Reading Rescue: An Effective Tutoring Intervention Model for Language-Minority Students Who Are Struggling Readers in First Grade

Linnea C. Ehri, Lois G. Dreyer, Bert Flugman and Alan Gross

Am Educ Res J 2007; 44; 414
DOI: 10.3102/0002831207302175

American Educational Research Journal
June 2007, Vol. 44, No. 2, pp. 414–448
DOI: 10.3102/0002831207302175

Reading Rescue: An Effective Tutoring Intervention Model for Language-Minority Students Who Are Struggling Readers in First Grade

Linnea C. Ehri
Lois G. Dreyer
Bert Flugman
Alan Gross
Graduate Center, City University of New York

The Reading Rescue tutoring intervention model was investigated with 64 low–socioeconomic status, language-minority first graders with reading difficulties. School staff provided tutoring in phonological awareness, systematic phonics, vocabulary, fluency, and reading comprehension. Tutored students made significantly greater gains reading words and comprehending text than controls, who received a small-group intervention ($d = 0.70$) or neither intervention ($d = 0.74$). The majority of tutored students reached average reading levels whereas the majority of controls did not. Paraprofessionals tutored students as effectively as reading specialists except in skills benefiting nonword decoding. Paraprofessionals required more sessions to achieve equivalent gains. Contrary to conventional wisdom, results suggest that students make greater gains when they read text at an independent level than at an instructional level.

KEYWORDS: beginning reading instruction, language-minority students, paraprofessionals, reading acquisition, struggling readers, tutoring

Currently there is much interest in developing effective ways to intervene early with young children who are at risk of difficulty in learning to read, especially low–socioeconomic status (SES), language-minority students from families where English is not the first language (August & Shanahan, 2006). Studies indicate that children who fail to acquire adequate reading skills in first grade often continue to have difficulties and may never catch up (Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997; Juel, 1988; Stanovich, 1986). Children who are at risk can be identified early (Vellutino et al., 1996). Results of the National Reading Panel (2000) report showed that early intervention was more effective than later remediation. There is much interest in

evaluating intervention models to provide schools with evidence as the basis for choosing programs to teach reading. In fact, some government initiatives have made funding contingent on schools' choosing reading programs that are evidence based. The main purpose of the present study was to obtain evidence regarding the effectiveness of a comprehensive tutoring intervention model, Reading Rescue (RES), when it is applied to teach language-minority struggling readers in first grade.

The best-known tutoring intervention model aimed at struggling readers in the early grades is Reading Recovery, developed by Clay (1985, 1993b). Reviews of the effectiveness of Reading Recovery have shown that in the majority of studies, students who received this tutoring made greater gains in reading achievement during first grade than students who did not receive the program (D'Agostino & Murphy, 2004; Elbaum, Vaughn, Hughes, & Moody, 2000; Shanahan & Barr, 1995). In the D'Agostino and Murphy (2004) meta-analysis, the average effect size on standardized reading achievement tests for low-achieving students was 0.32. Studies have also shown that if Reading Recovery instruction is enriched with more phonemic awareness or systematic phonics instruction, its impact on reading is enhanced (Greaney, Tunmer, & Chapman, 1997; Hatcher, Hulme, & Ellis, 1994; Iversen & Tunmer, 1993; Santa & Hoien, 1999). In these studies, effect sizes on reading ranged from 0.31 to 0.76. A question of interest in the present study was whether the RES model would prove as effective and whether this model would benefit students in high-poverty schools from families whose first language is not English. This population has been neglected in studies of Reading Recovery, which have included only a small number of language-minority students (Cox & Hopkins, 2006; Schwartz, 2005).

LINNEA C. EHRI is a Distinguished Professor of Educational Psychology in the PhD program in Educational Psychology, Graduate Center of the City University of New York, 365 Fifth Ave., New York, NY 10016-4309; e-mail: *lehri@gc.cuny.edu*. Her research interests involve reading and spelling processes, their course of development, and how to facilitate their acquisition in typically and atypically developing readers.

LOIS G. DREYER is a Professor Emerita of Southern Connecticut State University and a Research Scientist at the Center for Advanced Study in Education, Graduate Center of the City University of New York, 365 Fifth Ave., New York, NY 10016-4309; e-mail: *lgdreyer@optonline.net*. Her research focuses on beginning reading and spelling, the instructional needs of students experiencing difficulty, and reading assessment.

BERT FLUGMAN is the Director of the Center for Advanced Study in Education, Graduate Center of the City University of New York, 365 Fifth Ave., New York, NY 10016-4309; e-mail: *bflugman@gc.cuny.edu*. His primary research interest involves evaluating the effects of educational interventions.

ALAN GROSS is a Professor of Educational Psychology in the PhD program in Educational Psychology, Graduate Center of the City University of New York, 365 Fifth Ave., New York, NY 10016-4309; e-mail: *agross@gc.cuny.edu*. His research interests include applied statistics, educational measurement, and techniques for solving problems of missing data.

A reading intervention model is more encompassing than a reading program, which is focused on curriculum materials and instructional procedures. A model includes additional features specifying the qualifications and training of instructors; the students who are targeted and criteria for selecting them based on the objectives of instruction; observational tools, and tests to monitor students' progress and make instructional decisions; and the costs to schools.

The Reading Rescue model was implemented with the following features in the present study. The instructors were personnel already employed by schools, including paraprofessionals, reading specialists, and credentialed teachers. They received on-site training, supervision, and coaching. The students were low-SES, language-minority first graders reading near the bottom of their class as indicated by low scores on literacy tests. The instruction involved one-to-one tutoring. It was comprehensive in covering all five components identified as evidence-based and essential by the National Reading Panel (2000) (i.e., phonemic awareness, systematic phonics, fluency, vocabulary, and reading comprehension). Tutors kept records of students' performance, and they tested students to assess readiness for higher level lessons and texts. The costs to schools were relatively modest because existing school staff members rather than outside, specially trained teachers were used as tutors.

We expected several features of Reading Rescue instruction to contribute to its effectiveness, as shown by previous research on these features. Activities drawn from the Reading Recovery model were sentence writing and the rereading of little books (Clay, 1993b). The inclusion of systematic, sequential phonics instruction was expected to enhance its impact as shown by the National Reading Panel (Ehri, Nunes, Stahl, & Willows, 2001). Students were taught phonemic awareness, decoding, sight word reading, and spelling. They read little books (i.e., 8 to 24 pages) that were leveled (i.e., classified by number to indicate gradients of difficulty). The books contained both high-frequency and decodable words that were tied to a scope and sequence chart used for letter–sound instruction. These texts were expected to help students apply the letter–sound correspondences and decoding skills they had been taught (Juel & Roper-Schneider, 1985; Torgesen et al., 1999). The inclusion of vocabulary instruction was expected to make the program especially appropriate for low-SES, language-minority students (Biemiller, 1999). Questioning students in various ways about the books they read and distinguishing between text-explicit and text-implicit questions were expected to foster reading comprehension (Block & Pressley, 2002).

There are many reasons why tutoring in first grade has been especially attractive to educators as a means of preventing early reading failure (Clay, 1985). Instruction can be tailored to the difficulties of individual students. Tutors can easily monitor students' progress and the need for review. Compared to group instruction, tutoring allows more time for students to read, and their errors and misconceptions receive immediate feedback. Also, having the tutor's undivided attention is motivating. In a meta-analysis to assess the effectiveness of tutoring programs in reading for students at risk, Elbaum et al. (2000) reported a mean effect size of 0.41.

Although one-to-one tutoring has many advantages, not all studies have found that tutoring is more effective than small-group instruction. In a meta-analysis evaluating experiments on phonemic awareness instruction, Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh, and Shanahan (2001) found significantly smaller effect sizes for tutoring (d = 0.60) than for small groups (d = 1.38). In a meta-analysis of systematic phonics instruction, Ehri, Nunes, Stahl, and Willows (2001) found that the effect size for individual tutoring (d = 0.57) did not differ significantly from that for small-group instruction (d = 0.43). However, these comparisons were extracted from different studies. In their meta-analysis of tutoring studies with struggling readers, Elbaum et al. (2000) cite two unpublished doctoral dissertations that directly compared tutoring to small groups in the same study and found no differences, with effect sizes of –0.12 and 0.05. One purpose of the present study was to compare the effects of tutoring to the effects of small-group instruction.

The backgrounds of adult tutors have varied across studies. These have included teachers, paraprofessionals, college students, parents, and community volunteers (Elbaum et al.. 2000). None of these studies, however, has compared different types of tutors in the same study teaching the same program, so it is unclear whether they are equally effective. It is an issue among educators whether tutors without elementary education certification are as successful as credentialed teachers in providing adequate reading instruction. Organizations such as the International Reading Association regard credentialing and professional development as vital for adequate preparation in the teaching of reading (International Reading Association, 2006). This group's position implies that adults without this training would not be as effective. However, when the complexities of classroom management and decision making required of teachers are reduced, specifically, when non-credentialed adults provided with a well-structured reading program are taught how to use it to tutor individual students and are supervised, they may prove effective. The use of such tutors is especially valuable in schools serving low-SES populations. Such schools may have larger numbers of children reading below grade level, and their limited budgets may preclude hiring enough certified reading specialists to work with all the students who are needy. One purpose of the present study was to determine whether paraprofessionals who are trained and supervised can tutor individual students in reading as effectively as credentialed teachers and reading specialists.

Findings regarding the effectiveness of paraprofessionals in teaching reading are mixed. Wasik and Slavin (1993) reviewed studies that provided tutoring for struggling readers in first grade. Although studies utilizing certified teachers tended to show larger effects than those using paraprofessionals, the paraprofessionals and certified teachers did not teach the same programs. Moreover, the programs used by paraprofessionals were more focused and less comprehensive in the reading components taught than the programs used by teachers. Thus, these findings are not definitive.

In a recent study, Brown, Morris, and Fields (2005) compared the effectiveness of certified teachers and paraprofessionals. Both used the same

reading program, Next Steps (Morris, 1999), to tutor second- and third-grade struggling readers. The program included guided oral reading of instructional level text, word study, fluency training through rereadings of text, and listening comprehension. Results revealed that means favored students tutored by teachers to students tutored by paraprofessionals. Effect sizes were 0.36 (word recognition), 0.43 (oral passage reading), 0.27 (reading comprehension), and 0.79 (pseudoword reading). However, differences were statistically significant only on the pseudoword reading measure.

Although Brown et al.'s (2005) data indicate that paraprofessionals are almost as effective as teachers with older students, it is not clear whether the same results would hold for first graders. In the Wasik and Slavin (1993) review, paraprofessionals were less effective in tutoring at-risk first-grade readers than credentialed teachers. Teachers would be expected to possess a greater knowledge base that gives them an advantage over paraprofessionals in teaching beginning-level reading, which may be more difficult. One limitation of the Brown et al. study was that paraprofessionals were randomly assigned to students whereas classroom teachers tutored students from their own classes. The present study used complete random assignment to compare trained paraprofessionals to credentialed teachers and reading specialists. Because of the additional education and expertise of reading specialists, we expected them to be especially skilled in tutoring students. Apart from Brown et al. and the present study, there have been few, if any, studies that have directly compared the effectiveness of different types of tutors.

The Reading Rescue intervention model has been shown to be effective in previous conference presentations and unpublished reports. Studies by N. Hoover (1994, 1995), N. Hoover and Lane (2001), N. Hoover and Sullivan (1996), and Britt (2002) were field evaluations rather than controlled experiments. Results indicated that RES tutoring increased students' reading from pretest to posttest and raised scores from below-level to grade-level performance that was equivalent to average readers in their schools.

Muller and Davies (M&D; 2004) conducted a more controlled experiment with language-minority first graders who were urban, low-SES, struggling readers. Students were randomly assigned to receive either RES tutoring or regular classroom instruction. They found that tutoring produced higher reading scores on the Gates-MacGinitie Reading Tests (4th ed.; GMRT4; MacGinitie, MacGinitie, Maria, & Dreyer, 2000), with an effect size of 0.32. Their study was less extensive than ours in several respects: nature of control groups (i.e., use of one no-treatment control group drawn only from schools using RES by M&D vs. use of multiple control groups that received either an alternative treatment or no treatment in schools using as well as not using RES), standardized reading tests to measure outcomes (i.e., use of one group-administered test, the GMRT4, by M&D vs. use of this test plus two individually administered subtests of the Woodcock Reading Mastery Tests–Revised [WRMT-R; Woodcock, 1987, 1998]); and schools' experience with RES (i.e., half of the schools used by M&D were in the 1st year of RES implementation whereas none of our schools was this inexperienced). In

addition, we compared the reading performance of RES-tutored students to all first graders across eight schools, we examined the effectiveness of different types of tutors, and we analyzed tutoring events and their relationship to outcomes, whereas M&D did not do this. Results of the present study were intended to replicate and extend findings of these unpublished studies.

The participants in the present study were *language-minority students*, a term advocated by the National Literacy Panel on Language Minority Children and Youth (August & Shanahan, 2006). The term refers to students from homes where a language other than the predominant societal language is actively used. The students may be of limited second-language proficiency, bilingual, or essentially monolingual in the second language and proficient enough in the second language to be able to profit from classroom instruction conducted in that language. There is a dearth of controlled studies examining the effectiveness of literacy instruction with these students, according to the panel's report. Shanahan and Beck (2006) located six studies that provided instruction in multiple literacy elements to language-minority first graders. Effect sizes ranged from 0.20 to 1.15 ($M = 0.55$), indicating generally positive and moderate effects.

Several questions were addressed in the present study:

1. Will struggling first-grade language-minority readers who receive RES tutoring make greater gains in their reading achievement than struggling readers who do not receive tutoring but are enrolled in the same schools, and greater gains than struggling readers who are enrolled in comparable schools not implementing the RES intervention? Both types of control groups were included to take account of contamination effects on reading instruction possibly resulting from the presence of the RES program in the schools.
2. Will struggling language-minority readers who receive RES tutoring make greater gains in reading than struggling readers who receive a small-group intervention program adopted by the district?
3. Will RES tutoring increase students' reading achievement to average levels of performance based on nationally normed tests?
4. Will entry-level abilities of students and features of RES tutoring predict how much first graders improve in their reading as a result of tutoring?
5. Will all types of tutors, including reading specialists, credentialed teachers, and trained paraprofessionals, be equally effective at improving the reading achievement of their students?

Method

Participants

Participants were primarily language-minority students who were enrolled in first grade during the 2003-2004 school year. They came from five low-SES urban public schools that had implemented the RES tutoring intervention for 2 to 3 years and three comparable schools that had not implemented RES but hoped to do so in the future. All schools were located in the same district in a large metropolitan city. The majority of students, more than 90%, were from homes where Spanish was the first language, and 95% qualified for free or

reduced-price lunch. Students judged by teachers to have virtually no knowledge of English and students receiving English language learner services were excluded from the study. The remaining students were considered to have sufficient English to qualify, on the basis of teacher judgments. Students in special education at the beginning of first grade (i.e., those with speech, language, or physical disabilities or behavioral problems) were excluded by schools from the RES program because they were already receiving special services.

First graders in the five RES schools (N = 497) were group administered the RES Classwide Screening Assessments to identify lower performing students (N = 203) who were then tested individually on the RES pretests. Students who could name at least 17 letters but were unable to read a preprimer passage on an informal reading inventory formed the pool of students who qualified for the intervention (N = 190). Within each school, these candidates were rank ordered on the RES individual pretests, and adjacent scores were used to form 64 matched pairs. Between the time that pairs were formed and the intervention began, four of the five schools assigned some of these students to a small-group intervention mandated by the district, *Voyager Passport* (Voyager Expanded Learning, 2004). In these cases, the other member of the pair was assigned to receive RES. For the remaining pairs, members were randomly assigned to the RES and control groups. During the year, a few students left school (n = 8 RES and 9 controls). Replacements (i.e., 8 RES and 7 controls) were students drawn from the larger pool with the most similar pretest rankings.

A word of explanation is needed regarding selection criteria. Students are able to proceed more rapidly through the instructional phases of the RES program if they know most letters. Low-performing students who do not know enough letters in the fall are typically enrolled in the program later in the year after they have been taught letters in their classrooms. This step makes tutoring more efficient. However, because tutoring began late in the fall in the present study, it was not possible to tutor a second wave of students.

To address the possibility that the presence of RES in schools would affect the reading instruction in those schools, a second control group of first graders was drawn from three comparable schools not implementing the RES intervention. This group consisted of struggling readers who had similar scores to students in the RES group on the fall GMRT4 pretest and who completed both fall and spring GMRT4 pretests and posttests.

There were 64 students who received the RES intervention, 62 control students (C1) enrolled in schools offering the RES program, and 60 control students (C2) enrolled in non-RES schools. Students in the two control groups were reconfigured to distinguish those who received the district-mandated small-group intervention (n = 52) from those who did not receive this intervention (n = 70).

Tutoring was provided by 59 adults who were members of the school staff and had received formal RES training. The tutors included certified reading specialists with graduate degrees (n = 17), adults certified in other areas (guidance counseling, math, and social work; n = 15), and paraprofessionals (n = 27).

Those certified in counseling and math had completed at least 6 credits of coursework in reading and language arts. All paraprofessionals were high school graduates, many had bachelor's degrees or were working toward BAs and teacher certification, and a few had MAs in other fields. Half of the tutors were bilingual. All tutors had additional responsibilities within the school. For many, this included teaching the small-group intervention program. Tutors were randomly assigned to tutees within each school and remained with those students throughout the intervention. All but three tutors worked with only one RES student.

Materials and Procedures

The pretests, posttests, and tutoring intervention were administered by school personnel. Researchers had contact with RES trainers and with school coordinators who were liaisons to the study but had no direct contact with other tutors or their students.

Pretests. Group-administered and individually administered pretests were given to assess entry-level reading ability and to identify potential candidates for tutoring. Group tests were given in all eight schools. Individual tests were limited to the five RES schools because staff members in non-RES schools lacked the training to give the individual tests.

1. GMRT4. The GMRT4 Level BR (Beginning Reading; MacGinitie et al., 2000; MacGinitie, MacGinitie, Maria, & Dreyer, 2002) was administered to classrooms of first graders in all eight schools in November. This multiple-choice test measured students' knowledge of letter–sound correspondences and their ability to read high-frequency words. Teachers read aloud each question and answer choice and gave students sufficient time to respond. Raw scores were transformed to normal curve equivalent (NCE) scores based on national norms. NCEs are percentiles that have been transformed into a scale of equal units and are therefore appropriate for computing averages. The Kuder-Richardson 20 (K-R 20) reliability is reported to be 0.95. Scores were used to match students from the non-RES control schools to students in the RES group. Testing occurred before the interventions began.
2. Iowa Tests of Basic Skills (ITBS; H. Hoover et al., 2001, 2003) vocabulary test. The ITBS Level 6 vocabulary subtest was administered to classrooms of students in November. The teacher read the questions and answer choices aloud, and students recorded their answers. The K-R 20 reliability is reported to be 0.75.
3. RES Classwide Screening Assessments. The RES screen consisted of four group-administered tests adapted from Clay (1993a, 2002) that were given by classroom teachers in mid-September to students in RES schools in a print-free testing environment. Scores were used to identify low-performing candidates for the study. The tests assessed alphabetic knowledge (i.e., students wrote all the lower- and uppercase letters they knew), word writing (i.e., students had 10 minutes to write all the words they knew how to spell), and developmental spelling (i.e., 10 words were dictated and students wrote letters for the sounds they heard).

4. RES Individual Pretests. Tutors administered seven RES pretests individually to confirm students' candidacy for tutoring in early November. Students (a) named lowercase letters, (b) named uppercase letters, (c) gave sounds of the letters or examples of words beginning with the letter sounds, (d) blended spoken sounds to form words, (e) segmented words into phonemes, (f) read aloud nonwords, and (g) read sight words from graded lists of the informal reading inventory. Scores in these tasks were summed to yield an overall score. Students needed to know at least 17 letter names to qualify for RES tutoring.
5. Informal reading inventory. The Ekwall/Shanker Reading Inventory (4th ed.; Shanker & Ekwall, 2003) was administered in early November. Scores on the sight word task (see above) determined the proper starting point for reading the graded passages. Each passage was followed by comprehension questions. Two grade-level scores were calculated: word reading accuracy and comprehension. Students began with the preprimer passage and continued until their word reading accuracy dropped below 91% or their comprehension below 60%. Any student who read the primer or Grade 1 passage at a minimum of 95% accuracy and at least 60% comprehension was disqualified.

The tutoring intervention. The program was not guided by a teacher's manual with predetermined content specifying each day's lesson. Rather, tutors made instructional decisions that were guided by the RES scope and sequence of skills as well as by their analyses of students' performance on assessments, their written record of students' text reading, and their observations of students' response to instructional activities. Tutors were trained to teach explicitly and systematically, rather than incidentally, while being responsive to individual needs.

The intervention began in December, occurred during school hours, and consisted of "easing-in" sessions followed by regular instructional sessions and an "easing-out" period. During easing-in sessions, tutors administered the individual assessments and established rapport with students through such activities as drawing pictures and reading easy books together. Regular instruction consisted of a sequence of lessons organized into phases that gradually presented more challenging concepts and material as students made progress. Lessons taught fluency, phonological awareness and phonics, comprehension, and vocabulary development, with the amount of time for each dependent on the phase of the program.

To apply the strategies and skills acquired, students read fiction and nonfiction small, illustrated books from the *Ready Readers* series (Englebretson, Hiebert, & Juel, 2000). These books were used but without the supplementary materials accompanying the books. The books were organized by difficulty, and they systematically introduced and provided practice with high-frequency sight words as well as sound-symbol relationships in decodable words. The books provided practice in the phonics instruction that students received, and they were read and reread to build fluency. During the easing-out sessions, a final review of the application of learned skills and strategies was conducted and students' skills were assessed.

The following components were taught in this order during each session:

1. Fluency. Students started each lesson by rereading one or more familiar books. Beginning in Phase 1, oral reading accuracy, sight word knowledge, and confidence as a reader were stressed. In Phase 2, students practiced the application of blending, segmenting, and syllabication skills. Timed readings of familiar text were used to increase reading rate and reduce errors. In Phase 3 and beyond, reading speed and expression were emphasized.
2. Word analysis and comprehension strategies. Students reread orally a book that had been introduced to them during the previous lesson (see Component 5, below). A written record was taken indicating the student's word reading accuracy and use of word- and text-reading strategies. This served as a basis for deciding when to move to a higher level of difficulty. In selecting texts, tutors were advised to avoid books that were too easy or too hard and to maintain reading accuracy levels between 90% and 97%. In deciding which features of words to correct, tutors consulted a phonics elements chart. Errors on letter–sound relations that were less regular or were too advanced were not corrected.
3. Phonological awareness and word study. Multisensory approaches were used. A phonics elements chart guided instruction and was supported by text-reading practice in *Ready Readers*. The phonics chart grouped instruction in letter–sound correspondences by phases. Phase 1 covered basic consonants, short vowels, and consonant-*H* digraphs. Phase 2 covered the short–long vowel distinction, additional consonants, and *-er* and *-ed* endings. Phase 3 covered contractions, *r*-controlled vowels, additional vowel patterns, hard and soft *g* and *c*, and compound words. In addition, 18 to 23 phonograms were taught in each phase. Phonics was taught systematically and sequentially using both analytic and synthetic approaches. For example, students saw and heard a phonetically regular word; they repeated the word slowly, saying its separate sounds while pointing to its letters; and then they spelled the word by moving plastic letters to segment and represent the sounds. Additional activities were employed for students having difficulty, such as tracing letters while saying their sounds and using a mirror to monitor articulatory movements in words.

   Phonemic awareness, the ability to manipulate the smallest sounds in speech, was taught systematically as part of phonics instruction. During Phase 1, Elkonin (1973) boxes were used to teach phoneme segmentation. Students progressed from two-phoneme words to four-phoneme words, each illustrated by pictures so that pronunciations were recognized. Segmenting and blending were taught using fingers and magnetic letters. Students kept a personal book, the *Letters and Words I Know Book*. The names and sounds associated with unfamiliar letters were taught and practiced. In Phase 2, students were taught to decode by analogy using keywords. They focused on high-frequency words containing common phonograms. In late Phase 2 and beyond, tutors taught syllabication and other ways of dividing words to decode them, for example, common content words like *carrot*, *morning*, *talking*, and *bathtub*.
4. Writing to develop phonological awareness, phonics, and comprehension. In Phases 1 through 3, tutors assisted students in writing one or more sentences that the students had produced in conversation about the books they read. Attention was directed at spelling words correctly as well as comprehending the texts. When students were deemed ready in Phase 3, tutors taught students about text structure and how to organize information in texts they had read, for example, through the creation of semantic maps for expository text.

5. Comprehension and vocabulary development with the new book. In the initial phase, tutors introduced a new book by talking through it with students, page by page. As the books became longer, introductions focused more on making predictions based on the title and cover page. One purpose was to develop oral language by encouraging students to talk about the books and by explaining the meanings of new vocabulary words. These words were written in students' personal books, and the meanings were reviewed each time the book was read. Another goal was to spark the student's interest in the book. The tutor coached the student through the reading of the book and prompted the application of the reading skills and strategies that had been taught. Students were encouraged to decode unknown words by relying on their letter–sound knowledge and then cross-checking with meaning and pictures to confirm the identities of the words. Students were asked literal, inferential, and evaluative questions after the text was read. In Phase 3, students were taught the difference between text-explicit and text-implicit questions.

Graduation from RES. The RES program specifies multiple criteria for determining when students have met the goals of the program. Data on graduation criteria and decisions were not collected in the present study. Rather, the reading achievement of all the students who received tutoring was assessed at the end of first grade.

Training of tutors and coordinators. The tutors were taught how to provide the intervention by RES program staff. Training was guided by the *Reading Rescue Tutor Handbook* (N. Hoover, 2001). The initial phase consisted of 5 days of workshops delivered at the schools. The first 2 days at the beginning of the school year developed tutors' understanding of the reading process; the components of the intervention model; the purpose, administration, and scoring of the classwide screening assessments; and how to begin working with students. The next 3 days of training took place after tutoring had begun and focused on the lessons.

A school coordinator who was a member of the school's staff served as an on-site resource whose role was to provide modeling, feedback, and support to tutors; to help tutors form a peer coaching team; and to monitor implementation of the program. The coordinator had other literacy-related responsibilities within the school but did not have classroom teaching duties. Coordinators were in their 2nd or 3rd year of implementation. They had attended advanced training institutes. They were given all the materials necessary, including videotaped lessons, to continue the training of tutors in their school. They were expected to schedule regular staff development meetings. RES program staff returned to each school several times to provide continued training and technical assistance for tutors and coordinators and for quality assurance visits.

Assessment of tutoring. Tutors completed lesson record sheets as they tutored during each session. The sheets reminded tutors of assessments, lesson components, activities, and reading strategies to be taught and mastered in each phase. Tutors listed dates, materials used to teach components, and how students responded.

To assess characteristics of tutoring and tutors' adherence to procedures, two professionals thoroughly familiar with the program scored the records kept by tutors, whose names had been removed. The scorers analyzed various characteristics, including number of easing-in, easing-out, and regular tutoring sessions; the number of readings or rereadings of books; and reading accuracy levels, which consisted of the proportions of readings that were completed at a 98% to 100% accuracy level, at a 90% to 97% accuracy level, and at an accuracy level below 90%. These reading measures were taken from students' second exposure to the books when they read them independently while tutors kept a running record. The first exposure to the books had occurred during the previous session, when tutors coached students through the book (see Components 2 and 5 of the intervention, above).

To provide an index of tutors' adherence to the intervention model, judges rated the following six dimensions of the tutoring records on a 4-point scale to indicate whether expectations were met: number of books read, difficulty of the books, selection of elements for word study, features of the sentences written, instructional decisions made during sentence writing, and provision of tutoring on a regular basis on the days available. Scores were summed to provide an adherence measure, with a maximum score of 24. To illustrate, the criteria used to rate tutors' selection of sentences were as follows:

Rating of 4: Sentences chosen by the tutor for the student to write conformed to the recommended criteria with regard to their length, complexity, and lexical features.
Rating of 3: With few exceptions, sentences chosen by the tutor conformed to the recommended criteria.
Rating of 2: Many sentences chosen by the tutor conformed to the recommended criteria.
Rating of 1: For the majority of lessons, sentences chosen by the tutor did not reflect program recommendations, or many sentences were not recorded.

The two judges rated all of the records independently and were in agreement on 90% of their ratings. Disagreements were discussed and resolved. Nonconformance to the program corresponded to a rating of 1 on five dimensions and 2 on one dimension (total score of 7). Conformance to the program corresponded to ratings of 2 or higher on five dimensions and 3 or higher on one dimension, for a minimum adherence score of 13.

Small-group intervention. Some of the struggling readers in the control groups (C1 and C2) across the eight schools received instruction in a small-group intervention program, *Voyager Passport* (Voyager Expanded Learning, 2004). The small-group program was adopted by the school district as a supplemental intervention for first-grade struggling readers with the goal of bringing their reading up to grade level. This was the 1st year of implementation, in contrast to RES, in its 2nd or 3rd year. The program provided intensive, explicit, scripted instruction consisting of 130 lessons scheduled to be

taught daily for 26 weeks. Tutor training in the program was provided by the district and consisted of one session plus on-site assistance by a literacy coach. Because instruction was strictly controlled by a script, more extensive training was not regarded as necessary.

Sessions were conducted in groups (3 to 6 students) and lasted for 30 to 40 minutes. The program taught the same components as the RES program. These were phonemic awareness, phonics, daily story reading, vocabulary instruction that included multiple exposures to new words, reading and listening comprehension with both narrative and expository passages, and fluency. Phonics instruction was emphasized and included teaching letters and their sounds, decoding regular words, reading high-frequency irregular words, teaching various word types, word-building exercises, and spelling. Fluency included practice in speeded letter naming and word reading as well as text reading in a limited number of decodable books (not the same books used in RES). The small-group program included a vital-indicators-of-progress guide for assessment, which was scheduled to occur every fifth lesson. In the small-group program, parents were expected to provide home support with take-home readers. (Parents were not involved in the RES program.) Once students entered the small-group program, they remained for the duration of the school year. The small-group intervention was taught by many of the same staff members who also tutored in the RES program in the five RES schools. In the non-RES schools, it was taught by teachers and paraprofessionals who were not trained as RES tutors.

Classroom reading instruction. The first-grade curricula in reading and writing were uniform throughout the city. The instructional programs used were *Month-by-Month Phonics for First Grade* (Cunningham & Hall, 1997) and a readers and writers workshop. These programs did not follow a scope and sequence. Phonics instruction was taught opportunistically rather than systematically. Word-identification instruction involved teaching students to read new words by analogy to a set of keywords. The presence of a universal reading curriculum across the eight schools strengthened the design of the present study by controlling for the type of regular classroom reading instruction received by all the participants.

Posttests. Students in the RES and C1 groups completed several individually administered posttests in May: two WRMT-R subtests, given by trained reading specialists, and the RES posttests, given by tutors. Classroom teachers administered the GMRT4 to first graders in all eight schools and the RES classwide screen to first graders in the RES schools.

1. WRMT-R Word Identification Test. Form G of the WRMT-R required students to read a list of real words that increased in difficulty. The median split-half reliability of this measure is reported to be 0.98 (Woodcock, 1987, 1998).
2. WRMT-R Word Attack Test. Students read a list of decodable nonwords. The median split-half reliability of this measure is reported to be 0.94 (Woodcock, 1987, 1998).

3. RES individual posttests. The seven subtests of the RES individual pretests were repeated as posttests along with the Ekwall/Shanker Reading Inventory (4th ed.; Shanker & Ekwall, 2003). (See descriptions above.) The maximum score on the RES posttests was higher, 223, because the sight word test included 110 words. The highest passage given to any student in the posttest phase was the third-grade passage.
4. GMRT4. The GMRT4, Level 1 (MacGinitie et al., 2000, 2002), was group administered by classroom teachers. Questions were read and answered by students independently. The test consisted of two subtests whose scores were treated separately as well as combined into a total score. The K-R 20 reliability of the total score is reported as 0.96.
   a. Word Decoding. There were 43 items. On each, students identified which of four words corresponded to a picture. The foils resembled the correct words in appearance or sound. Knowledge of letter–sound correspondences and letter sequences was required to make correct choices. K-R 20 reliability of this subtest is reported as 0.94 for the spring of Grade 1.
   b. Reading Comprehension. Students read passages covering both fiction and nonfiction and answered each comprehension question by choosing from a panel of pictures the one that illustrated or answered a question about a segment of the text. K-R 20 reliability of this subtest is reported as 0.93 for spring of Grade 1.

Design and Analyses

Performance of the group receiving RES tutoring was compared to performance of two different configurations of the same control students. The first configuration divided the students into a control group drawn from the same five schools as the RES students, referred to as C1, and a control group drawn from the three schools that did not offer RES, C2. The purpose of including C2 was to check the possibility that the presence of RES and the professional development it offered to schools might affect instruction provided to control students in those schools. The second configuration divided the students into those receiving the small-group intervention (SG) and those who did not receive this intervention (NI). Students in SG and NI groups came from all eight schools.

ANOVAs were applied to fall pretest scores to determine whether the groups differed in entry-level reading skills. To assess effects of the RES and small-group interventions, ANCOVAs were applied to spring posttest scores with GMRT4 pretest scores as the covariate. The dependent variables included word-decoding and reading-comprehension scores on the GMRT4, word-reading and nonword-reading scores on the WRMT-R, and performance on the RES program tests. In some analyses, data were missing for a few students. Statistical tests were conducted on actual scores.

To compare the three types of tutors—reading specialists, credentialed staff, and paraprofessionals—the performance and tutoring experiences of their students were subjected to ANOVAs. To determine whether features of the tutoring instruction predicted students' reading achievement at the end of first grade, regression analyses were conducted.

Results

Equivalence of RES Treatment and Control Groups on the Pretests

Table 1 shows the mean pretest performance of students who received RES tutoring and control students. From mean percentiles on the fall GMRT4 reading test and ITBS vocabulary test, it is apparent that participants were struggling readers at risk for future reading difficulties and had English-language vocabularies well below average compared to national norms for first graders.

The ANOVAs of pretest scores with treatment group as the independent variable revealed that the three groups were similar in age and gender. They did not differ on the vocabulary test, but they did differ on the GMRT4 test. A post hoc Bonferroni test indicated that the RES and C2 groups performed equivalently and significantly better than the C1 group. This was unexpected because RES and C1 students had been matched on their performance on the RES individual tests. One possible reason for poorer performance of the C1 group is that four of five RES schools partially undermined random assignment by selecting some of our participants for the small-group intervention.

To check on whether C1 students were poorer readers than RES students on the other literacy pretests, mean scores were compared. ANOVAs revealed that the groups did not differ on any of the three group-administered RES class-wide screen tests (all ps $> .05$) or on any of the seven individually administered RES pretests (all ps $> .05$). Correlations between the RES pretests and the GMRT4 were as high as $r = .69$ (developmental spelling) and $r = .67$ (sight word reading), indicating that the RES tests were not lacking in sensitivity. These findings raise doubt that the RES and C1 groups actually differed in their reading level in the fall.

Nevertheless, to take account of the possibility of initial differences in ability favoring the RES and C2 groups over the C1 group, fall scores on the GMRT4 were used as a covariate in ANCOVAs of spring posttest scores. This analysis and choice of a covariate to adjust for initial differences in reading ability were considered legitimate for several reasons: The GMRT4 test has high reliability, it was administered to all participants at the same time of year before tutoring began, and it was strongly correlated with the next level in the GMRT4 test series given as a posttest ($r = .68$, $p < .001$, $N = 844$).

Use of ANCOVA requires that the slopes of the regression lines predicting posttest scores from pretest scores in each of the treatment and control groups being compared are shown to be homogeneous. We tested the assumption of homogeneity of slope in each of the ANCOVAs reported by using the following procedure. We first set up a regression model (posttest score was the dependent variable) that contained the following predictor variables: dummy variables coding group membership, the pretest, and the cross-products of the dummy variables with the pretest. We then tested the significance of the regression weights for the interaction terms. None of the F tests was significant at $p < .05$. Thus, these findings support the appropriateness of applying ANCOVAs to our data.

Table 1

Characteristics, Mean Performance, and Standard Deviations on Fall Pretests and Spring Posttests for Reading Rescue (RES) Students, Control Students in RES Schools (C1), and Control Students in Non-RES Schools (C2)

Characteristic or Test	RES (n = 64)	C1 (n = 62)	C2 (n = 60)	Post Hoc Comparison (Bonferroni)	F Stat[a]	Number of Students[b] (RES/C1/C2)
Age (months)	75.3 (3.9)	74.3 (5.4)	74.2 (3.6)	—	1.26 n.s.	64/62/59
Gender (female/male)	30/34	30/32	30/28	—	< 1 n.s.	64/62/58
Fall pretests[c]						
GMRT4 NCE	29.2 (9.2)	25.3 (9.1)	32.5 (8.5)	RES = C2 > C1	9.91**	63/61/60
Percentile	16	12	20			
ITBS vocabulary SS	125.3 (8.5)	125.6 (8.5)	128.3 (9.0)	—	2.18 n.s.	64/62/59
Percentile	21	22	27			
Spring posttests[c]						
GMRT4 word decoding NCE	49.9 (15.9)	42.1(13.5)	35.2 (11.8)	RES > C1 > C2	19.41**	62/60/60
Percentile	50	36	25			
GMRT4 reading comprehension NCE	43.5 (12.9)	37.8 (13.8)	35.3 (11.2)	RES > C1 = C2	7.50**	62/60/60
Percentile	38	29	25			
GMRT4 total NCE	47.5 (12.9)	41.2 (12.1)	36.4 (10.4)	RES > C1 = C2	15.47**	62/60/60
Percentile	45	33	28			
WRMT-R word reading	35.1 (9.3)	27.5 (11.8)	—	RES > C1	19.14**	62/61/NA
Grade equivalent	1.9	1.7				
WRMT-R word attack	14.7 (9.0)	9.4 (7.4)	—	RES > C1	12.93**	62/61/NA
Grade equivalent	1.9	1.5				

Note. GMRT4 = Gates-MacGinitie Reading Tests (4th ed.; MacGinitie, MacGinitie, Maria, & Dreyer, 2000); ITBS = Iowa Tests of Basic Skills (H. Hoover et al., 2001, 2003); SS = standard score; NCE = normal curve equivalent; WRMT-R = Woodcock Reading Mastery Tests–Revised (Woodcock, 1987, 1998).

a. ANOVAs were conducted on fall pretest scores. ANCOVAs were conducted on posttest scores with fall GMRT4 scores as the covariate. Posttest means are adjusted values, and posttest standard deviations (shown in parentheses) are unadjusted.

b. Missing data account for the varying number of students contributing to each analysis.

c. In the fall, Level BR (Grade 1 Beginning Reading) of the GMRT4 was administered. In the spring, Level 1 (Grade 1) of the GMRT4 was administered. All tests were group administered except the WRMT-R tests, which were individually administered.

*p < .05. **p < .01. n.s. = not statistically significant.

Comparison of RES Treatment to C1 and C2 Control Groups on Posttests

Results of the ANCOVAs applied to posttest measures revealed significant main effects of the treatment group. Table 1 reports means and test statistics on the nationally normed tests. Bonferroni post hoc comparisons on GMRT4 performance revealed that RES-tutored students decoded significantly more words and comprehended text significantly better than both of the control groups. In turn, the C1 group, located in RES schools, decoded significantly more words than the C2 group, located in non-RES schools, but comprehension of the two groups did not differ. This finding of a difference suggests that the presence of the RES program in schools may have improved the decoding instruction that struggling readers in these schools received.

On the WRMT-R tests that were individually administered, the RES-tutored group read significantly more words and decoded significantly more nonwords than the C1 group. Mean grade-equivalent scores on the WRMT-R tests in Table 1 show that RES students were equally skilled at reading words and nonwords, an indication of good phonics skills. In contrast, C1 students were weaker in reading nonwords than real words.

Performance of the RES-tutored and C1 control groups on posttests from the RES batteries are shown in Table 2. ANCOVAs were conducted with fall GMRT4 scores as the covariate. The groups did not differ on four of the five tests of letter knowledge where means were close to ceiling, indicating that students in both groups knew most of the letter names and sounds. However, significant main effects were observed on other posttests. RES-tutored students outperformed control students in writing lowercase letters, segmenting and blending phonemes, reading sight words and pseudowords, generating plausible spellings of words, and reading and comprehending text. These findings provide additional evidence that struggling readers who received RES tutoring read significantly better at the end of first grade than struggling readers who did not receive this program.

Effect sizes were calculated by dividing the difference between the RES and C1 adjusted means by the pooled standard deviation on each measure. Results on the group-administered GMRT4 were 0.53 (decoding), 0.43 (comprehension), and 0.50 (total score). Results on the individually administered WMRT-R were 0.72 (word reading) and 0.65 (word attack). Results on the RES individual tests are given in Table 2. On tasks that assessed word and text reading, effect sizes ranged from 0.61 to 1.09. According to Cohen's (1988) rule of thumb, our effects sizes range from moderate to large. Moreover, they are equal to or higher than effect sizes in other reading instruction studies. In the evaluation of RES tutoring with first graders, M&D (2004) reported an effect size of 0.32 on the total GMRT4 test. In its meta-analysis of systematic phonics instruction, the National Reading Panel reported an average effect size of 0.57 in studies that involved tutoring and an average effect size of 0.74 in studies involving first graders at risk (Ehri et al., 2001). In Elbaum et al.'s (2000) meta-analysis of tutoring studies with at risk readers, the mean effect size was 0.41.

Table 2

Posttest Means, Standard Deviations, and Test Statistics on the Group-Administered Reading Rescue (RES) Classwide Screen and the Individual RES Tests for the RES Group and the C1 Control Group

Test	RES (n = 64)	C1 (n = 62)	F Stat[a]	Number of Students[b] (RES/C1)	Effect Size[c]
Classwide screen tests					
Write uppercase letters (26 max)	22.8 (4.0)	21.8 (5.8)	1.08 n.s.	62/60	0.20
Write lowercase letters (26 max)	23.3 (3.6)	20.9 (7.3)	5.48*	62/60	0.44
Write words	26.2 (14.9)	23.3 (13.3)	1.34 n.s.	62/60	0.21
Developmental spelling (40 max)	27.6 (5.8)	24.6 (7.9)	5.66*	62/60	0.44
RES individual tests					
Name lowercase letters (26 max)	25.7 (.69)	25.4 (1.0)	2.42 n.s.	61/59	0.36
Name uppercase letters (26 max)	25.8 (.50)	25.9 (.72)	< 1 n.s.	61/59	–0.16
Give sounds of letters (26 max)	25.5 (.92)	25.1 (1.6)	2.72 n.s.	61/59	0.32
Phoneme blending (10 max)	9.72 (.54)	9.34 (1.2)	5.22*	61/59	0.44
Phoneme segmentation (10 max)	8.78 (1.6)	7.03 (2.7)	18.23**	61/59	0.81
Read pseudowords (15 max)	11.37 (3.1)	7.62 (3.8)	34.00**	61/59	1.09
Read sight words (110 max)	37.13 (13.9)	28.77 (13.7)	11.30**	61/59	0.61
Sum of scores (223 max)	144.0 (17.5)	129.1 (18.6)	20.97**	61/59	0.83
Informal reading inventory					
Oral text accuracy (3.5 max)	2.1 (1.3)	0.9 (1.1)	33.54**	61/59	1.00
Comprehension (3.5 max)	2.1 (1.2)	1.1 (1.0)	33.16**	62/59	0.96

a. ANCOVAs were conducted on spring posttest scores with fall Gates-MacGinitie Reading Tests (4th ed.; MacGinitie, MacGinitie, Maria, & Dreyer, 2000) scores as the covariate. Posttest means are adjusted values, and posttest standard deviations (shown in parentheses) are unadjusted.

b. Missing data account for the varying number of students contributing to each analysis.

c. Effect size was calculated by subtracting adjusted mean of C1 group from adjusted mean of RES group and dividing by the unadjusted pooled standard deviations.

*p < .05. **p < .01. n.s. = not statistically significant.

Another question of interest was whether RES tutoring would raise students' reading achievement from below average in the fall to an average level in the spring based on national test norms. We regarded the minimum average level as performance at the 40th percentile, although this level is higher than that used by other researchers who have considered the 30th percentile as the lower limit (e.g., Torgesen, Rashotte, Alexander, Alexander, & MacPhee, 2003). From Table 1, it is apparent that on the GMRT4, mean performance of RES students was at the 50th percentile in decoding words, thus exceeding the minimum average level. RES students performed at the 38th percentile on the reading comprehension test. Although this is slightly below the 40th percentile, the fact that students reached this level of achievement on a grade-appropriate test is impressive given that they were primarily low-SES, language-minority students with low vocabulary scores. On the word and nonword reading subtests of the WRMT-R, RES-tutored students' mean grade-equivalent scores placed them at the 1.9 grade level, which is the level expected at the end of first grade. These findings show that RES tutoring did indeed raise the means of the group from below average to average levels based on national norms. In contrast, both control groups performed well below the average range (see Table 1).

To determine what proportion of the struggling readers in each group reached average levels by the end of the year on the GMRT4 posttests, scores of individual students were examined. As shown in Table 3, before tutoring began, very few struggling readers scored at or above the 40th percentile across the groups—only 5% to 9%. However, at the end of the year, a much greater proportion of RES-tutored students' than C1 or C2 students' scores reached average levels. When RES students were compared to the entire sample of first graders (i.e., the "WS" group in Table 3), a greater proportion of RES-tutored students reached average levels on the decoding test, and about the same proportions of RES and whole-sample students reached average levels on the comprehension test. These findings show that the RES intervention brought many more of the struggling readers to average levels than would have occurred without this program. In fact, the record of success for RES students was comparable to, if not better than, the record for the entire sample of first graders.

Supplementary Comparison of RES Treatment Group to a Larger Control Group

Because the RES and C1 groups differed on one of the pretests, this left open the possibility that the groups reflected different populations of struggling readers. To address this possibility, a larger control group was created from the five schools implementing the tutoring program. This group included not only the 62 C1 struggling readers but also 62 struggling readers who qualified to receive tutoring but were not selected for the RES or C1 group. Comparison of entry-level performance of the RES group to this larger control group revealed no statistically significant differences on any of the pretests, including the GMRT4 (all ps $> .05$). The only pretest that approached significance was the RES individually administered phoneme segmentation test ($p < .07$), with

Table 3

Proportion (%) of Students in the Reading Rescue Intervention (RES), Small-Group Intervention (SG), Comparison Control Groups (C1, C2, and NI) and Whole First-Grade Sample (WS) Who Performed At or Above the 40th Percentile on the Fall and Spring Gates-MacGinitie Reading Tests (4th ed.; GMRT4)

GMRT4 Tests	RES (n = 64)	C1 (n = 62)	C2 (n = 60)	SG (n = 52)	NI (n = 70)	WS (N = 844)
Fall pretest	5	5	7	9	3	35
Spring posttests						
Decoding	68	35	23	27	31	58
Reading comprehension	52	29	23	24	28	51
Total score	60	34	23	27	30	53

Note. The four comparison groups do not represent four independent samples of struggling readers. C1 (Control 1 from RES schools) and C2 (Control 2 from non-RES schools) consist of 126 students who were reconfigured to form the SG and NI (neither-intervention) groups. The WS group consists of all the first graders, including those in the treatment and control groups.

means favoring the control group. Thus, entry-level reading capabilities did not distinguish the RES group from this larger control group. ANCOVAs of posttest scores with the fall GMRT4 as the covariate revealed significant main effects of treatment, with the RES group outperforming the control group on all three GMRT4 measures (decoding, reading comprehension, total score; all $ps < .001$). These findings replicate those above and show that the positive effects of the RES tutoring intervention held when compared to all struggling readers who were eligible for the program but did not receive it.

Comparison of Students Receiving the Tutoring Intervention, Small-Group Intervention, and Neither Intervention

Students in the C1 and C2 control groups were reconfigured to distinguish those who received a small-group intervention (SG) from those who received neither intervention (NI). Performance of these groups was compared to that of students receiving RES tutoring. ANOVAs of pretest means reported in Table 4 revealed a significant main effect of treatment group on the GMRT4 but no significant effect on the ITBS vocabulary test. Bonferroni post hoc comparisons on the GMRT4 showed that the RES group did not differ significantly from the SG or NI groups, but the NI group performed significantly better than the SG group. These findings indicate that even though some of the schools assigned struggling readers to the SG group and thus partially undermined random assignment of matched pairs, the reading level of students receiving the small-group intervention was not significantly lower than that of students receiving the tutoring intervention.

Table 4

Characteristics and Statistics on Fall Pretests and Spring Posttests for Reading Rescue (RES) Students, Small-Group Intervention (SG) Students, and Neither-Intervention (NI) Control Students

Characteristic or Test	RES (n = 64)	SG (n = 52)	NI (n = 70)	Post Hoc Comparison (Bonferroni)	F Stat[a]	Number of Students[b] (RES/SG/NI)
Age (months)	75.3 (3.9)	74.4 (4.0)	74.1 (5.1)	—	1.30 n.s.	64/52/69
Gender (female/male)	30/34	25/26	35/34		< 1	64/51/69
Fall pretests[c]						
GMRT4 NCE	29.2 (9.2)	26.4 (9.2)	30.8 (9.4)	NI > SG; RES = SG, NI	3.40*	63/52/69
Percentile	16	13	18			
ITBS vocabulary SS	125.3 (8.5)	125.7 (7.9)	127.7 (9.4)	—	1.51 n.s.	64/52/69
Percentile	21	22	26			
Spring posttests[c]						
GMRT4 word decoding						
NCE	49.9 (15.9)	38.1 (12.9)	39.1 (12.4)	RES > SG = NI	14.90**	62/52/68
Percentile	50	28	30			
GMRT4 reading comprehension						
NCE	43.5 (12.9)	37.4 (12.5)	36.0 (12.7)	RES > SG = NI	7.05**	62/52/68
Percentile	38	28	26			
GMRT4 total						
NCE	47.5 (12.9)	38.9 (11.7)	38.7 (10.9)	RES > SG = NI	12.51**	62/52/68
Percentile	45	30	31			

Note. GMRT4 = Gates-MacGinitie Reading Tests (4th ed.; MacGinitie, MacGinitie, Maria, & Dreyer, 2000); ITBS = Iowa Tests of Basic Skills (H. Hoover et al., 2001, 2003); SS = standard score; NCE = normal curve equivalent.

a. ANOVAs were conducted on pretest scores. ANCOVAs were conducted on spring posttest scores with fall GMRT4 scores as the covariate. Posttest means are adjusted values, and posttest standard deviations (shown in parentheses) are unadjusted.

b. Missing data account for the varying number of students contributing to each analysis.

c. In the fall, Level BR (Grade 1 Beginning Reading) of the GMRT4 was administered. In the spring, Level 1 (Grade 1) was administered.

*p < .05. **p < .01. n.s. = not statistically significant.

To adjust for initial pretest differences statistically, ANCOVAs were conducted on outcomes with fall GMRT4 scores as the covariate. The groups were compared only on the GMRT4 posttest because of incomplete data for the SG and NI groups on the other measures. Results revealed main effects of treatment (see Table 4). Bonferroni post hoc pairwise comparisons showed that in all three analyses, RES-tutored students significantly outperformed the SG and NI groups, who did not differ. The effects size favoring RES over SG was 0.70 and RES over NI was 0.74. These are comparable, if not superior, to effect sizes in other studies (see above). These findings indicate that struggling readers who received RES tutoring decoded words and comprehended text more effectively than struggling readers who received the small-group intervention as well as readers who received neither intervention.

The small-group intervention was not effective in raising performance above that of students who did not receive this intervention. Effect sizes comparing mean performance of the SG and NI groups were very small and nonsignificant, ranging from –.08 on the word decoding measure to +.11 on the reading comprehension measure. Moreover, the mean performance of both groups was well below average based on national norms (see Table 4). These findings were somewhat surprising because the small-group intervention program was comprehensive in its coverage of beginning reading skills, with a strong emphasis on systematic phonics instruction, and it held a national reputation as a strong program.

We also assessed the proportion of RES, SG, and NI students whose scores rose to average levels (i.e., at or above the 40th percentile) on the GMRT4 at the end of first grade. As shown in Table 3, although proportions rose substantially from fall to spring, many fewer SG and NI children than RES-tutored children scored within the average range at the end of the school year.

In sum, these findings show that RES tutoring helped struggling first-grade readers progress in their reading more than the small-group intervention program and more than receipt of neither intervention. In contrast, the small-group intervention did not improve the reading achievement of students over that of students not receiving this intervention.

Tutors and Characteristics of Tutoring

During each tutoring session, the tutors kept records. These were used to assess the occurrence of several features of tutoring. Mean values are reported in Table 5. The mean number of total tutoring sessions was 49.6, with a standard deviation of 18, indicating that individual students differed substantially in the amount of tutoring they received. The amount of text reading and rereading completed by students was extensive, $M = 164$ texts, with a large standard deviation, also indicating much variability. It is not surprising that the correlation between the number of sessions and the number of books read was high, $r = .79$, $p < .01$.

The program recommends that books be selected to enable students to practice reading at their instructional level, indicated by text-reading accuracy

Table 5

Mean Performance of Reading Rescue Students on Tutoring Measures and Reading Outcomes as a Function of Whether the Tutor Was a Reading Specialist (R), Other Certified Teacher (O), or Paraprofessional (P)

Tutoring Event or Outcome	R (n = 17)	O (n = 15)	P (n = 26)	Mean	F Stat	Bonferroni Post Hoc
Outcomes						
GMRT4 total[a]	47.0 (14.9)	50.3 (13.8)	46.5 (11.0)	47.7 (12.9)	< 1 n.s.	
WRMT-R word[b]	35.2 (11.3)	37.4 (10.8)	36.3 (6.8)	36.3 (9.3)	< 1 n.s.	
WRMT-R nonword[b]	19.6 (10.5)	16.6 (9.4)	10.8 (5.4)	15.0 (9.0)	6.39**	R > P; O = R, P
Events						
Easing-in sessions	1.31 (1.8)	3.31 (2.7)	2.4 (2.2)	2.3 (2.3)	2.91 n.s.	
Regular sessions[c]	33.9 (16.2)	41.5 (14.8)	55.7 (13.6)	46.0 (17.3)	11.76**	P > R = O
Easing-out sessions	1.0 (1.7)	1.9 (2.4)	1.2 (1.4)	1.3 (1.8)	< 1 n.s.	
Books read/reread	128.1 (54.8)	154.2 (52.6)	191.7 (59.7)	164.3 (62.2)	6.50**	P > R; O = P, R
Text accuracy 98%-100%	54% (.26)	48% (.27)	48% (.23)	50% (.25)	< 1 n.s.	
Text accuracy 90%-97%	37% (.17)	33% (.15)	40% (.13)	37% (.14)	1.14 n.s.	
Text accuracy < 90%	13% (.11)	19% (.19)	14% (.15)	15% (.15)	< 1 n.s.	
Adherence to program[d]	14.4 (2.7)	12.6 (3.3)	13.4 (3.4)	13.5 (3.2)	1.15 n.s.	

Note. Standard deviations are shown in parentheses. GMRT4 = Gates-MacGinitie Reading Tests (4th ed.; MacGinitie, MacGinitie, Maria, & Dreyer, 2000); WRMT-R = Woodcock Reading Mastery Tests–Revised (Woodcock, 1987, 1998).

a. GMRT4 means are normal curve equivalent scores.

b. WRMT-R means are raw scores.

c. Number of regular sessions does not include easing-in and easing-out sessions.

d. Adherence to program was based on judgments of tutors' records with 13 the minimum acceptable and 24 the maximum.

*p < .05. **p < .01. n.s. = not statistically significant.

scores ranging from 90% to 97%. To examine accuracy levels, running records were evaluated. These were recorded during the first independent reading of newly introduced books after the books had been previewed and coached during the previous session. As indicated by means in Table 5, only 37% of the books were read at the instructional level. More books, 50%, were read at an independent level (98% to 100% accuracy), and only 15% of the readings were at a frustration level (below 90% accuracy). This reveals that the largest proportion of books were read at accuracy levels higher than that recommended by the program.

The records that tutors kept were used to assess their adherence to the program. Raters evaluated adherence on six dimensions, such as the number of books read in each session, the difficulty level of the books, the appropriateness of the sentences written, and whether tutoring was provided on a regular basis. Ratings were summed to provide an overall score (maximum of 24). A total score of 13 was considered the minimum level indicating full adherence to recommended practice. As seen in Table 5, the mean value, 13.5, shows that tutors on average were performing slightly above this level. The small standard deviation (3.2) indicates that there was limited variation in scores. Inspection of scores showed that 63% of the tutors displayed full adherence to program procedures. Two dimensions displayed the strongest adherence, indicating that tutors were best at selecting the appropriate number of books to read and having students write appropriate sentences that contained sight words and words with the letter–sound relations being learned. These findings indicate that the majority of the tutors did conform to expected instructional procedures.

To determine whether tutors' adherence to the program influenced how well their students performed on outcome measures, correlations were examined between adherence scores and the various posttests. Results revealed little relationship, with values ranging from $r = 0.06$ to $r = 0.20$, all ps $> .05$, perhaps because tutors showed little variation in their adherence to the program.

Three types of tutors worked with students: reading specialists, other credentialed personnel, and paraprofessionals. To determine whether the three types differed in their tutoring effectiveness, ANOVAs were conducted with type of tutor as the independent variable and various measures taken on their students as the dependent variables. All but 3 of the 59 tutors taught only 1 of the 64 RES students, so most data points emerged from different tutors.

The three types of tutors did not differ in the reading level of their students when tutoring began, as indicated by a nonsignificant main effect of tutor type on the GMRT4 fall pretest ($F < 1$, $p > .05$). This is as expected, because students were randomly assigned to tutors. ANCOVAs of reading outcomes with fall GMRT4 as the covariate revealed no main effects of tutor type on the GMRT4 posttests, the WMRT-R word identification posttest, or the RES program posttests (all Fs < 1). However, a difference was detected on the WRMT-R nonword test assessing students' ability to decode unfamiliar words. As shown in Table 5, Bonferroni post hoc analyses revealed that

reading specialists were significantly better at teaching students to decode than paraprofessionals ($p < .003$). The effect size was 1.11.

Measures of events that occurred during tutoring were subjected to ANOVAs with tutor type as the independent variable. From Table 5, it is apparent that tutors differed significantly in the number of regular sessions they conducted with students. Post hoc analyses revealed that the paraprofessionals conducted significantly more sessions than the other two types of tutors. However, the tutors did not differ in the number of easing-in or easing-out sessions. In addition, the tutors differed significantly in the number of books that they had students read and reread. Paraprofessionals exceeded reading specialists in this regard, which is not surprising because they provided more tutoring sessions. Other tutoring events were examined as well, but as shown in Table 5, no other effects of tutor type were detected (all ps $> .05$). The fact that reading specialists' students attained the same reading level in fewer tutoring sessions than paraprofessionals' students suggests that reading specialists were more efficient in their tutoring.

From these findings, we conclude that paraprofessionals tutored as effectively as the credentialed teachers who were not reading specialists. Moreover, paraprofessionals tutored as well as the reading specialists in strengthening students' word reading and text comprehension skills but not as well in teaching the skills that benefit nonword decoding. In addition, paraprofessionals took more sessions to work with students than reading specialists, suggesting that they were less efficient in carrying out the tutoring.

Prediction of Reading Achievement for Students Receiving RES Tutoring

The next question of interest was whether entry-level abilities of tutored students and characteristics of tutoring might explain why some students improved more in reading than others. The outcome measure was the total GMRT4 spring score, which was highly correlated with the other reading outcome measures (i.e., rs ranged from .62 to .93), indicating that it was indicative of general reading performance. Correlation coefficients between this outcome and each of the other variables involving students' entry-level skills and tutoring characteristics revealed several statistically significant values, reported in column 1 of Table 6. The significant predictors included performance on the fall GMRT4 pretest, letter writing on the RES classwide screen, student RES program individual student pretests (total score), the proportions of books read at the three accuracy levels, and the number of easing-in sessions. The two accuracy levels, 90% to 97% and < 90%, were each strongly and negatively correlated with the 98% to 100% accuracy measure (see Table 6), indicating that all three measures reflect one factor, text reading accuracy. This factor is best represented by the 98% to 100% accuracy measure. The number of easing-in sessions correlated negatively with the GMRT4. One reason may be that students with more severe reading problems often took longer to assess because tutors spaced out the tests to avert a sense of failure. None of the other student or tutoring variables was correlated significantly with reading achievement (all ps $> .05$).

Table 6

Entry-Level Characteristics of Tutored Students and Features of Tutoring That Were Significantly Correlated With GMRT4 Posttest Scores

Measure	1	2	3	4	5	6	7	8	*M*	*SD*
1. GMRT4 posttest	—								47.4	12.8
Pretests										
2. GMRT4	.46**	—							29.2	9.2
3. RES individual	.43**	.41**	—						89.6	15.0
4. Write uppercase letters	.26*	.17	.15	—					16.0	6.2
5. Write lowercase letters	.26*	.28*	.12	.57**	—				11.3	6.8
Features of tutoring										
6. Text accuracy 98%-100%	.66**	.49**	.41**	.22	.22	—			49%	.24
7. Text accuracy 90%-97%	−.56**	−.24	−.40**	−.19	−.05	−.76**	—		38%	.15
8. Text accuracy < 90%	−.45**	−.53**	−.36**	−.10	−.26	−.77**	.14	—	15%	.15
9. Easing-in sessions	−.32*	−.02	−.05	−.10	−.38**	−.19	.11	.07	2.3	2.2

Note. There were between 56 and 64 observations contributing to each correlation as a result of a few missing values. GMRT4 = Gates-MacGinitie Reading Tests (4th ed.; MacGinitie, MacGinitie, Maria, & Dreyer, 2000); RES = Reading Rescue.
*$p < .05$. **$p < .01$.

Regression analyses were conducted with the spring GMRT4 total score as the reading outcome. The predictors were the four pretest measures, the 98% to 100% text accuracy measure, and the easing-in measure. When all of the predictors were entered as a single set, only the 98% to 100% text accuracy variable explained significant variance in the model, with intercept = 21.3, slope = 25.5, $p < .001$. When the accuracy variable was entered first, none of the other variables explained any significant additional variance (all ps $> .05$). When the accuracy variable was entered last, it explained significant additional variance ($p < .001$), with r^2 increasing from .36 to .51, and contributions of each of the other predictors were nonsignificant (all ps $> .05$). These findings indicate that high text-reading accuracy during tutoring was the strongest predictor and the only unique predictor of students' reading achievement at the end of first grade.

The strong relationship between high text-reading accuracy during tutoring and reading is open to at least two interpretations. One is that the relationship may reflect the effects of tutoring. Children whose tutors enabled them to read a greater proportion of texts at high accuracy levels became better readers by the end of the year than children whose tutors did not do this. Higher levels might have been achieved either by tutors' selecting easier texts or by tutors previewing and coaching students more effectively through the texts during the previous session, when the books were introduced. The other interpretation is that students who entered the study as better readers were able to read texts at higher levels of accuracy during tutoring than the weaker readers and also achieved higher reading scores at the end of first grade. Findings of the regression analyses appear to rule out the second interpretation. Pretest reading skills of students taken when they began tutoring did not explain variance on the reading posttest once the impact of the text accuracy tutoring variable was considered. Also, these students had very limited reading skills when tutoring began, so their growth in reading resulted primarily from the instruction that they received, not from entry-level reading skill. These findings favor the explanation that students' experiences of learning to read texts at high accuracy levels during tutoring influenced how well they read as a result of the tutoring.

Discussion

Findings of the present study offered strong support for the effectiveness of the Reading Rescue intervention model. First-grade, language-minority struggling readers who received RES tutoring made significantly greater improvement in reading than language-minority struggling readers who did not receive this tutoring but were enrolled in the same schools, and also greater improvement than struggling readers who were enrolled in comparable schools not using the program. Also, students receiving RES tutoring outperformed students receiving a commercially produced small-group intervention program mandated by the school district for struggling readers. An advantage occurred despite the fact that both intervention programs were comprehensive

and provided instruction in the same components. Greater gains in reading by tutored students were observed on standardized measures of word and pseudoword reading as well as reading comprehension. In fact, RES tutoring raised the majority of students' reading from below-average to average levels. This contrasted with many fewer students reaching average levels in the comparison groups. The RES program proved effective with language-minority students who have been underrepresented in prior controlled studies of literacy instruction. Effect sizes with this group were comparable to, if not higher than, effect sizes found in tutoring studies in general.

The reading achievement of students who received Reading Rescue tutoring appeared to be explained primarily by one aspect of their tutoring experience—reading texts at a high level of accuracy, between 98% and 100%. This finding needs more careful consideration because it appears to contradict a commonly accepted view about the quality of text reading that is most effective for building fluency. Rasinski (2003) stated this shared wisdom in his book *The Fluent Reader*: "As with any direct-teaching activity, the greatest gain will occur when the difficulty of the material is at the student's instructional level (i.e., 90-95 percent accuracy in word recognition)—neither too hard nor too easy, just right!" (p. 63) In fact, the developer of the RES tutoring model adopted this view and recommended it to tutors. However, present findings suggest otherwise, at least for beginning readers.

Following program guidelines, tutors in the present study coached students through a book during one session and then in the next session had students reread the book, this time independently, and recorded their accuracy. Results revealed that the proportion of texts read at an instructional level (90% to 97% accuracy) was strongly and *negatively* correlated with positive growth in reading, whereas the proportion of texts read at an independent level (98% to 100% accuracy) was strongly and positively correlated with reading growth. In fact, the latter, high accuracy measure figured more prominently than any other variable in explaining variance on the reading posttest. Even when students' entry-level reading skill was controlled statistically, this tutoring variable still explained significant variance. Before concluding that reading text at high accuracy levels is important for promoting growth in struggling readers, we need to consider alternative interpretations.

In the present study, the high accuracy levels that proved important occurred on text that was read independently after prior coaching. It may be that conventional wisdom regarding recommended accuracy levels (i.e., 90% to 97%) applies to text read without prior coaching. Reading text cold is certainly susceptible to more errors. What difference this makes deserves further research.

Monitoring high text accuracy levels may have proven important for reading achievement because it reflected the impact of the preview session for improving students' word reading. According to theories of word learning (Ehri, 1992; Perfetti, 1992; Rack, Hulme, Snowling, & Wightman, 1994; Share, 1999, 2004b), readers remember how to read words when they

possess knowledge of grapheme–phoneme relations that provide the connections securing the written words to their pronunciations and meanings in memory. Perhaps more thorough coaching by tutors to help students use decoding skill to secure words in memory as they previewed the decodable texts throughout tutoring was an important factor strengthening students' word- and text-reading accuracy and their general reading ability. This possibility needs further study.

Another factor possibly contributing to the effectiveness of the Reading Rescue model merits consideration, specifically, the criteria used to qualify struggling readers for the RES program. Rather than selecting all first graders at the bottom of the class, the program selected those who knew at least 17 letters yet received low scores on other literacy tasks and lacked any text-reading ability. Many studies have shown that letter name knowledge is the best single kindergarten predictor of success in learning to read (see review by Scarborough, 2001). If students have not learned the shapes and names or sounds of letters, it is impossible for them to acquire word decoding skill. Share (2004a) showed that if students already knew the names of letters, it was much easier for them to learn their sounds because most names contain their sounds. Letter names or sounds are typically taught in kindergarten, so first graders are expected to know them. Selecting students who already possess substantial letter knowledge reduces the need for RES tutors to spend time teaching letters and allows instruction to focus primarily on the application of letter knowledge to acquire reading and spelling skills. According to the model, low-performing readers who lack sufficient letter knowledge at the beginning of first grade are admitted for RES tutoring later in the year once they have learned enough letters to qualify. This more refined tailoring of program characteristics to suit student capabilities very likely enhanced its effectiveness.

Although having letter knowledge when they began Reading Rescue tutoring may have enabled students to make better progress, it is important to note that this factor does not explain why RES students outperformed control students at the end of the year. In the present study, control students met the letter knowledge criteria as well when they were selected for the study. This suggests that the nature of the instruction as it *capitalized* on entry-level letter knowledge was of primary importance in explaining the greater success of RES students in learning to read.

One purpose of the present study was to assess the effectiveness of various types of tutors. Findings indicate that paraprofessionals delivered RES tutoring as effectively as reading specialists and credentialed teachers except in two respects. Paraprofessionals were less successful in affecting students' ability to decode pseudowords than reading specialists were. Paraprofessionals provided more tutoring sessions than the other tutors and, as a result, had students read and reread more books. The fact that paraprofessionals' students were just as good at reading real words and comprehending text as students taught by the other tutors suggests that paraprofessionals were less efficient in their tutoring.

It is interesting that one of these findings replicates and extends that reported by Brown et al. (2005), who found that students tutored by paraprofessionals did not differ from students tutored by teachers on tests of word and text reading but only on the WRMT-R pseudoword reading test, the same test used in the present study. Although both studies show agreement, confidence in the present finding is greater because tutors here were randomly assigned to students whereas tutors in the Brown et al. study were not.

What might explain the shortcoming in paraprofessionals' tutoring that was observed in both studies? Perhaps paraprofessionals spent less time teaching decoding than the reading specialists did. In the present study, paraprofessionals had their students read more books than other tutors. Perhaps paraprofessionals lacked the background knowledge to teach phonics effectively, or perhaps they did not consider decoding as important to teach. According to Moats (2000), a high level of knowledge about the alphabetic system is needed to teach systematic phonics programs effectively. These findings and their explanation deserve further study to determine whether more extensive staff development in decoding instruction is needed to better prepare paraprofessionals as RES tutors.

Findings of the present study indicated that tutoring was more effective than small-group instruction for teaching reading to struggling readers, despite the fact that the skills taught were similar in the two programs. Tutoring may have proven more successful because of the greater ease of adapting instruction to the needs of individuals and the greater amount of reading practice with feedback that was possible, or because of specific features of the RES program, such as the use of the *Ready Reader* decodable books. Alternatively, circumstances specific to this study may have created an advantage for tutoring. The RES program had been in use for 2 to 3 years, whereas the small-group program was in its 1st year of implementation, so tutors and staff were more experienced delivering RES than the small-group intervention. Also, RES provided more training for tutors than did the small-group program, so this may have provided an advantage. On the other hand, new programs can benefit from a novelty effect in the 1st year, and this would have benefited the small-group program. Previous studies have not consistently shown stronger effects of tutoring. Because of the above uncertainties, results of the present study offer only tentative support for the greater effectiveness of tutoring over small-group instruction.

The author of the Reading Rescue program suggests that implementing the program in schools may raise the reading achievement of even those first graders who are not tutored (N. Hoover, 1996). This general benefit may result from the professional development and tutoring experience that teachers and paraprofessionals on the staff receive. Some evidence for this was detected in the present study. Control students enrolled in RES schools outperformed control students in non-RES schools on the GMRT4 word decoding posttest. The possibility that participation in the RES program improved teachers' ability to teach decoding outside of the RES program in their classrooms merits systematic study.

Possible limitations of the present study need to be considered. Because we were unable to assign all the students randomly to the tutored and control groups, our design was quasi-experimental. However, we took steps to match students in assigning them to the groups, and we adjusted for any entry-level differences in our statistical analyses of outcomes. In addition, we compared the tutored group to several different comparison groups and found an advantage for tutoring in all cases.

A second possible limitation is that our paraprofessionals were relatively well educated, with many holding college degrees. Our findings may not generalize to paraprofessionals with less than a high school diploma. This possibility awaits attention.

A third limitation is that only short-term effects of the RES tutoring model were tested at the end of the year when instruction ended. Of equal interest is whether the effects of tutoring in first grade are lasting and maintain students' reading at average levels in subsequent grades. This matter awaits study.

A fourth limitation bears on the comparison of tutoring to small-group intervention. These two forms of instruction were confounded with other variables. The two systematic phonics programs were similar but not identical. RES tutors received more training and were more experienced using their program than small-group tutors were. The time that students spent receiving small-group instruction was not monitored, so comparability to RES time is unknown. As a result, it is not clear that the instructional delivery unit was the critical factor explaining the difference in students' performance following the interventions, so any conclusion about the greater effectiveness of tutoring over small-group instruction remains tentative.

Present findings do make it clear, however, that the small-group intervention was not effective, whatever the cause. Struggling readers who received this intervention performed no better than struggling readers who did not receive it. This came as a surprise. The program held a strong reputation and was mandated by the district. The program was scripted and explicit regarding the steps for delivering instruction. Reading specialists as well as paraprofessionals served as the teachers. Sessions were conducted on a regular basis at all the schools. Because this program is used widely, uncertainty about its effectiveness merits further study. One possibility to explore is that some of the bottom-level students in the small groups did not know many letters and this necessitated spending time teaching letters, hence reducing the time spent teaching other reading skills needed by students who already knew their letters, that is, control students who were assigned to the small-group sample in the present study.

Despite some limitations, the present study has many strengths. This was a well-designed, field-based research study, not all that common in the literature and never perfectly executed because of the capriciousness of school environments. The study was conducted in several typical urban, low-SES elementary schools with language-minority children, a population that has been neglected in controlled studies of literacy instruction. Procedures of the intervention programs were implemented by school personnel with

few if any modifications introduced by researchers. As a result, the study has high external validity. Random assignment was employed with tutors, so the internal validity of conclusions regarding their effectiveness is high. Few if any studies have compared different types of tutors or have analyzed their differential effectiveness as we have done in this study. Although random placement of students in treatment and control groups was not fully implemented, statistical adjustments applied to multiple outcome measures maintained the internal validity of the study.

Findings advance our understanding of the nature and impact of early intervention, and they carry important implications for practice. Results offer evidence for the effectiveness of a specific tutoring intervention model's providing comprehensive instruction in multiple components of reading needed by struggling beginning readers. The model proved effective with the neediest students, those in urban, high-poverty schools with large numbers of language-minority students. The core of the intervention model is not a commercial program but rather involves a service offered by a nonprofit foundation whose staff helps schools institute a tutoring program by using teachers and paraprofessionals already employed at the school. As a result, implementation costs are substantially less than programs, such as Reading Recovery, that require adding credentialed, specially trained teachers to the staff. The present study contributes by documenting the effectiveness of this approach to reading intervention. In addition, findings show generally how effective intervention models for struggling readers in first grade might be structured.

Note

Reading Rescue was developed in cooperation with the University of Florida and is sponsored by a charitable, not-for-profit organization, the Literacy Trust. This study was an independent evaluation conducted with grants from the Literacy Trust and Pearson Learning. All aspects of the research were under the control of investigators who were neither employed by nor affiliated with the funding agencies. For more detailed information about the Reading Rescue intervention model, contact the Literacy Trust, www.literacytrust.org. We express gratitude to the following colleagues for their assistance: Nora Hoover, Kathy Kaufman, Alba Langenthal, Tara Mastrorilli, Belinda Nix, Henry Park, and Benedict Silverman. We thank the district and school personnel, including principals, teachers, and paraprofessionals, for their cooperation and assistance in the conduct of our study.

References

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.

Biemiller, A. (1999). *Language and reading success: From reading research to practice: A series for teachers* (Vol. 5). Cambridge, MA: Brookline.

Block, C., & Pressley, M. (Eds.) (2002). *Comprehension instruction: Research-based best practices*. New York: Guilford.

Britt, L. (2002). *Reading Rescue in District Six community schools of the New York City public school system* (Technical Report). Gainesville, FL: Literacy Trust.

Brown, K., Morris, D., & Fields, M. (2005). Intervention after Grade 1: Serving increased numbers of struggling readers effectively. *Journal of Literacy Research*, *37*, 61–94.

Clay, M. (1985). *The early detection of reading difficulties* (3rd ed.). Portsmouth, NH: Heinemann.

Clay, M. (1993a). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.

Clay, M. (1993b). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.

Clay, M. (2002). *An observation survey of early literacy achievement* (2nd ed). Portsmouth, NH: Heinemann.

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cox, B., & Hopkins, C. (2006). Building on theoretical principles gleaned from Reading Recovery to inform classroom practice. *Reading Research Quarterly*, *41*, 254–267.

Cunningham, P., & Hall, D. (1997). *Month-by-month phonics for first grade: Systematic, multilevel instruction*. Greensboro, NC: Carson-Dellosa.

D'Agostino, J., & Murphy, J. (2004). A meta-analysis of Reading Recovery in United States schools. *Educational Evaluation and Policy Analysis*, *26*, 23–38.

Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–143). Hillsdale, NJ: Lawrence Erlbaum.

Ehri, L., Nunes, S., Stahl, S., & Willows, D. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393–447.

Ehri, L., Nunes, S., Willows, D., Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*, 250–287.

Elbaum, B., Vaughn, S., Hughes, M., & Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*, 605–619.

Elkonin, D. (1973). U.S.S.R. In J. Downing (Ed.), *Comparative reading: Cross-national studies of behavior and processes in reading and writing* (pp. 551–579). New York: Macmillan.

Englebretson, R., Hiebert, E., & Juel, C. (2000). *Ready readers*. Parsippany, NJ: Modern Curriculum.

Foorman, B., Francis, D., Shaywitz, S., Shaywitz, B., & Fletcher, J. (1997). The case for early reading intervention. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 243–264). Mahwah, NJ: Lawrence Erlbaum.

Greany, K., Tunmer, W., & Chapman, J. (1997). Effects of rime-based orthographic analogy training on the word recognition skills of children with reading disability. *Journal of Educational Psychology*, *89*, 645–651.

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

Hoover, H., Dunbar, S., Frisbie, D., Oberley, K., Bray, G., Naylor, R., et al. (2003). *Iowa Tests of Basic Skills complete/core battery norms and score conversions*. Itasca, IL: Riverside.

Hoover, H., Dunbar, S., Frisbie, D., Oberley, K., Ordman, V., Naylor, R., et al. (2001). *The Iowa Tests of Basic Skills*. Itasca, IL: Riverside.

Hoover, N. (1994). *Reading Rescue for elementary readers: An early literacy intervention program 1993-1994 end-of-pilot year report* (Tech. Report No. 1). Gainesville, FL: University of Florida.

Hoover, N. (1995). *Reading Rescue for elementary readers: An early literacy intervention program 1994-1995 end-of-second year report on student progress at Interlachen Elementary School, Putnam County, Florida. Gainesville, Florida* (Tech. Report No. 2). Gainesville, FL: University of Florida.

Hoover, N. (1996, April). *Primary teachers learn to teach at-risk students to read: Empowerment and improvement through a school-based inservice model*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Hoover, N. (2001). *Reading Rescue: A handbook for tutors*. Gainesville, FL: Literacy Trust.

Hoover, N., & Lane, H. (2001, May). *Preventing literacy failure through cost-effective tutoring: A report on outcomes in high-poverty and rural schools*. Paper presented at the annual meeting of the International Reading Association, San Francisco.

Hoover, N., & Sullivan, L. (1996, April). *Teaching at-risk children to read: One school's successful implementation of cost-effective tutoring*. Paper presented at the annual meeting of the American Educational Research Association, New York.

International Reading Association. (2006). *Position statements on (1) teacher preparation and (2) excellent reading teachers*. Available from http://www.reading.org/resources/issues/positions.html

Iversen, S., & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology*, *85*, 112–126.

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

Juel, C., & Roper-Schneider, D. (1985). The influence of basal readers on first grade reading. *Reading Research Quarterly*, *20*, 134–152.

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L. (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Itasca, IL: Riverside.

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L. (2002). *Gates-MacGinitie Reading Tests, 4th Edition Technical Report*. Itasca, IL: Riverside.

Moats, L. (2000). *Speech to print: Language essentials for teachers*. Baltimore, MD: Brookes.

Morris, D. (1999). *The Howard Street tutoring manual: Teaching at-risk readers in the primary grades*. New York: Guilford.

Muller, P., & Davies, R. (2004). *Reading Rescue in inner city schools: An experimental study examining reading outcomes of a one-on-one tutoring intervention*. Bloomington: Indiana University, the Indiana Evaluation Center.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Rockville, MD: National Institute of Child Health and Human Development Clearinghouse.

Perfetti, C. (1992). The representation problem in reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale, NJ: Lawrence Erlbaum.

Rack, J., Hulme, C., Snowling, M., & Wightman, J. (1994). The role of phonology in young children learning to read words: The direct-mapping hypothesis. *Journal of Experimental Child Psychology*, *57*, 42–71.

Rasinski, T. (2003). *The fluent reader: Oral reading strategies for building word recognition, fluency, and comprehension*. New York: Scholastic.

Santa, C., & Hoien, T. (1999). An assessment of Early Steps: A program for early intervention of reading problems. *Reading Research Quarterly*, *34*, 54–79.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York: Guilford.

Schwartz, R. (2005). Literacy learning of at-risk first-grade students in the Reading Recovery early intervention. *Journal of Educational Psychology*, *97*, 257–267.

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

Shanahan, T., & Beck, I. (2006). Effective literacy teaching for English-language learners. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth* (pp. 415–488). Mahwah, NJ: Lawrence Erlbaum.

Shanker, J., & Ekwall, E.(2003). *Ekwall/Shanker Reading Inventory* (4th ed). Needham Heights, MA: Allyn & Bacon.

Share, D. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*, 95–129.

Share, D. (2004a). Knowing letter names and learning letter sounds: A causal connection. *Journal of Experimental Child Psychology*, *88*, 213–233.

Share, D. (2004b). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, *87*, 267–298.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–406.

Torgesen, J., Rashotte, C., Alexander, A., Alexander, J., & MacPhee, K. (2003). Progress toward understanding the instructional conditions necessary for remediating reading difficulties in older children. In B. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 275–297). Baltimore, MD: York.

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

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

Voyager Expanded Learning. (2004). *Voyager Passport*. Dallas, TX: Author.

Wasik, B., & Slavin, R. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly*, *28*, 178–200.

Woodcock, R. (1987). *Woodcock Reading Mastery Tests–Revised*. Circle Pines, MN: American Guidance Service.

Woodcock, R. (l998). *Woodcock Reading Mastery Tests–Revised Examiner's Manual Normative Update*. Circle Pines, MN: American Guidance Service.

Manuscript received April 4, 2006
Revision received August 3, 2006
Accepted September 11, 2006