether their unresponsiveness was due primarily to reading deficits or to severe language problems. Third, this study did not include a no-treatment control group, which would have been useful in determining whether the three treatments were more beneficial for unresponsive readers than more traditional classroom instruction. Finally, our pretreatment measures were administered before PALS began in October. Modified PALS and Tutoring did not begin until January. Whereas we used the students’ December Dolch level as a covariate in comparing posttest scores, additional information about group equivalence on other reading related measures immediately before the additional treatments were implemented would have strengthened our analyses.

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

Implications for Research and Practice

Further research is needed to explore features of early intervention that should be in place to maximize the learning of struggling readers. For example, standardized instructional programs should be compared with instruction tailored to individual needs. Characteristics of children unresponsive to treatment should continue to be studied closely. It may be important to include components that address important child characteristics such as attention, motivation, and behavior. The ideal size of instructional group (e.g., whole class vs. small group vs. one-on-one) should also be further explored, and interventions that are supplemental to regular instruction should be compared with interventions that take the place of regular instruction. In addition, the length of intervention needed to produce strong and stable growth must be investigated. And follow-up studies are needed to better understand long-term benefits of early reading intervention.

Finally, researchers should examine ways of implementing instruction that is effective but also feasible, given finite school resources. Interventions that can be implemented by classroom teachers, paraprofessionals, parents, and school volunteers are likely to be more accessible, and thus more widely beneficial, than those that require special training and many hours outside the general classroom. Someday, perhaps, we will know the conditions necessary for all children to learn to read. However, if and when that day arrives, the work of early intervention researchers will still not be complete. A critical goal of future research should be not only to develop the most effective interventions, but also to determine the most appropriate and efficient means of delivering them.

We offer several implications for practice based on findings from this study and previous research. First, whereas we suggest all teachers implement evidence-based instruction that meets a wide range of needs, we also caution that such instruction should not be viewed as a "cure-all." Ongoing progress-monitoring of students, especially those at risk for, or identified as having, reading disabilities, is critical for determining whether an instructional strategy is beneficial. We suggest that teachers use frequent curriculum-based measures to ensure that their students are making progress; to identify students who are discrepant from their peers in performance level and growth rate; and to make changes to group or individual instruction when students are not progressing as expected. Second, when modifications are made to general education instruction, they should be implemented with fidelity, and concurrently with ongoing progress monitoring. Finally, when evidence-based instruction and modifications in general education fail to meet students' unique learning needs, we encourage practitioners to consider more intensive, individualized instruction. Recent research suggests that such "special-education-like" instruction is the best response to children most at-risk for reading failure.

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

References

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*, 300-315.

Berninger, V. W., Abbott, R. D., Vermeulen, K., Ogier, S., Brooksher, R., Zook, D., & Lemos, Z. (2002). Comparison of faster and slower responders to early intervention in reading: Differentiating features of their language profiles. *Learning Disability Quarterly, 25*, 59-76.

Berninger, V. W., Abbott, R. D., Zook, D., Ogier, S., Lemos-Britton, Z., & Brooksher, R. (1999). Early intervention for reading disabilities: Teaching the alphabet principle in a connectionist framework. *Journal of Learning Disabilities, 27*, 491-503.

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

Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R., & McGraw, C. K. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing: An Interdisciplinary Journal, 11*, 239-273.

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.

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

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

Deno, S. L. (l985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219-232.

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

Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly, 27*, 13-26.

Foorman, B. R. Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 68*, 70-74.

Fuchs, D., & Fuchs, L. S. (1994). Inclusive schools movement and the radicalization of special education reform. *Exceptional Children, 60*, 294-309.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.

learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34*, 174-206.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N., Braun, M., & O'Connor, R. (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., Svenson, E., Yen, L., Al Otaiba, S., Yang, N., McMaster, K. N., Prentice, K., Kazdan, S., & Saenz, L. (2001). Peer-assisted learning strategies in reading: Extensions for kindergarten, first grade, and high school. *Remedial and Special Education, 22*, 15-21.

Fuchs, D., Fuchs, L. S., Yen, L., McMaster, K., Svenson, E., Yang, N., Young, C., Morgan, P., Gilbert, T., Jaspers, J., Jernigan, M., Yoon, E., & King, S. (2001). Developing first-grade reading fluency through peer mediation. OSEP Project Report.

Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice, 18,* 172-186.

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., Fuchs, D., & Hamlett, C. L. (1989). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. *Journal of Educational Research, 83*, 103-111.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, 20-29.

Good, R. H., & Kaminski, R. A. (Eds.). (2001). *Dynamic indicators of basic early literacy skills* (5th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H., Kaminski, R., & Shinn, M. (1999, January). *Growth and development indicators: From development to refinement and back again.* Paper presented at the Pacific Coast Research Conference, San Diego, CA.

Good, R. H., Kaminski, R. A., Shinn, M., Bratten, J., Shinn, M., & Laimon, L. (2002). *Technical adequacy and decision making utility of DIBELS* (Technical Report). Manuscript in preparation. Eugene, OR: University of Oregon.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*, 41-57.

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

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

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

Mathes, P. M., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade read-

ers: Responding to the needs of diversity. *Reading Research Quarterly, 33*, 62-94.

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

President's Commission on Excellence in Special Education. (July 1, 2002). *A New Era: Revitalizing Special Education for Children and their Families.* Retrieved December 1, 2002, from http://www.ed.gov/inits/commissionsboards/whspecialeducation/index.html

Psychological Corporation. (1992). *Wechsler individual achievement test.* San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich, Inc.

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

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

Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93,* 735-749.

Speece, D. L., Case, L. P., & Molloy, D. E. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research and Practice, 18,* 147-156.

Torgesen, J. K. (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., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*, 33-58, 78.

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

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

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., Gersten, R., & Chard, D. J. (2000). The underlying message in LD intervention research: Findings from re-

search syntheses. *Exceptional Children, 67*, 99-114.

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.

Vaughn, S., & Schumm, J. S. (1995). Responsible inclusion for students with learning disabilities. *Journal of Learning Disabilities, 28*, 264-270, 290.

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

Wilson, D. B. (1996). *Effect Size Determination Program (Version 1.0)* [Computer Program]. Nashville, TN: Vanderbilt University.

Woodcock, R. W. (1987). *Woodcock reading mastery test-Revised*. Circle Pines, MN: American Guidance Service.

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

McMaster, K. L., Fuchs, D., Fuchs, L. S., and Compton, D. L.(2003, December). *Responding to Nonresponders: An Experimental Field Trial of Identification and Intervention Methods.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.