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

KRISTEN L. McMASTER
University of Minnesota

DOUGLAS FUCHS
LYNN S. FUCHS
DONALD L. COMPTON
Peabody College of Vanderbilt University

ABSTRACT: First graders (N = 323) participated in an evidence-based classwide reading program (Peer-Assisted Learning Strategies; PALS). A dual-discrepancy approach was used to identify 56 children whose reading performance and growth rates were substantially below those of average readers, indicating they were not responding sufficiently to PALS. This approach reliably distinguished among unresponsive at-risk, responsive at-risk, and average-performing readers. Nonresponders were assigned randomly to one of three increasingly individualized treatments: PALS, Modified PALS, or tutoring by an adult. No statistically significant between-group differences on reading-related measures were found. Effect sizes (between .30 and .50) comparing groups and proportions of nonresponders following treatment suggest that tutoring was most promising for reducing unresponsiveness.

Reading research over the past 20 years has greatly advanced our understanding of reading problems. We know that many 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; Share & Stanovich, 1995). Moreover, a number of researchers (e.g., Blachman, Ball, Black, & Tangel, 1994; Byrne & Fielding-Barnsley, 1991; Fuchs, Fuchs, Thompson, Svenson, Yen, Al Otaiba, et al., 2001; Torgesen, Wagner, & Rashotte, 1997) 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 well to even the most effective interventions. Researchers have reported that 20% to 30% of children at risk for reading difficulties (see Torgesen, 2000), and more than 50% of children with disabilities (e.g., Fuchs, Fuchs, Thompson, et al., 2001; Torgesen et al., 2001) do not respond adequately to generally effective early reading intervention. Such students have been dubbed nonresponders. Recently, responsiveness-to-intervention has been suggested as an alternative 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, 2003; O'Connor, 2000; Speece & Case, 2001; Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino et al.).

IDENTIFYING NONRESPONDERS

It is generally agreed that nonresponders to reading interventions are students who do not make adequate reading progress despite participating in evidenced-based instruction. However, there is little agreement about what constitutes "adequate reading progress" (Torgesen, 2000). In an extensive review of research addressing responsiveness to reading instruction, Al Otaiba and Fuchs (2002) showed that researchers have used one of two types of indicators of reading progress: performance level or growth rate. In terms of performance level, researchers have defined unresponsiveness as performance below the 10th percentile to below the 50th percentile on a given measure (e.g., Foorman, Francis, Fletcher, Schatschneider, & Mehta, 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) or limited growth (e.g., Vellutino et al.).

There are serious conceptual problems related to performance-level or growth-rate-only approaches. For example, a child's performance level may be low even though he or she is making important growth. Without considering his or her growth rate, it is difficult to determine whether the child is responsive to intervention. 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 he or 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 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). Others are testing its utility by comparing it to alternative procedures like median split, normalized, and benchmark scores (see D. Fuchs, Fuchs, & Compton, 2004; L. Fuchs, 2003). Continued research is needed to determine the technical adequacy and utility of the dual-discrepancy approach.

TREATING NONRESPONDERS

A second question important to the study of nonresponders is how to address such children's difficulty with learning to read. A number of reading intervention researchers reported proportions of nonresponders (e.g., Berninger et al., 1999; Foorman et al., 1998; Torgesen et al., 1999). However, only a few attempted multiphased interventions in an attempt to decrease rates of unresponsiveness (e.g., Case et al., 2003; O'Connor, 2000;
Vaughn et al., 2003; Vellutino et al., 1996). In this subgroup of ambitious studies, the first phase consisted of instruction for students identified as at risk for reading failure, and subsequent phases involved instruction directed at nonresponders in the previous phase. Some of these studies defined unresponsiveness in the context of general education instruction; others in the context of more intensive, small-group instruction.

**Special Education-Like Approaches**

Several researchers 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 reading intervention from well-trained teachers. Vellutino et al., for example, implemented an intensive, one-to-one tutoring intervention for 13 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 meet exit criteria after another 10 weeks received a third phase. After 30 weeks of intervention, less than 23% of the original at-risk sample had not met the exit criteria.

Whereas researchers working within this special education-like framework demonstrated that many poor readers improve when intensive intervention is in place, they have not examined whether or how general educators might modify instruction, curricula, or materials to accommodate the 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 the best ways to meet their needs.

We offer two reasons for this view. First, as Vellutino et al. (1996) suggested, some students’ reading difficulties are no doubt due to inadequate instruction rather than to 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, 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. (2003) did precisely this. O’Connor implemented four increasingly intensive levels of reading interventions to beginning readers. Intervention at the first level was an evidence-based, whole-class, phonological awareness program conducted by kindergarten teachers. Unresponsive kindergartners 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.

Similarly, Case et al. (2003) defined unresponsiveness as the difference between poor readers’ and their peers’ growth rates and performance levels in the mainstream classroom. The researchers then worked with the classroom teachers of nonresponders to implement research-based phonological awareness and phonics instruction, partner reading, or computer programs. Students who received the classroom interventions made greater reading gains than those who did not (Speece et al., 2003). Like O’Connor’s (2000) findings, these results suggest that evidence-based,
classroom intervention may be beneficial for some nonresponders.

**Comparing Levels of Service Delivery**

The special education-like approach described earlier is important to research examining responsiveness to intervention, in part because it identifies the most difficult-to-teach students in relatively few steps: Nonresponders 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 curriculum, it is likely that most schools will encourage teachers to implement evidence-based classroom instruction and modifications before removing lagging students for more individualized instruction (Vaughn, Gersten, & Chard, 2000). Moreover, schools typically do not have the resources for intensive interventions for all at-risk children. Thus, we believe that students' responsiveness to general education instruction (modified or unmodified) should be included in research on nonresponders. As indicated, researchers working within a general education framework have begun to do this. What is 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 of generally effective classroom instruction, and (c) more individualized one-to-one intervention in strengthening struggling readers' achievement.

**Purpose of the Study**

The purpose of this study was twofold. First, we further explored the validity of the dual-discrepancy approach (e.g., 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 an evidence-based classwide intervention, Peer-Assisted Learning Strategies (PALS; e.g., D. Fuchs et al., 2001). 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.

More comprehensive examinations of unresponsiveness should begin by determining whether modifying instruction in the general education classroom is effective.

Second, we compared the achievement of nonresponders who continued receiving PALS to nonresponders who received one of two treatments modeled after part of the continuum of services available to students 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 instruction or modifications of such instruction (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. Although 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., Torgesen et al., 1999; 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 pro-
gram from October 2000 to April 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 who volunteered to participate 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 PALS (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, (b) monitoring these 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 screened using a Rapid Letter Naming (RLN) test, an effective predictor of future reading achievement (Torgesen et al., 1997). Within each class, students were rank-ordered from highest to lowest based on their RLN scores. Adjustments to these rankings were made based on teacher judgment. Using the adjusted rankings, we identified 4 average-performing and the 8 lowest-performing students in each class. By identifying the 8 lowest-performing students in each class, we deliberately oversampled the number of at-risk children to reduce the possibility of false negatives. These lowest-performing students were considered at risk for unresponsiveness to PALS (n = 176). The average-performing students served as a comparison group (n = 88).

**Monitoring Progress.** From October to December 2000, the at-risk and average-performing students' reading progress was monitored weekly. Monitoring measures included criterion-referenced chapter tests of PALS sounds and words and two word-level CBM measures. These were Nonword Fluency (NWF) probes from the Dynamic Indicators of Basic Early Literacy Skills (Good & Kaminski, 2001), and Dolch word probes consisting of high-frequency preprimer, primer, and first-grade-level words.

**Identifying Nonresponders.** Nonresponders were identified in January 2001, after students participated in PALS for three 35-min sessions per week for 7 weeks. Although 7 weeks is a relatively brief period, we reasoned that waiting longer may have resulted in students falling further behind, making remediation more difficult. After 7 weeks, complete monitoring data were available for 166 at-risk and 87 average students (reflecting 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 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. We calculated slope using least squares regression between monitoring scores and calendar days.

Second, we calculated means and standard deviations (SD) of the average-performing students' CBM levels and slopes. Third, all students from each classroom were identified who had scored 90% or less, or who were the lowest scoring in their class, on the last PALS chapter test, (n = 97). Fourth, z-scores were calculated separately on the CBM levels and slopes of these 97 students, using the means and SDs of the average-performing students. Finally, we examined the z-scores of each student. Students scoring more than 0.50 SD below the average readers in terms of both level and slope on the CBM measures were identified as nonresponders (n = 66).

**Assigning Nonresponders to Groups.** Within the Standard PALS and PALS + Fluency classes, the 66 nonresponders were stratified by low (n = 28) versus very low (n = 38) status based on CBM levels and slopes. Then, 22 students were assigned randomly to PALS; 22 to Modified PALS; and 22 to Tutoring. Each of these groups reflected the same proportion of low versus very low students,
and the same proportion participating in Standard PALS versus PALS+ Fluency treatments. Because of attrition, at the end of the study, there were 21, 15, and 20 students in the PALS, Modified PALS, and Tutoring treatments, respectively, for a total of 56 nonresponders.

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 \( F[2, 53] = 1.09, p = 0.34 \), NWF level \( F[2, 53] = 0.24, p = 0.79 \), and NWF slope \( F[2, 53] = 0.94, p = 0.40 \) before Modified PALS and Tutoring began. The three groups were also compared on demographic variables including sex, ethnicity, English Language Learner (ELL) status, Title I status, and special education status. Chi-square tests indicated no statistically significant differences across conditions on these variables (all \( p > 0.05 \)).

**Treatments**

**PALS.** First-Grade PALS was developed by researchers at Vanderbilt University (D. Fuchs & Fuchs, in press; Fuchs, Fuchs, Svenson, Yen, Thompson, McMaster, et al., 2001; Mathes, Howard, Allen, & Fuchs, 1998). PALS is a structured peer tutoring program that emphasizes phonological awareness, decoding, and fluency, all skills that have been demonstrated to be important for successful beginning reading programs (e.g., Blachman, Tangel, Ball, Black, & McGraw, 1999; Hatcher, Hulme, & Ellis, 1994; Torgesen et al., 1997). Results of large-scale experimental studies indicate that First Grade PALS helps develop beginning reading skills for a majority of low-, average-, and high-achievers, and for many children with disabilities (e.g., Fuchs, Fuchs, Svenson, et al., 2001; Mathes et al., 1998). PALS has proved effective for many students in schools with high proportions of minority children and children living in poverty, as well as in schools with predominantly white, middle-class populations.

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. Standard PALS and PALS + Fluency were implemented 3 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. PALS + Fluency included 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 in the same amount of time. The First Grade PALS manual (Fuchs, Fuchs, Svenson et al., 2001) may be obtained by visiting http://www.peerassistedlearningstrategies.net or by e-mailing PALS@vanderbilt.edu.

**Modified PALS.** Modified PALS was conducted in the classroom during the scheduled PALS time for 35 min, 3 times per week. The teachers selected Coaches who could read the PALS lessons independently and worked well with lower performing students. Modified PALS activities were similar to PALS in several ways. The activities included letter–sound recognition, decoding, sight word recognition, and reading short stories. The Coach and Reader roles were reciprocal (although the Coach read from the regular PALS lesson), and involved using similar prompts and correction procedures. However, Modified PALS incorporated three important alterations. First, fewer sounds and words were introduced at one time, and the Readers worked on lessons that matched their skill level. Second, the Coach modeled the sounds and words for the Reader before providing the Reader opportunities to read without a model. Third, greater emphasis was placed on phonological awareness and decoding.
ing skill. During the Sound it Out activity, the Coach modeled three steps with each word: segmenting, blending, and decoding. The Reader repeated the Coach at each step, and then conducted the three steps independently.

Tutoring. Tutoring took place three times per week for 35 min per session and substituted for PALS. As in PALS and Modified PALS, Tutoring included letter-sound recognition, decoding, sight word recognition, and reading short stories. However, Tutoring was designed to more closely resemble a special education-like pull-out approach. The activities were based on research indicating that small-group or one-to-one direct instruction in phonological awareness and decoding have a positive effect on the performance of struggling readers and students with disabilities (e.g., Blachman et al., 1999; Hatcher et al., 1994; Torgesen et al., 1997).

Tutored students received one-to-one tutoring from a trained research assistant 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, a motivational component was added. 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. At the end of the activity, the student counted the number of sounds or words mastered, and graphed this number in relation to the goal.

PRE- AND POSTTEST MEASURES

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

Rapid Naming. There were two rapid naming measures given to assess letter and sound naming skill. This first, RLN, was developed for use in a previous PALS study (Fuchs, Fuchs, Thompson, Svenson, Yen, Al Otaiba, et al., 2001) and 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 1 min. 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 (Fuchs, Fuchs, Thompson, 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 1 min.

Phonological Awareness. The ability to segment words into phonemes correlates well 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., Fuchs, Fuchs, Thompson, et al., 2001) was administered. Students were told to say the sounds in each word provided. The score was the number of phonemes expressed correctly in 1 min.

A blending task, developed previously (Fuchs, Fuchs, Thompson, 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." The score was the number of words blended correctly in 1 min.

Reading Words. The Word Identification (Word ID) 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 ID 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. The score was the number of words written correctly. The WIAT correlates well (.70s to .80s) with other 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 1 min.

**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 at posttest. The passages were traditional folktales. Students were instructed to read the stories quickly and correctly. The score was recorded as the number of words read correctly in 1 min. Test-retest reliability for the fluency measure ranged from .93 to .96, and concurrent validity with the comprehension subtest of the Stanford Achievement Test (SAT) was .91 (L. Fuchs, Fuchs, & Maxwell, 1988). The passages were followed by 10 open-ended comprehension questions. For number of questions answered correctly, test–retest reliability was .92; the correlation with the SAT was .82 (L. Fuchs, Fuchs, & Maxwell, 1988).

**Monitoring Measures**

**Chapter Tests.** Research staff developed seven chapter tests. These tests looked very similar to PALS lessons and covered sounds, decodable words, and sight words presented in PALS before the test was administered. The tests were cumulative and untimed. The score was recorded as the percentage of sounds and words read correctly.

**Dolch Probes.** Research staff developed Dolch probes for this study. These probes were 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 1 min.

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

**Procedures**

**Test Administration Training.** Research staff were two full-time project coordinators and eight graduate students in special education, counseling, or educational administration. All had experience working in schools. They were taught during several training sessions to administer the pre- and posttreatment and monitoring measures. 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.

**Pre- and Posttreatment Testing.** PALS participants were tested individually in a quiet location (such as a library or conference room) in two sessions. During the first session, the examiner spent several minutes establishing rapport with the student. The RLN 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. At posttest, 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. Staff did not test students whom they had tutored, to avoid examiner bias.

**Monitoring Progress.** For the first 7 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 Standard PALS and PALS + Fluency teachers 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 trained nonresponders and their Coaches to conduct Modified PALS and provided weekly technical assistance. Tutors worked one-on-one with their assigned students. The PALS, Modified PALS, and Tutoring treatments were implemented for an additional 13 weeks (January to April).

Treatment Fidelity. For PALS, Modified PALS, and Tutoring, a checklist of teacher and student behaviors was developed to assess treatment fidelity. Fidelity checks of the 22 PALS classrooms were conducted in December and March. 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%. The most common reasons for low fidelity scores were that Coaches did not consistently model sounds and words for the Readers or they did not correct the Readers' mistakes.

Because of scheduling problems, it was not possible 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 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.

RESULTS

IDENTIFYING NONRESPONDERS

One purpose of this study was to explore a dual-discrepancy approach to identifying nonresponders to 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 while minimizing false positives and false negatives.

As indicated, we first identified average readers and a risk pool of low-performing readers, using RLN 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 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 7 weeks of PALS. To determine criterion validity and test–retest reliability, Pearson rs were calculated among the monitoring levels and posttreatment scores. Dolch levels correlated strongly with Word ID (.82), Word Attack (.52), near-transfer fluency (.92), far-transfer fluency (.93), and comprehension (.73; all p's < .01). NWF levels also correlated well with Word ID (.65), Word Attack (.51), near-transfer fluency (.78), far-transfer fluency (.80), and comprehension (.54; all p's < .01). Test–retest correlations were .88 for Dolch and .87 for NWF. Results suggest these 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, CBM levels and slopes of these three groups were compared using ANOVAs. Not surprisingly, there were statistically significant differences among the three
TABLE 1
Comparisons of At-Risk and Average-Performing Students on Pretreatment Measures

| Measure          | At-Risk * (n = 174) | Average * (n = 87) | F   (df) | ES *
|------------------|---------------------|--------------------|--------|-------
|                  | M       (SD)        | M       (SD)        |        |       |
| Rapid Letter Naming | 34.02 (13.40) | 48.75 (10.37) | 80.74* (259) | 1.29 |
| Rapid Letter Sound  | 20.60 (13.96) | 31.75 (11.51) | 41.36* (259) | .90  |
| Segmentation       | 20.40 (14.68) | 26.14 (12.61) | 9.70* (259)  | .40  |
| Word ID            | 9.20 (9.54)  | 17.63 (9.75)  | 44.63* (259) | .84  |
| Word Attack        | 3.45 (5.27)  | 6.92 (4.74)   | 26.87* (259) | .72  |
| Blending           | 12.44 (7.47) | 16.52 (6.66) | 18.58* (259) | .49  |
| Spelling           | 8.83 (3.48)  | 11.63 (3.02) | 40.49* (258) | .88  |

Note. ES = Effect size. Effect sizes were calculated by dividing mean difference by the SD of the average performers.

At-Risk data were not available for 2 at-risk students and 1 average-performing student. *p < .01

TABLE 2
Comparisons of At-Risk Nonresponders, At-Risk Responders, and Average-Performers in December

<table>
<thead>
<tr>
<th>Measures</th>
<th>Responders * vs. Nonresponders</th>
<th>Average Performers * vs. Nonresponders</th>
<th>Average Performers vs. Responders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Difference</td>
<td>SE</td>
<td>ES</td>
</tr>
<tr>
<td>Dolch</td>
<td>8.41**</td>
<td>1.86</td>
<td>0.85</td>
</tr>
<tr>
<td>Level</td>
<td>0.57**</td>
<td>0.15</td>
<td>0.57</td>
</tr>
<tr>
<td>Slope</td>
<td>17.34**</td>
<td>2.17</td>
<td>1.35</td>
</tr>
<tr>
<td>NWF</td>
<td>0.97*</td>
<td>0.34</td>
<td>0.41</td>
</tr>
</tbody>
</table>

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

*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 454 Summer 2005
2 (PALS program x treatment x nonresponder status) ANOVAs. There were no statistically significant pretreatment main effects or interactions.

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 differences, we were concerned that possible between-group differences may not have been detected because of low statistical power. We used ANCOVA with December Dolch level as the covariate because it was a more proximate pretreatment measure than those administered in October. The homogeneity of regression assumption was met for these analyses. There were no statistically significant main effects or interactions of interest. To further explore possible systematic posttreatment differences between the two PALS programs (Standard PALS and PALS + Fluency), their respective means, SDs, and effect-size comparisons were examined for each treatment condition (PALS, Modified PALS, and Tutoring). No discernable patterns were found. Complete results from these analyses may be obtained from the first author.

Because of no statistically significant differences between Standard PALS and PALS + Fluency, and no pattern of effect sizes suggesting systematic differences involving the two PALS programs, we simplified the design of our analysis to one between-group ANCOVAs. Table 3 shows means and SDs of the pre- and posttreatment and growth scores of the three treatment groups on all measures. No statistically significant between-group differences were found on any of these measures.

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 moderate differences.

Because statistical analyses were 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. Small-to-moderate effects, reported in Table 4, were found favoring (a) Tutoring over PALS on Word ID, Word Attack, blending, and comprehension; (b) Tutoring over Modified PALS on Word ID, blending, and spelling; and (c) Modified PALS over PALS on Rapid Letter Sound, Word Attack, and comprehension.

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 persistent nonresponders following the 13 weeks of additional treatment. As in December, students were identified as persistent nonresponders if they scored more than 0.50 SD below average performers' levels and slopes on the CBM measures. Accordingly, 81% of PALS students, 80% of Modified PALS students, and 50% of Tutoring students remained unresponsive to treatment at the end of the study. Chi square analyses indicated that the proportion of nonresponders to PALS was statistically significantly higher than the proportion of nonresponders to Tutoring ($\chi^2 = 4.36, p < .05$). No other statistically significant differences were found (all $p$s > .05).

Overall, 70% of the initial nonresponders were persistent nonresponders. 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 Standard PALS or PALS + Fluency students ($n = 323$). Chi square analyses indicated no statistically significant differences involving sex, race, Title I status, or special education status between students who responded to one of the three treatments and persistent nonresponders, or among persistent nonresponders across the three treatments (all $p$s > .05).
<table>
<thead>
<tr>
<th>Measures</th>
<th>PALS (n = 21)</th>
<th>Modified PALS (n = 15)</th>
<th>Tutoring (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td><strong>Pre- and Posttest</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid Letter Naming</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>30.14 (7.64)</td>
<td>27.60 (12.84)</td>
<td>27.30 (12.83)</td>
</tr>
<tr>
<td>Post</td>
<td>47.43 (17.11)</td>
<td>44.47 (19.23)</td>
<td>45.80 (19.19)</td>
</tr>
<tr>
<td>Growth</td>
<td>17.28 (13.96)</td>
<td>16.87 (13.50)</td>
<td>18.50 (15.83)</td>
</tr>
<tr>
<td>Rapid Letter Sound</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>13.29 (8.44)</td>
<td>18.20 (15.62)</td>
<td>12.90 (9.42)</td>
</tr>
<tr>
<td>Post</td>
<td>41.71 (11.15)</td>
<td>48.27 (22.75)</td>
<td>45.50 (15.54)</td>
</tr>
<tr>
<td>Growth</td>
<td>28.42 (9.11)</td>
<td>30.07 (16.44)</td>
<td>32.60 (13.59)</td>
</tr>
<tr>
<td>Segmentation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>13.62 (13.00)</td>
<td>18.53 (13.19)</td>
<td>15.05 (13.68)</td>
</tr>
<tr>
<td>Post</td>
<td>35.52 (13.39)</td>
<td>33.87 (14.91)</td>
<td>35.25 (14.04)</td>
</tr>
<tr>
<td>Growth</td>
<td>21.90 (12.40)</td>
<td>15.33 (12.61)</td>
<td>20.20 (13.03)</td>
</tr>
<tr>
<td>Word ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>2.86 (2.85)</td>
<td>5.40 (5.53)</td>
<td>4.65 (6.05)</td>
</tr>
<tr>
<td>Post</td>
<td>20.62 (9.56)</td>
<td>20.47 (10.13)</td>
<td>25.60 (9.51)</td>
</tr>
<tr>
<td>Growth</td>
<td>17.76 (9.74)</td>
<td>15.07 (8.33)</td>
<td>20.95 (9.78)</td>
</tr>
<tr>
<td>Word Attack</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>1.10 (1.87)</td>
<td>2.27 (3.08)</td>
<td>1.40 (2.46)</td>
</tr>
<tr>
<td>Post</td>
<td>6.71 (5.26)</td>
<td>8.67 (5.55)</td>
<td>8.95 (5.36)</td>
</tr>
<tr>
<td>Growth</td>
<td>5.62 (5.10)</td>
<td>6.40 (4.45)</td>
<td>7.55 (4.99)</td>
</tr>
<tr>
<td>Blending</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>9.67 (7.07)</td>
<td>10.00 (6.01)</td>
<td>10.75 (7.01)</td>
</tr>
<tr>
<td>Post</td>
<td>19.38 (8.54)</td>
<td>19.00 (5.94)</td>
<td>22.70 (7.14)</td>
</tr>
<tr>
<td>Growth</td>
<td>9.71 (6.70)</td>
<td>9.00 (6.23)</td>
<td>11.95 (7.56)</td>
</tr>
<tr>
<td>Spelling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>6.86 (2.01)</td>
<td>7.47 (2.59)</td>
<td>6.95 (3.33)</td>
</tr>
<tr>
<td>Post</td>
<td>12.67 (3.37)</td>
<td>11.27 (3.61)</td>
<td>12.45 (2.86)</td>
</tr>
<tr>
<td>Growth</td>
<td>5.81 (3.28)</td>
<td>3.80 (2.75)</td>
<td>5.50 (3.75)</td>
</tr>
<tr>
<td>Posttest Only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Near-Transfer Fluency</td>
<td>18.26 (12.55)</td>
<td>19.07 (12.71)</td>
<td>22.20 (10.68)</td>
</tr>
<tr>
<td>Far-Transfer Fluency</td>
<td>19.29 (12.22)</td>
<td>19.90 (13.69)</td>
<td>22.95 (10.08)</td>
</tr>
<tr>
<td>Comprehension</td>
<td>0.38 (0.55)</td>
<td>0.70 (0.92)</td>
<td>0.68 (0.82)</td>
</tr>
</tbody>
</table>
### TABLE 4
Posttreatment Effect Sizes by Nonresponder Treatment

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

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

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

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Dual-Discrepancy</th>
<th>Performance-Level-Only</th>
<th>Growth-Rate-Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>PALS (n = 21)</td>
<td>17 (81)</td>
<td>8 (38)</td>
<td>21 (100)</td>
</tr>
<tr>
<td>Modified PALS (n = 15)</td>
<td>12 (80)</td>
<td>8 (53)</td>
<td>15 (100)</td>
</tr>
<tr>
<td>Tutoring (n = 20)</td>
<td>10 (50)</td>
<td>9 (45)</td>
<td>20 (100)</td>
</tr>
<tr>
<td>Total nonresponders (N = 56)</td>
<td>39 (70)</td>
<td>25 (45)</td>
<td>56 (100)</td>
</tr>
<tr>
<td>Total risk pool (N = 174)</td>
<td>39 (22)</td>
<td>25 (14)</td>
<td>56 (32)</td>
</tr>
<tr>
<td>Total PALS participants (N = 323)</td>
<td>39 (12)</td>
<td>25 (8)</td>
<td>56 (17)</td>
</tr>
</tbody>
</table>

"Dual-discrepancy was determined by slopes and levels > 0.50 SD below average on the Dolch and/or Nonword Fluency measures. "Percentile rank was determined by scores below the 30th percentile on the WRMT-R Word ID and/or Word Attack subtests. "Criterion level was determined by < 40 correct words per min on the Near-Transfer and/or Far-Transfer Fluency measures. "No growth was determined by a gain of 0 words (or negative growth) on Word ID and/or Word Attack. "Limited growth was determined by a gain of 10 words or less on Word ID and/or a gain of 5 words or less on Word Attack."
Performance-Level-Only and Growth-Rate-Only Approaches. To determine whether the procedures we used to identify nonresponders yielded similar proportions as procedures used by other researchers, two alternative approaches to identifying nonresponders (performance-level-only and growth-rate-only) were explored. Performance levels and growth rates on the Word ID and Word Attack subtests and the near- and far-transfer fluency measures were examined (see Berninger et al., 1999; Fuchs, Fuchs, Thompson, et al., 2001; Torgesen et al., 1999; Vellutino et al., 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 approaches, we first determined how many of the initial 56 nonresponders for whom we had 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 ID and Word Attack subtests or limited growth (less than 10 words gained on the Word ID 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 criterion (see 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. However, 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 as 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 et al., 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 utility of a dual-discrepancy approach to identifying nonresponders (see L. Fuchs & Fuchs, 1998). Findings show that our CBM measures were reliable, valid indicators of students’ reading skill, and that the dual-discrepancy approach identified students performing reliably and substantially 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).

Findings also suggest that dual-discrepancy holds promise as a better method of identification than performance-level and growth-rate-only approaches. In this study, the performance-level-only approach would identify a student as a nonresponder who scored below the 30th percentile on the Word ID or Word Attack subtests. The dual-discrepancy approach would not identify such a student as a nonresponder if he or 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 average-performing peers, he or she probably would not need additional instruction. Our dual-discrepancy approach provided both perfor-
mance-level and growth-rate criteria (based on average students' performance) needed to make such decisions.

Nevertheless, we offer three important caveats about our approach. First, just as performance-level and growth-rate criteria used by previous researchers are arbitrary, our discrepancy criterion of 0.50 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. Further, some may argue that established grade-level benchmarks are needed to determine whether students are progressing at a sufficient rate to succeed in school. Our approach did not make use of such benchmarks. Finally, nonresponders 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, fluency, or comprehension to identify nonresponders. Continued research is needed to determine which measures are most appropriate, and to establish performance levels and growth rates that serve as indicators of whether sufficient progress is made, 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, it is likely that the treatments' relatively low intensity and short duration were simply insufficient to yield dramatic gains in students with severe reading difficulties. Although we viewed the Tutoring intervention as a special education-like approach, it was still only a supplement to general classroom instruction and was not as intensive as we would expect traditional pullout instruction to be. Others who have implemented special education-like interventions (e.g., Foor- man et al., 1998; Torgesen et al., 2001; Vellutino et al., 1996) conducted treatments for similar students totaling 65 hours to more than 300 hours, some spanning 2 years or more. These interventions have often been implemented by highly trained teachers or reading specialists. Our intention was to develop a more practical, less costly approach to individualized instruction; for some students, however, it was not enough.

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. Perhaps nonresponders would benefit from qualitatively different activities in addition to an increased focus on individualization. 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. Because effect sizes were not statistically significant, they should be viewed cautiously. However, the reliably smaller proportion of students who were persistently unresponsive to Tutoring as compared to PALS, together with a pattern of effect sizes favoring Tutoring suggests that Tutoring may be the more promising treatment approach.

Assuming this is the case, Tutoring and Modified PALS deserve some additional comment that may be helpful in considering future directions for research and practice. 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 assis-

---

**Findings also suggest that dual-discrepancy holds promise as a better method of identification than performance-level and growth-rate-only approaches.**
tant was trained to ensure that immediate, corrective 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 adults are able to provide.

Another feature of 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, adult tutoring that leads to a stronger desire to engage in reading activities—leading eventually, perhaps, to increased learning. Future researchers may wish to investigate this aspect of individualized tutoring through more systematic observation and interviewing.

We also hoped that Modified PALS would benefit nonresponders and show 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 unsuccessful in substantially 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, 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 not always conducted as intended. In addition, making modifications to a 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 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 sufficient 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 made 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 this 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 conventional classroom instruction. Finally, our pretreatment measures were administered before PALS began in October. Modified PALS and Tutoring did not begin until January. Although 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.

**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. Group size (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. Follow-up studies are needed to better understand long-term benefits of early reading intervention. Finally, researchers should examine ways of implementing instruction that are effective but also feasible, given finite school resources. Interventions that can be implemented by classroom teachers, assistants, 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.

We also offer several implications for practice based on our findings and previous research. First, although we suggest all teachers implement evidence-based classroom instruction, 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 frequently monitor student progress to identify students discrepant from their peers in performance level and growth rate and 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. 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 emphasizing phonological awareness and decoding skill. Recent research suggests that such special education-like instruction is the best response to children most at risk for reading failure.

**References**


**ABOUT THE AUTHORS**

**KRISTEN L. MCMASTER** (CEC Chapter #190), Assistant Professor, Educational Psychology, University of Minnesota, Minneapolis.

**DOUGLAS FUCHS** (CEC Chapter #185), Professor, Special Education; **LYNN S. FUCHS** (CEC Chapter #185), Professor, Special Education; and **DONALD L. COMPTON** (CEC TN Federa-

tion), Assistant Professor, Special Education, Peabody College of Vanderbilt University, Nashville, Tennessee.

Address correspondence to Kristen McMaster, University of Minnesota, 233 Burton Hall, 178 Pillsbury Dr. S.E., Minneapolis, MN 55455. (email: mcmas004@umn.edu)

This research was supported in part by Grant #H324D000033 from the Office of Special Education Programs in the U.S. Department of Education and by Core Grant HD 15052 from the National Institute of Child Health and Development, both to Vanderbilt University. The article does not necessarily reflect the position or policy of the funding agencies, and no official endorsement by them should be inferred.

The authors wish to thank the following people for their invaluable help in implementing this research: Tracy Gilbert, Jennifer Jaspers, Maryann Jernigan, Sophia King, Paul Morgan, Nancy Yang, Loulee Yen, and Carea Young.

Manuscript received November 2003; revision accepted May 2004.

**INDEX OF ADVERTISERS**

AGS, cover 2
Captioned Media Program, p. 400
Collaborative Information Centers, p. 414
Council for Exceptional Children, p. 360, 377, 378, 430, 464, 483, 484, 486-487, 490-491, 496, cover 3, cover 4
Cumberland Therapy Services, p. 353
East Carolina University, p 485
Houghton Mifflin, p. 430
Hunter College, p. 377
Lower Kuskokwin School District, p. 430
Making the Grade, p. 482
National University, p 377
Seattle University, p. 492
Special Education Service Agency, p. 495
Wadsworth, p. 485

Exceptional Children 463
Copyright of Exceptional Children is the property of Council for Exceptional Children and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.