Exceptional Children

Vol. 76, No. 4, pp. 394-416.

Effectiveness of a Supplemental Early Reading Intervention Scaled Up in Multiple Schools

CAROLYN A. DENTON
University of Texas Health Science Center at Houston

KIM NIMON
University of North Texas

PATRICIA G. MATHES
Southern Methodist University

ELIZABETH A. SWANSON
University of Texas at Austin

CAROLINE KETHLEY
Southern Methodist University

TERRI B. KURZ
University of Texas at Austin

MINYI SHIH
California State University, Los Angeles

ABSTRACT: *This effectiveness study examined a supplemental reading intervention that may be appropriate as one component of a response-to-intervention (RTI) system. First-grade students in 31 schools who were at risk for reading difficulties were randomly assigned to receive Responsive Reading Instruction (RRI; Denton, 2001; Denton & Hocker, 2006;* n *= 182) or typical school practice (TSP;* n *=240). About 43% of the TSP students received an alternate school-provided supplemental reading intervention. Results indicated that the RRI group had significantly higher outcomes than the TSP group on multiple measures of reading. About 91% of RRI students and 79% of TSP students met word reading criteria for adequate intervention response, but considerably fewer met a fluency benchmark.*

The well documented gap between educational research and practice has increasing relevance in light of the emphasis in federal legislation—such as the No Child Left Behind Act and the 2004 reauthorization of the Individuals With Disabilities Education Improvement Act (IDEA)—on universally high standards for all students, including those with disabilities. IDEA emphasizes the critical role of early intervening services and authorizes the use of data documenting students' responsiveness to high-quality, research-based intervention in the identification of students with learning disabilities (LD). The Council for Exceptional Children (CEC; 2007) position statement on such response to intervention (RTI) approaches states that RTI "must be viewed as a schoolwide initiative, with special education as an explicit part of the framework . . . to identify and address the academic and behavioral needs of learners" (p. 1), noting that the implementation of a comprehensive RTI model "may reduce the number of students referred for special education, promote

effective early intervention, provide diagnostic information to consider in the identification of a disability, and/or may reduce the impact of a disability on a child's academic progress" (p. 2). Since the reauthorization of IDEA in 2004, RTI models have been increasingly implemented across the country. In a recent nationwide survey of special education administrators and other district personnel (Spectrum K12, 2009), 71% of respondents indicated that their school districts were either piloting or implementing RTI models (margin of error 4.6%).

Many students who currently receive intervention within RTI models receive this assistance because of reading difficulties (Haager, Klingner, & Vaughn, 2007; Jimerson, Burns, & VanDerHeyden, 2007; Spectrum K12, 2009). Preventative RTI reading models most often consist of three tiers of intervention (Batsche et al., 2006; Spectrum K12, 2009), in which Tier 1 is high-quality classroom reading instruction delivered to all students. Students with inadequate RTI in Tier 1 receive Tier 2 supplemental intervention in addition to classroom instruction, provided either by their classroom teachers or other interventionists. Students who continue to struggle receive Tier 3 intervention delivered with higher intensity (e.g., smaller groups, longer sessions, longer duration, highly qualified teachers), sometimes provided within special education (CEC, 2007). This study examined one Tier 2 reading intervention that may be appropriate for use within such an RTI model.

RESEARCH ON EARLY READING INTERVENTION

Based on a comprehensive review of research, the National Research Council proposed that many reading difficulties could be prevented when young children receive effective instruction (Snow, Burns, & Griffin, 1998). Considerable research evidence has documented strong positive effects of preventative instructional interventions for students in the early grades who are at risk for serious reading difficulties or disabilities (e.g., Blachman et al., 2004; Denton, Fletcher, Anthony, & Francis, 2006; Lovett et al., 2000; Mathes et al., 2005; Simmons, Kame'enui, Stoolmiller, Coyne, & Harn, 2003; Torgesen et al., 2001; Torgesen, Wagner, Rashotte, et al., 1999; Vaughn, Cirino et al., 2006; Vaughn, Mathes et al., 2006; Vellutino et al., 1996). Wanzek and Vaughn (2007) synthesized studies of reading interventions having at least 100 sessions, and found that most reported effect sizes in the moderate to large range and that these effects were larger in kindergarten and Grade 1 than in the upper elementary grades.

IDEA emphasizes that students' RTI must be evaluated when they are provided with high-quality instruction using research-validated programs and approaches. About 75% of the respondents to the Spectrum K12 (2009) survey indicated that their need for evidence-based instructional resources was a significant obstacle to their implementation of RTI models. There has been a substantial research focus in recent years on the efficacy of reading interventions (e.g., Wanzek & Vaughn, 2007), but some have questioned whether reading interventions implemented in controlled studies can produce similar results in "real schools" (e.g., D. Fuchs, Stecker, & Fuchs, 2008). D. Fuchs et al. suggest that this has particular relevance for the implementation of RTI models. If the potential of the RTI approach is to be realized, instructional interventions with evidence of efficacy in controlled studies must also be effective when scaled up in natural school contexts.

If the potential of the RTI approach is to be realized, instructional interventions with evidence of efficacy in controlled studies must also be effective when scaled up in natural school contexts.

STUDIES OF EFFICACY AND EFFECTIVENESS

This idea of "scaling up" evidence-based educational practices in multiple schools is related to the concept of efficacy and effectiveness studies, often cited in medical and clinical research (e.g., Hunsley & Lee, 2007; Smith et al., 2007) and more recently in educational research (e.g., Raudenbush, 2007). Raudenbush describes a two-

stage process. First, *efficacy* studies demonstrate that the innovation has positive effects on student outcomes when researchers take steps to assure that it is implemented with very high treatment fidelity and quality. If the results of efficacy trials are positive, studies of *effectiveness* are conducted to demonstrate that the innovation can produce similarly strong effects when implemented in field settings where resources, teacher qualifications, and the quality and intensity of implementation will vary. Raudenbush suggests that effectiveness studies are essential if treatments are to become widely accessible, particularly in high-poverty schools where resources like those described in efficacy studies are often unavailable.

PURPOSE OF THE STUDY

The purpose of this study was to determine whether a first-grade small-group reading intervention with evidence of efficacy for at-risk readers would demonstrate effectiveness when implemented in multiple schools with reduced researcher control. The efficacy of the intervention had been previously demonstrated in a randomized, controlled field trial in six schools in a single school district (Mathes et al., 2005). In the current study, we examined its effects on student outcomes when implemented in 31 schools with varying contexts. Within the Raudenbush (2007) model, this study was a test of intervention effectiveness. We did not evaluate a schoolwide comprehensive RTI implementation, but the effectiveness of a reading intervention approach that may be appropriate as one component of an RTI model.

Efficacy of the Intervention

Mathes and her colleagues (2005) investigated the efficacy of two small-group supplemental interventions for students at risk for reading difficulties in first grade. The two interventions differed in theoretical underpinnings and in aspects of instructional delivery, but both were based on elements that have research-demonstrated effectiveness for the instruction of students who struggle to learn to read. One of these interventions was Responsive Reading Instruction (RRI; Denton, 2001; Denton & Hocker, 2006), the intervention that is the subject of the current study.

In the efficacy study, conducted in six schools in a large urban school district, Mathes et al. (2005) identified a group of students at risk for serious reading difficulties at the beginning of first grade. All students received enhanced classroom reading instruction, in which classroom teachers received professional development and were provided with monthly graphs of participating students' progress in oral reading fluency. Students were randomly assigned within classrooms to one of three supplemental intervention conditions: (a) typical school practice (i.e., whatever services schools typically provided to struggling readers but no researcher-provided intervention), (b) the RRI (Denton, 2001; Denton & Hocker, 2006) intervention, or (c) a second researcher-provided small-group intervention (Mathes & Torgesen, 2005). Supplemental intervention in the last two conditions was provided daily for 40 min in groups of 3 to 4 students for about thirty weeks in addition to regular classroom reading instruction. Three certified teachers selected, trained, coached, and closely supervised by the researchers implemented RRI. Mathes et al. found that students in the RRI group performed better at posttest in phonological awareness, word reading, spelling (Cohorts 1 and 2 combined), and oral reading fluency (Cohort 2) than students who had received only classroom reading instruction without supplemental researcher-provided intervention. At posttest, about 93% of students who received RRI demonstrated adequate intervention response, based on a criterion of performance at or above the 30th percentile on the Basic Reading Skills composite of the Woodcock-Johnson III Tests of Achievement (WJ-III; Woodcock, McGrew, & Mather, 2001). Although the Mathes et al. study demonstrated positive effects associated with RRI when implementation was standardized and of high quality, what was unknown was whether the intervention approach would be effective when implemented in various school settings with reduced researcher control.

Research Questions

The research questions addressed in the current study were:

1. Are there significant differences in phonemic awareness, word identification, phonemic decoding, spelling, reading comprehension, and oral reading fluency for first grade students at risk for reading difficulties who receive the research intervention and those who receive the reading instruction and intervention typically provided in their schools?
2. What percentage of students in each group demonstrates adequate RTI?

METHOD

Context and Participants

Schools. The study was conducted in 31 elementary schools from 16 school districts in a southwestern state over a 2-year period, with two cohorts of first-grade students. Three of the 31 schools participated only in Year 1 of the study, 20 participated only during Year 2, and 8 participated in both Years 1 and 2. Seventeen of the schools were located in rural areas or small towns, 12 were in suburbs of larger cities, and 2 were located in an urban area. The smallest school served only 26 first-grade students schoolwide, whereas the largest served 183 first-grade students. The school districts spanned a distance of approximately 260 miles north to south and 110 miles east to west. On average, 53.3% of students in the schools had economic disadvantage (*SD*, 20.1%); the percentage ranged from 5.6% to 91.9% across schools.

Three schools (all rural) withdrew from the study during Year 1 and one (suburban) withdrew during Year 2. On average, 54% of students in these schools had economic disadvantage, ranging from 33% to 71%. According to research staff field notes, one school withdrew because it was accepted into the Reading First program, which provided them with federal assistance but required that they implement only the reading programs they had described in their Reading First grant. The other three schools withdrew because the teachers who had agreed to provide RRI in the study requested that they be released from this responsibility. None of these teachers was a full-time reading interventionist, and all had been asked by their administrators to provide the intervention in addition to the duties associated with their primary job assignments.

Students. In each of the 2 years of the study, the research team screened all first-grade students in the participating schools to identify a pool of students who were at risk for serious reading difficulties. Using a multistep screening process researchers assessed students with (a) letter sounds, blending phonemes, and word reading screens from the Texas Primary Reading Inventory (TPRI; Foorman, Fletcher, & Francis, 2004); (b) a word reading task derived from the WJ-III Letter-Word Identification subtest (Woodcock et al., 2001); and (c) a first-grade oral reading fluency passage. In the first year of the study, students were eligible for inclusion if they failed the TPRI screen measures, had a score of 8 or less on the word-reading measure, and had oral reading fluency rates of 8 correct words per minute or less. In Year 2, the cutoffs for word reading and oral reading fluency were lowered to scores of 3 and 5, respectively. To account for potential effects of this change in criteria, we examined the outcome data for cohort effects.

In both years of the study, we excluded students from the sample who received their primary reading instruction in a language other than English, those who had severe mental retardation or emotional disturbance, and those with developmental disabilities such as autism. We included students with mild to moderate disabilities (e.g. LD, speech and language disorder, behavior disorder) who received their primary reading instruction in the general education classroom.

After identifying the pool of eligible at-risk students at each school, we randomly selected 6 to 26 students within each school and randomly assigned them to one of two groups: RRI or typical school practice (TSP). The range in number of students assigned from school to school was due to differing numbers of at-risk first-grade students and available RRI teachers. In all, 507 students were randomized to the study conditions, 222 to RRI and 285 to TSP. During implementation, 40 RRI (18%) and 45 TSP (16%) students left the study; the large majority attended the four schools that withdrew from the study. A chi-square analysis indicated no significant differences in levels of attrition between groups. Within both

TABLE 1

Pretest Word Reading Scores, Students Who Left Compared to Final Sample

	Attrition Sample				*Final Sample*				*Group Difference*
Group	n	M	SD	EM	n	M	SD	EM	(p)
TSP	45	377.28	18.35	377.16	240	381.38	14.65	381.33	.11
RRI	40	383.02	16.16	383.26	182	383.90	13.88	383.72	.86

Note. Word reading scores are W scores from the WJ-III Letter-Word Identification subtest (Woodcock, McGrew, & Mather, 2001). EM = estimated means from multilevel model; TSP = typical school practice; RRI = Responsive Reading Instruction.

the RRI and TSP groups, students who left the study did not differ significantly from those who completed it on pretest word reading ability (see Table 1). After attrition, 182 RRI students and 240 TSP students had complete data and were included in the analysis.

Tests of significant between-group differences in demographic data for the final sample revealed only one: the percentage of males (54%) assigned to the RRI group was lower than that assigned to the TSP group (67%). The mean ages of children in the RRI and TSP groups were 78.80 months (*SD* = 4.4) and 78.96 months (*SD* = 4.2), respectively. According to school report, 104 (43%) of the TSP students received an alternative school-provided reading intervention (not RRI) in addition to their regular first-grade classroom reading instruction. Four students (2%) in the RRI group and 10 (4%) in the TSP group were identified by their teachers as being served in special education (RRI group = 1 LD, 2 speech and language, 1 not reported; TSP group = 2 LD, 4 speech and language, 1 other health impairment, 3 not reported). According to teacher-completed Strengths and Weakness of ADHD Symptoms and Normal Behavior rating scales (Swanson et al., n.d.), the following percentages of students in the RRI and TSP groups exhibited symptoms associated with attention deficit hyperactivity disorder subtypes: combined type: 2.0% and 2.8%, respectively; primarily inattentive type: 1.6% and .4%, respectively; primarily hyperactive/impulsive type: 5.3% and 5.5%, respectively. Table 2 includes other demographic information for the sample.

Intervention Teachers. The principal at each school selected one or more certified teachers, who were currently employed at the school, to provide RRI to at least one group of first-grade students each day, for a total of 40 teachers over the 2-year study. Fourteen teachers joined the study in Year 1. In Year 2, 26 new RRI teachers joined the study. Six teachers participated in both years of the study. The interventionists had a variety of primary job assignments on their campuses, with 31 full-time reading teachers, 4 classroom teachers, 2 librarians, 1 special educator, 1 physical education coach who had never taught reading, and 1 educational assistant (i.e., paraprofessional) with teaching credentials. In the case of those who were not full-time reading teachers, interventionists maintained their current primary job assignments while also providing RRI to one or two groups of children per day. Full-time reading specialists substituted RRI for the first-grade intervention they had formerly implemented.

All but one of the teachers held elementary education teaching credentials, and several were certified in other areas as well, including English as a second language/bilingual education (15), early childhood education (13), reading specialist/master reading teacher (13), special education (12), and secondary education (8). Teachers' total years of experience as professional educators ranged from 2 to 31 years (*M* = 17.15; *SD* = 8.38), and they had from 0 to 21 years of experience as reading interventionists (*M* = 4.93; *SD* = 5.12). Seven teachers had previous training in an explicit phonics-emphasis intervention (i.e., alphabetic phonics, see Cox, 1985; Reading Mastery, e.g., Engelmann & Bruner, 1988; Corrective Reading, e.g., Engelmann, Carnine, & Johnson, 1988; Wilson Reading System, Wilson, 2004), and 15 had previously been trained in the Reading Recovery program (Clay, 1993, 2005), a first-

TABLE 2

Student Demographic Information by Group

	RRI		*TSP*	
Characteristic	n	*%*	n	*%*
Ethnicity[a]				
Caucasian and Asian	76	42	99	41
African American	33	18	32	13
Hispanic	73	40	109	46
Gender[b]				
Male	98	54	160	67
Female	84	46	80	33
Cohort[a]				
Year 1	53	29	61	25
Year 2	129	71	179	75
ESL received[a]	9	7	16	9
Special education received[a]	4	2	10	4

Note. RRI = Responsive Reading Instruction; TSP = typical school practice; ESL = English as a second language instruction.

[a]No statistically significant group differences. [b]$\chi^2(1) = 7.160$, $p = .007$, $\Phi = .130$.

grade reading intervention that requires an intensive 1-year teacher training course.

Description of the Intervention

Intervention was provided using a prepublication version of RRI. RRI consists of five lesson components designed to be implemented within a 40-min lesson. Teachers provide direct, explicit instruction in phonics skills and text-reading strategies, as well as modeling and instructional scaffolding as students apply these skills and strategies while reading and writing connected text. As RRI teachers plan daily lessons, they select from a menu of activities for each of five lesson components based on diagnostic assessments of student strengths and needs. Procedures for each of these activities are described in detail in the RRI handbook (Denton & Hocker, 2006). This study implemented the following RRI components:

Word Work (10 min). Teachers used a variety of instructional activities to provide explicit and systematic instruction and practice in phonemic awareness, letter-sound correspondences, sight word recognition, phonemic decoding, and spelling.

Print Concepts/Fluency and Assessment (10 min). In the earliest lessons, students received explicit instruction in basic print concepts (e.g., left-to-right directionality; print-related vocabulary such as the meanings of the term*s word* and *letter)* as needed. Later, the focus shifted to the development of fluent reading through teacher modeling, repeated oral reading with feedback, and partner reading. This lesson segment also included the individual assessment of one student in the group while the others practiced reading orally individually or in pairs. Teachers assessed each student at least once per week to observe individual strengths and needs and monitor progress over time. Teachers were given a choice of diagnostic curriculum-based assessment tools to use as appropriate (e.g., letter-sound inventory, sight word inventory, oral reading fluency assessment).

Supported Reading (10 min). Teachers provided scaffolding, feedback, and instruction as students read increasingly difficult text. Students were taught to use letter-sound relationships ("sounding out" words) as the primary strategy for identifying unknown words and were discouraged from using pictures and context to identify words. Text was selected at students' instructional reading levels. Teachers in this study used primarily leveled text not intended to be decodable using instructed letters and words, although they included books

specifically designed to provide practice in the application of specific phonics elements. Daily text reading included integrated comprehension instruction: Before reading, teachers provided a brief introduction, setting a purpose for reading with a guiding question focused on a particular comprehension skill (e.g., recalling key facts, identifying the problem in a story). During reading, teachers briefly called students' attention to the guiding question when appropriate, and students and teacher discussed the question after reading. Finally, during the Supported Writing segment of the lesson, students wrote a complete sentence in response to the teacher's question related to the comprehension focus.

Teachers assessed each student at least once per week to observe individual strengths and needs and monitor progress over time.

Supported Writing (10 min). One student in the group composed a complete sentence in response to a question from the teacher, and all students recorded the same sentence. With teacher scaffolding, students applied their knowledge of orthographic patterns and phonemic analysis to spell words and record them accurately. In later lessons, students learned to edit their writing.

Procedures

Our goal was to implement the RRI intervention as in the Mathes et al. (2005) efficacy study, with important differences. In particular, we intended to reduce the influence and control of the research team over the implementation of the intervention in order to evaluate its effectiveness in natural school environments. However, the fact that we were conducting research necessitated some intrusion into school procedures. To clarify the roles of the research team, we contrast implementation in the efficacy and effectiveness studies.

In both studies, we implemented procedures that were essential to the research designs. To assure consistency in identifying student participants, we screened all first-grade students in participating schools and applied preestablished criteria to identify a group of at-risk readers, and we randomly assigned them to intervention condition. We received informed parental consent for study participation. In both studies, we required that all RRI students remain in intervention for 25 to 30 weeks. Finally, in both studies, we collected teacher and student data, although data collection was reduced in the effectiveness study. We collected data on teachers' fidelity of implementation of RRI in both studies. In the efficacy study, teachers received detailed feedback following these observations in order to assure high-fidelity implementation; in the effectiveness study we collected fidelity data to document the intervention's implementation, but did not provide feedback to teachers.

There were also important aspects of intervention implementation that were controlled in the efficacy study, but not in the current effectiveness study. In the efficacy study, Mathes et al. (2005) took measures to assure that the intervention was implemented with high fidelity and adequate intensity. For example, they selected highly qualified teachers whom they provided with large amounts of professional development and coaching support, and they worked closely with schools to assure that the intervention was implemented on a consistent schedule. In the effectiveness study we did not follow this approach. Another key contrast between the efficacy and effectiveness studies was that classroom reading teachers in the Mathes et al. study received professional development designed to support their implementation of Tier 1 classroom instruction, whereas classroom teachers in the effectiveness study did not receive this kind of support. Moreover, the efficacy study took place in a single school district in only six schools that were similar in many regards (e.g., none were Title I-eligible); the effectiveness study was implemented in schools with more diverse characteristics that were located in 16 different school districts. Table 3 illustrates other similarities and differences between implementation in the efficacy study and the current effectiveness study.

Program Manualization. In preparation for the current study, we developed a detailed teacher's manual describing implementation of the RRI program (Denton, 2001), including instructions for using student assessment data to plan lessons and step-by-step descriptions of the procedures for implementing each of the activities

TABLE 3

Comparison of Procedures for Efficacy and Effectiveness Studies

Element	*Efficacy Study*[a]	*Effectiveness Study*
Schools	6 in 1 urban school district; none Title 1-eligible	31 rural, suburban, and urban, 16 school districts; diverse socioeconomic status
Tier 1	Classroom reading instruction enhanced by comprehensive districtwide PD effort; researchers provided 2 days of PD to classroom reading teachers	None
Teachers	3 highly skilled, experienced teachers; recruited, employed, and supervised by researchers; all taught both years of study; all full-time reading teachers	Not controlled by researchers; 40 regular school district teachers with variable experience and expertise; selected, employed, and supervised by schools; variety of primary assignments
PD	42 hours preintervention Year 1, 12 hours in Year 2; monthly half-day sessions throughout the school year	18 hr preintervention; 6 hr for continuing teachers Year 2; 2 follow-up days during school year
Coaching	All received on-site coaching from program developer as often as needed to maintain high implementation fidelity; modeling and feedback with each observation	1/3 received monthly on-site coaching; another 1/3 received some level of computer-based support; No feedback provided after fidelity observations
Students	Researchers screened, identified, and randomly assigned students to treatment or comparison	Researchers screened, identified, and randomly assigned students to treatment or comparison
Materials	Researchers provided all instructional materials, including leveled book library; intervention manual lacked detail but was supplemented heavily by handouts and on-site coaching and modeling	Researchers provided detailed RRI intervention manual, all forms, some instructional materials, video with RRI instructional activities; schools provided other materials including book libraries
Group size	3 to 4 students	Not controlled by researchers; requested 3 to 4 students
Schedule	40 min per day, 5 days per week, 30 weeks; strong agreements with schools to assure intervention was provided regularly	Not controlled by researchers; requested 40 min per day, 5 days per week, 25 weeks; implementation highly variable across schools
Data collection	Frequent observations of interventionists; required teacher record keeping; 5 waves of standardized testing; CBM every 3 weeks; CBM graphs given to intervention and classroom teachers	Less frequent observations; minimal teacher record keeping; 2 waves of standardized testing; CBM every 4 weeks; CBM graphs given to intervention and classroom teachers

Note. PD = professional development; RRI = Responsive Reading Instruction (Denton, 2001; Denton & Hocker, 2006); CBM = curriculum-based measurement for progress monitoring.

[a]Mathes et al., 2005.

in the program. Each participating teacher received a copy of this manual, the RRI handbook (Denton & Hocker, 2006), and a compact disc of video clips demonstrating the implementation of most of the RRI activities.

Teacher Professional Development. Teachers in this effectiveness study received less professional development (PD) and coaching than those in the efficacy study. In general, we attempted to model the kind of support a school might receive if it contracted with a program publisher or developer to provide training and facilitate implementation during the first 1 or 2 years of implementing a new program. In the current study, new teachers participated in approximately 18 hr of PD over 3 days at the beginning of the school year, provided by program developers and one experienced RRI teacher. It focused on (a) the research basis for the program, (b) RRI instructional procedures, and (c) the use of assessment results in lesson planning, and included videotape modeling and mock teaching sessions to practice procedures. In Year 2, continuing teachers received about 6 hr (1 day) of PD at the beginning of the school year. In each of the 2 years, all teachers participated in 12 hr of follow-up PD over 2 days—1 in late fall and another early spring. These sessions were designed to provide more in-depth knowledge of the RRI procedures and to address teachers' needs.

Coaching Support. Two experienced former RRI teachers provided coaching to some of the intervention teachers who participated in this study. One served as a coach over both years of the study; the other was an RRI teacher during Year 1 and a coach during Year 2. Each RRI teacher in the study had been randomly assigned (by school) to receive support in one of three formats: (a) traditional on-site coaching (n = 15), (b) virtual (technology-based) coaching (n = 13), or (c) on-demand coaching (essentially a control condition in which teachers could contact a coach with questions or request visits from a coach but rarely did; n = 12) as part of a separate study of the effects of technology-based coaching. Within each condition, the actual participation of teachers in coaching varied and in no condition did it approximate the level of coaching received by teachers in the Mathes et al. (2005) study. An analysis of the effects of coaching condition on outcomes for treatment group students indicated only small differences in phonemic awareness (Denton, Mathes, & Shih, 2007). Thus, teachers in the three coaching conditions were combined for analyses in this study.

Alternate Interventions Provided to TSP Students. Forty-three percent (n = 104) of the students assigned to the TSP group received a school-provided alternative intervention as part of their schools' typical practice. These interventions varied in focus, intensity, and mode of delivery. In some of the shortest and least intensive interventions students read a book silently and answered a short set of comprehension questions during 20-min sessions. Other comparison students participated in a computer-based 20-min early reading skills intervention led by a teaching assistant. On the opposite extreme, more intensive interventions employed the use of one-on-one skills-based instruction for 40-min time periods. Most TSP interventions were more balanced, with students in small groups receiving mainly skills-based instruction.

Fidelity of Implementation and Intervention Dosage. To evaluate the degree that the RRI intervention was provided as designed, two RRI coaches and one project coordinator observed intervention teachers and students for one entire teaching session approximately once every 9 weeks for a total of three fidelity observations per teacher per year using a fidelity measure closely aligned with the intervention. The coaches did not observe their own teachers. Teachers did not receive feedback following these fidelity observations. Interobserver reliability was established by co-viewing and independently rating lessons in schools, a procedure repeated with each cycle of data collection. Training of observers continued until they reached at least 80% agreement with the program developers.

As described previously, the RRI lesson is divided into five components designed to be delivered during four 10-min segments. Fidelity was rated on a 4-point Likert scale from 0 to 3 (expected but not observed, low, medium, high) for each activity observed within these 5 components; for each component, observers rated items addressing the domains of (a) adherence to program procedures, (b) implementation quality, and (c) students' on-task behaviors during the intervention. Observers also evaluated global quality indi-

cators (e.g., students are seated appropriately; teacher provides appropriate corrective feedback) using a yes/no choice format for the overall lesson.

On average, teachers implemented the RRI intervention with adequate fidelity, although there was variation across teachers. The mean overall fidelity rating for all RRI teachers across all observations was 80.14% (*SD* = 8.43%; range, 57.33%–94.58%). We also calculated a mean rating across all observations for each teacher for each of the three domains listed above. On the 0-to-3 scale (with 0 meaning the element was expected but not observed and 3 being the highest rating), the mean ratings were 2.08 (*SD* = .47) for adherence to RRI procedures; 2.26 (*SD* = .38) for quality of implementation; and 2.61 (*SD* = .22) for students' on-task behavior.

Dosage of the intervention received by students was calculated for Year 2 of the study using student attendance data recorded by intervention teachers. Teachers recorded the actual numbers of lessons provided to each student in the RRI group, noting both student absences and days on which intervention was not provided by the teacher for some reason. Year 1 attendance data were not deemed reliable as many teachers did not record or submit these data during the first semester of the study, so dosage analysis was possible only for the second study cohort. When schools committed to participation in the study, they agreed to implement the RRI intervention daily for 40 min for at least 25 weeks, which would result in a total of approximately 75 hr of instruction, allowing for holidays and so forth. According to teachers' attendance records, Cohort 2 students actually received a mean of 62.16 hr of instruction (*SD* = 8.08), and dosage ranged from 42.67 hr to 75.33 hr across schools.

Other Instructional Observations. Two times during each school year both the RRI intervention teachers and the teachers who provided alternate reading intervention to TSP students were observed using the Instructional Content Emphasis-Revised (ICE-R; Edmonds & Briggs, 2003). The ICE-R was used to systematically categorize and code the content of reading and language arts instruction. It produces data related to several dimensions of instruction, including the amount of time allocated to instructional components (e.g., phonics and word work) and subcomponents (e.g., letter-sound instruction). Interrater reliability on the instrument was obtained using videotaped samples, where observers coded the observation independently and attained at least 80% agreement with the ICE-R trainer.

Alternative intervention provided to TSP students was observed for a total of 1,914 min over 63 observations spanning the 2-year study duration. These alternative instruction periods lasted, on average, 30.4 min. RRI groups were observed with the ICE-R (Edmonds & Briggs, 2003) for a total of 2,114 min over 50 observations over 2 years. RRI periods lasted, on average, 42.3 min. Table 4 summarizes the proportion of lesson time spent in various instructional emphases and text reading across all observations in the TSP alternative intervention sessions and during RRI lessons. Most notably, observations indicated that teachers who provided alternative intervention to TSP students spent a larger proportion of their lesson time instructing students in phonological awareness and phonics skills than did teachers in the RRI condition, but RRI students spent relatively more time reading and writing connected text.

Student Measures

Because the early development of reading skills demands increasing proficiency in word reading, phonemic decoding, reading fluency, and reading comprehension, these domains are routinely measured in reading intervention research. In this study, we administered a comprehensive battery that assessed each of these aspects of reading, along with spelling and phonological awareness, an important predictor of early reading development. These were the same domains measured in the efficacy study.

Pretest and Posttest Measures. At pretest and posttest, we measured phonological awareness using the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) Blending Words, Blending Nonwords, and Segmenting Words subtests. We assessed word identification (timed and untimed) with the Test of Word Reading Efficiency Sight Word Efficiency subtest (TOWRE; Torgesen, Wagner, & Rashotte, 1999) and the WJ-III

TABLE 4

Observed Percentages of Instructional Time Engaged in Various Instructional Emphases and in Text Reading

Activities	*RRI*	*TSP*
Instructional emphases		
Phonological awareness	2.1%	5.8%
Letter knowledge	0.1%	1.7%
Word study/phonics	22.9%	41.7%
Spelling	0.5%	1.3%
Oral language development	0.0%	0.1%
Fluency	30.4%	13.5%
Comprehension	13.2%	10.0%
Writing or language arts	15.2%	8.1%
Text reading across all instructional emphases		
Teacher text reading	5.4%	5.2%
Student text reading	42.3%	25.2%

Note. RRI = Responsive Reading Instruction (Denton, 2001; Denton & Hocker, 2006) intervention; TSP = typical school practice intervention.

Letter-Word Identification subtest (Woodcock et al., 2001), as well as a researcher-developed word list. Phonemic decoding (timed and untimed) was measured with the TOWRE Phonemic Decoding Efficiency subtest and the WJ-III Word Attack subtest. We assessed reading comprehension using WJ-III Passage Comprehension, and spelling with WJ-III Spelling. Here we do not report on the CTOPP Blending Nonwords subtest or researcher-developed word list to reduce redundancy and multiple contrasts in the analyses.

In CTOPP Blending Words (Wagner et al., 1999) students are orally presented with word or nonword parts and are asked to put the parts together to make a real word (e.g., t-r-a-p: trap); in Segmenting Words, students are orally given a word and asked to say the word one sound at a time (e.g., say the sounds in stamp: s-t-a-m-p). Most average internal consistency and alternate forms reliability coefficients exceed .80. The test/retest (time sampling) coefficients range from .70 to .92.

The WJ-III (Woodcock et al., 2001) is a nationally standardized, individually administered battery of achievement tests. Letter-Word Identification evaluates a student's ability to read individual words of increasing difficulty in isolation, and Word Attack examines the ability to apply phonics and structural analysis skills in order to read nonsense words, simulating the real-life task of decoding unknown words. The Basic Reading composite, composed of these two subtest scores, is widely used as a norm referenced indicator of RTI. It has reliability (split-half) of .91 to .97 in the age range of interest. Passage Comprehension is a measure of reading comprehension at the sentence level. The participant reads a sentence or short passage and fills in missing words. This subtest is a frequent outcome measure in reading intervention studies and has split-half reliability of .83 to .96 in the age range of the study. In WJ-III Spelling, students hear a word and write the correct spelling; reliability (split-half) is .87 to .91 in the age range of interest.

The TOWRE (Torgesen et al., 1999) is an individually administered test of speeded word reading commonly used as an outcome in reading intervention studies. Subtests assess word reading efficiency and phonemic decoding efficiency, reflecting speeded reading of real words, and pseudowords. The participant reads a list of 104 words or nonwords as fast as possible. The number read correctly within 45 sec is recorded. Test-retest reliabilities range from .83 to .92 for each subtest in the age range of interest.

Continuous Progress Monitoring. We monitored students' development of fluent reading of connected text every 4 weeks using the Comprehensive Monitoring of Early Reading Skills (Mathes & Torgesen, 2008) Oral Reading Fluency (ORF) measure. The measure consists of a set of graded text passages of approximately equivalent difficulty, written specifically to conform to grade-level conventions in terms of word and sentence length and complexity. At each data collection point, students read a first-grade end-of-year passage for 1 min. The total number of words read, errors, and correct words read in one minute were recorded. ORF assessment is a type of curriculum-based measurement (CBM) commonly used to assess progress in reading development. ORF CBM has been found to have high validity and reliability for measurement of student reading progress, including high concurrent validity with standardized measures of reading comprehension for elementary school students (L.

Fuchs, Fuchs, Hosp, & Jenkins, 2001; Marston, 1989). In the current study, research assistants traveled to the schools to individually administer ORF assessments. They were trained to proficiency and established interobserver agreement of at least 90% prior to data collection.

Data Analysis

To address the first research question, group differences were modeled with multilevel techniques, where students were nested within schools and measurement occasions were nested within students in schools as appropriate. With multilevel techniques, variables are entered into models as either fixed or random effects. For example, nesting students within schools and defining the school-level intercept term as a random effect allows variance in school-level intercepts to be modeled, which in turn estimates the between-school variance in student performance. Variances for variables that are defined as fixed effects are assumed to be zero.

Following the step-up modeling strategy (West, Welch, & Galecki, 2007), examination of unconditional models indicated that the data did not support modeling students within classes within schools. Across the majority of the study's variables, including a random intercept for classes within schools, yielded ill-fitting models with nonpositive definite matrices. In cases where random class effects could be modeled, results from likelihood ratio tests indicated that the intercept for classes nested within schools could be omitted. Similarly, examination of unconditional models identified no need to model cluster effects of intervention students within tutors or tutor groups.

For pretest-posttest measures, differences between groups were estimated by analyzing two-level hierarchical linear models, where students were nested within schools. For the ORF continuous progress monitoring measure, differences between groups were estimated by analyzing three-level linear growth models where measurement occasions were nested within students in schools. By centering the intercept at the last measurement occasion, we were able to estimate the extent to which groups differed at the last phase of assessment as well as the shape and rate of growth.

We included student-level variables in the analyses. In all analyses, Group (RRI or TSP) served as a student-level factor. For all measures, Cohort was considered a potential covariate because the study was conducted over 2 years with two different cohorts of students. This was particularly important given the different eligibility criteria applied to the two cohorts in the study. In order to compare this study's results to Mathes et al. (2005), we considered Ethnicity and Gender as potential covariates and only included covariates in the final model if they were statistically significant. Covariates were grand mean centered.

We tested each student measurement outcome separately because a univariate approach was consistent with our interest in comparing our findings to those in prior reading intervention studies. In keeping with Mathes et al. (2005), we set alpha at .05. However, we report exact probabilities so that the reader can draw conclusions about statistical significance. With the exception of the WJ-III (Woodcock et al., 2001) measures, analyses were conducted on raw scores. For the WJ-III measures, analyses were conducted on W scores, continuously scaled standard scores available for the WJ-III.

We also evaluated group differences by examining effect sizes. Following Mathes et al. (2005) and Tymms (2004), effect sizes for Group differences on posttest measures (Δ) were calculated by dividing the group coefficient by the square root of the residual. Group difference effect sizes in polynomial trends (δ) were calculated by dividing the group difference of the polynomial trend by the square root of the residual (Mathes et al.). Computations for effect sizes of between-school variance followed Taylor, Pearson, Peterson, and Rodriguez's (2005) example and were computed as conditional Intraclass Correlation Coefficients (ρ), where ρ = between-school variance/total variance.

To address the second research question, we calculated the percentage of students assigned to each condition who met the preestablished benchmark of a grade-based standard score of 93 (30th percentile) or more on the WJ-III Basic Skills composite (Woodcock et al., 2001). This benchmark has been used in reading intervention

studies to denote adequate RTI and represents the theory that development of automaticity in word identification in the early stages of reading acquisition leads to fluent reading, which in turn facilitates comprehension. Most important, it was the benchmark used in the Mathes et al.'s efficacy study (2005), and we were interested in comparing the response rates of children in the two studies. We also evaluated the percentage of students in each condition who met a year-end ORF benchmark of 35 words correct per minute (wcpm), the 30th percentile in the spring of first grade for DIBELS ORF passages (Good, Wallin, Simmons, Kame'enui, & Kaminski, 2002) and the benchmark used by Denton and Mathes (2003) in their discussion of the Mathes et al. (2005) study.

RESULTS

Preanalyses

We conducted preanalyses to examine the initial equivalence of group differences on all reading measures. School-level intercepts were allowed to vary and were included as random-effects for all measures. Group was included as a fixed-effect factor for all variables. Results indicated that there were no statistically significant group differences in students' initial performance on any of the study's baseline measures. Effect sizes for pretest group and between-school differences, estimated means resulting from the multilevel models, and descriptive statistics for all measures are reported in Table 5.

End-of-Year Student Achievement

Table 6 outlines the results of the final models used to estimate group differences at posttest. For all measures, school-level intercepts were allowed to vary and included as random effects. Group was included as a fixed-effect factor for all variables. Cohort was included as a fixed-effect covariate in final model estimates for measures other than CTOPP Blending Words and Segmenting Words (Wagner et al., 1999), as well as WJ III Passage Comprehension (Woodcock et al., 2001). Ethnicity was included as a fixed-effect covariate in final model estimates for WJ-III Passage Comprehension and CTOPP Segmenting Words. Gender was included as a fixed-effect covariate in the final model estimate for WJ-III Spelling. There were no instances where any covariate interacted with Group.

Results indicated statistically and practically significant group differences favoring the treatment group for all measures other than CTOPP Segmenting Words (Wagner et al., 1999). Table 7 reports effect sizes for group and school-level differences, estimated means resulting from the multilevel models, and descriptive statistics for all posttest measures.

Continuous Progress Monitoring Measures

Table 8 reports the results of the final models used to estimate growth for the ORF continuous progress monitoring measure. ORF was best modeled with a linear trend with random intercept and slope terms at the student level and a random intercept term at the school level. In addition to the student-level grouping factor, the final model included a fixed-effect covariate for Cohort. The final model indicates that Group was a statistically and practically significant factor for both the student-level intercept (Δ = 1.48) and slope, or students' rate of growth over time (δ = .41), indicating that the treatment group had a significantly higher end-of-year status and rate of growth in ORF than the TSP group. Comparing the amount of variance in the school-level intercept to the total variance indicates that the final model explained almost all between-school variation in ORF.

Response to Intervention

Because it was the benchmark applied by Mathes et al. (2005), we used a standard score performance of 93 or above on the posttest WJ-III Basic Reading Skills composite score (Woodcock et al., 2001) as a benchmark to determine students' adequate RTI. An examination of posttest results showed that 91% of students in the treatment group met this benchmark, whereas 79% in the TSP group demonstrated adequate response in word reading skills. RTI was considerably weaker when measured using the ORF benchmark of 35

TABLE 5

Baseline Observed Score Means and Standard Deviations, Estimated Means, and Effect Sizes

	RRI (n = *182)*			*TSP* (n = *240)*			*Group*	*School Effects*
Measure	M	SD	EM[a]	M	SD	EM[a]	*Effects*[b]	(ρ)
CTOPP Blending Words	8.35	3.73	8.14	8.33	3.69	8.10	.01	.19
CTOPP Segmenting Words	4.39	3.65	4.35	4.25	3.85	4.18	.05	.13
TOWRE Sight Word Efficiency	5.23	4.29	5.20	4.62	4.16	4.60	.14	.02
TOWRE Phonemic Decoding	1.94	2.87	1.91	1.85	2.76	1.82	.04	.05
WJ-III Letter Word ID W	383.90	13.89	383.66	381.38	14.66	381.15	.18	.04
WJ-III Letter Word ID SS	94.62	9.44		93.15	8.95			
WJ-III Word Attack W	434.08	21.24	434.02	431.95	21.33	431.77	.11	.07
WJ-III Word Attack SS	98.92	14.22		97.33	13.36			
WJ-III Passage Comp W	409.77	17.72	409.90	409.11	16.98	409.21	.04	.03
WJ-III Passage Comp SS	81.81	11.29		81.04	10.95			
WJ-III Spelling W	426.79	12.36	426.87	425.59	12.93	425.66	.10	.03
WJ-III Spelling SS	97.31	9.37		96.28	9.23			
Oral reading fluency	6.92	4.75	6.91	5.70	3.88	5.72	.29	.08

Note. RRI = Responsive Reading Instruction (Denton, 2001; Denton & Hocker, 2006); TSP = typical school practice; CTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999); TOWRE = Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999); WJ-III = Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001); SS = WJ-III standard scores; W = WJ-III W scores. Raw scores reported except for WJ-III Scores.

[a]Estimated means from multilevel model. [b]Standardized mean differences.

wcpm: 48% of the RRI students and 28% of the TSP students met this benchmark.

DISCUSSION

The purpose of this study was to determine whether a first-grade reading intervention that had demonstrated efficacy in a previous study would be associated with positive student outcomes when implemented in multiple schools with reduced researcher influence. Specifically, we sought to evaluate the effectiveness and generalizability of Responseive Reading Instruction (RRI; Denton, 2001; Denton & Hocker, 2006), when implemented in 31 schools by 40 teachers with various levels of previous training and experience. The research questions addressed the comparison of student outcomes between a group of at-risk readers who received the RRI intervention (RRI group) and a group of at-risk readers who received the instruction and interventions typically provided by their schools (TSP group) on reading and reading-related measures and the percentage of students in each group who demonstrated adequate RTI.

Comparison of Outcomes for the RRI and TSP Groups

The RRI group demonstrated significantly higher outcomes than the TSP group on timed and untimed word reading, timed and untimed phonemic decoding, spelling, ORF, and reading comprehension. Results were mixed for measures of phonological awareness. These effects were evident even though 43% of the students in the TSP group received an alternate supplemental reading intervention provided by their schools. Effect sizes were mostly in the moderate range and have practical significance in light of the persistent gaps between able and less able readers in our schools (e.g., Lee, Grigg, & Donahue, 2007).

The fact that the one variable for which the RRI students in this study did not have significantly higher performance than the TSP comparison students was a measure of phonemic awareness may reflect the fact that teachers who

TABLE 6

End-of-Year Observed Score Means and Standard Deviations, Estimated Means, and Effect Sizes

Measure	*RRI* (n = *182*)			*TSP* (n = *240*)			*Group Effects*[d]	*School Effects* (ρ)
	M	SD	EM[a]	M	SD	EM[a]		
CTOPP Blending Words	14.21	3.04	14.27	13.43	3.27	13.43	.27	.13
CTOPP Segmenting Words	9.69	2.83	9.70[b]	9.33	3.61	9.31[b]	.12	.06
TOWRE Sight Word Efficiency	27.92	10.67	27.86[a]	23.28	10.34	23.21[a]	.47	.11
TOWRE Sight Word Efficiency SS	95.19	9.75		91.29	10.19			
TOWRE Phonemic Decoding	10.63	6.21	10.59[a]	8.14	5.64	8.09[a]	.44	.05
TOWRE Phonemic Decoding SS	95.62	9.00		92.11	9.32			
WJ III Letter Word ID W	438.08	18.04	438.00[a]	424.58	20.72	424.40[a]	.72	.07
WJ III Letter Word ID SS	104.58	10.67		98.75	11.10			
WJ III Word Attack W	473.33	17.54	473.32[a]	465.38	18.96	465.15[a]	.46	.06
WJ III Word Attack SS	108.25	11.76		103.63	11.95			
WJ III Passage Comp W	455.78	14.06	455.61[b]	447.34	17.23	447.40[b]	.53	.06
WJ III Passage Comp SS	96.53	9.97		91.43	10.90			
WJ III Spelling W	457.96	12.43	457.80[a,c]	449.72	14.63	449.90[a,c]	.63	.07
WJ III Spelling SS	105.60	11.18		100.15	11.52			
Oral reading fluency	31.35	18.68	32.01[a]	25.03	17.12	24.71[a]	.45	<.01

Note. RRI = Responsive Reading Instruction (Denton, 2001; Denton & Hocker, 2006); TSP = typical school practice; EM = estimated means from multilevel model; CTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999); TOWRE = Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999); WJ-III = Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001). Raw scores reported except when otherwise noted. SS = standard scores; W = WJ III W scores.
[a]Covaried by cohort. [b]Covaried by ethnicity. [c]Covaried by gender. [d]Standardized Mean Differences.

provided alternate Tier 2 interventions as part as typical school practice (received by about 40% of the total TSP group) spent a larger portion of their time engaged in phonics and phonemic awareness instruction than RRI teachers. However, despite this attention to phonics skills instruction, the RRI students outperformed the TSP students in word reading and phonemic decoding. Although RRI teachers spent less time teaching phonics skills than TSP intervention teachers, their students spent more lesson time engaged in text reading than those of the TSP teachers. The format of supported text reading may have promoted generalization and internalization of the phonics skills the RRI students were learning. Each day, one RRI student from each group read a section of unfamiliar text alone while the teacher provided explicit modeling, prompting, and scaffolding to encourage the student to apply strategies and skills while reading connected text.

The comparatively stronger performance of RRI students in reading comprehension may also be related to this text reading format. In RRI, daily comprehension instruction was integrated into text reading. As the RRI students read a section of unfamiliar text, teachers (a) provided a daily comprehension focus by presenting a single guiding question before reading, (b) discussed the question during and after reading, and (c) supported students as they composed and wrote a complete sentence responding to the question in their journals. In addition, as RRI students were reading, teachers directly taught and prompted them to self-monitor their comprehension and self-correct errors when the text did not make sense. Consistently following this routine every day over the extended intervention period appears

TABLE 7

Difference in End-of-Year Achievement

	CTOPP Blending Words			*CTOPP Segmenting Words*			*TOWRE Sight Word Efficiency*			*TOWRE Phonemic Decoding Efficiency*		
Effects	*Estimate*	SE	p	*Estimate*	SE	p	*Estimate*	SE	p	*Estimate*	SE	p
Fixed effects												
Intercept	13.43	.28	<.001	9.31	.25	<.001	23.21	.91	<.001	8.09	.44	<.001
Group (RRI)	.80	.29	0.007	.39	.31	0.218	4.65	.98	<.001	2.5	.56	<.001
Cohort (2)							2.77	1.28	0.032	1.58	.69	0.025
Ethnicity (AA)				–1.20	.49	0.014						
Ethnicity (Hispanic)				–.57	.11	0.107						
Random effects												
Residual	8.88	.64	<.001	10.04	.72	<.001	97.65	7.00	<.001			
Intercept[a]	1.27	.53	0.016	.60	.40	0.125	11.81	5.29	0.026	32.78	2.36	<.001

	WJ III Letter Word Identification			*WJ III Word Attack*			*WJ III Passage Comprehension*			*WJ III Spelling*		
Effects	*Estimate*	SE	p	*Estimate*	SE	p	*Estimate*	SE	p	*Estimate*	SE	p
Fixed effects												
Intercept	424.40	1.56	<.001	465.10	1.43	<.001	447.40	1.22	<.001	449.90	1.08	<.001
Group (RRI)	13.59	1.86	<.001	8.08	1.75	<.001	8.18	1.52	<.001	7.91	1.31	<.001
Cohort (2)	5.77	2.36	0.015	5.05	2.20	0.023				3.58	1.66	0.032
Ethnicity (AA)							–.35	2.36	0.883			
Ethnicity (Hispanic)							–4.05	1.72	0.019			
Gender (Female)										3.14	1.36	.021
Random effects												
Residual	353.01	25.32	<.001	314.50	22.74	<.001	236.16	16.89	<.001	171.49	12.35	<.001
Intercept[a]	26.76	14.94	0.073	20.41	13.59	0.133	13.83	8.52	0.104	12.33	7.43	.097

Note. CTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999); TOWRE = Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999); RRI = Responsive Reading Instruction (Denton, 2001; Denton & Hocker, 2006); AA = African American;WJ-III = Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001).

[a]Between-school variance. Cohort, Gender, and Ethnicity were only included as covariates if the main effect was statistically significant.

to have supported the development of reading comprehension in the RRI students.

Snow et al. (1998) emphasized the importance of providing both explicit instruction in the alphabetic principle and extended opportunities to read and write connected text. These recommendations have been echoed in subsequent research syntheses and meta-analyses (e.g., Ehri, 2003; National Reading Panel, 2000; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). In general, the results of the current study are aligned with this conclusion.

Response to Intervention

We had expected that a smaller percentage of students would meet the benchmark for adequate RTI in the current scaling-up study than in the efficacy trial (Mathes et al., 2005). Our findings related to this hypothesis were mixed. When considering the criteria of performance at or above the 30th percentile on the WJ-III Basic Reading Composite (Woodcock et al., 2001), intervention response rates in the two studies were comparable: 91% of the RRI students in the current study met this benchmark, whereas Mathes et al. reported that 93% of the students who received RRI in addition to enhanced classroom reading instruction did so.

However, when applying the ORF benchmark of 35 wcpm at the end of Grade 1, 77% of RRI students in Mathes et al. (2005) would be considered adequate responders (Denton & Mathes, 2003), whereas only 48% of the RRI students in the current study met this benchmark. A comparison of group means in the current study to ORF norms confirmed that, although RRI students in this study performed significantly better in ORF than TSP students, neither group's mean score was at or above average. The RRI group's year-end mean was 31 wcpm, and the TSP group mean was 25 wcpm. National norms place the year-end mean fluency rate for Grade 1 at 53 to 54 wcpm (Good et al., 2002; Hasbrouck & Tindal, 2006). According to the Good et al. norms, the RRI group performed, on average, at the 25th percentile, whereas the TSP group mean was at the 18th percentile. Interestingly, posttest standard score means for the RRI group on the TOWRE (Torgesen et al., 1999) word and nonword reading fluency measures (both about 95) indicated that the treatment students performed in the average range in these domains.

TABLE 8

Growth Model Analyses for Oral Reading Fluency Data

Effects	*Estimate*	SE	p
Fixed effects			
Intercept	24.69	1.06	<.001
Group (RRI)	7.34	1.58	<.001
Slope	–3.21	.16	<.001
Slope × Group	–.96	.25	<.001
Slope × Cohort (2)	1.19	.08	<.001
Random effects			
School intercept	1.18		
Student intercept	247.97		
Student slope	5.58		
Student residual	24.39		

Note. RRI = Responsive Reading Instruction (Denton, 2001; Denton & Hocker, 2006).

Comparison of Outcomes to the Efficacy Study

Contrary to our expectations, results indicated more positive outcomes on several measures in the current study than those reported for students who received RRI in the Mathes et al. (2005) efficacy study. Mathes and colleagues reported that RRI students had statistically significantly higher outcomes relative to comparison group students on tests of phonological awareness, timed and untimed word reading, ORF (for one cohort), and spelling, but not in phonemic decoding or reading comprehension. RRI students in the current study demonstrated significantly higher scores than comparison students in all these domains except on one phonological awareness subtest.

Effect Sizes. We compared effect sizes reported in the two studies on the WJ-III (Woodcock et al., 2001) and TOWRE (Torgesen et al., 1999) subtests and on ORF. In general, effects associated with RRI (Denton, 2001; Denton & Hocker, 2006) in the current study were more robust than in the efficacy study on untimed measures and on phonemic decoding fluency. In

contrast, the efficacy study was associated with stronger effects in oral reading fluency measured in text passages and word lists. For WJ-III Letter-Word Identification, Word Attack, Passage Comprehension, and Spelling, Mathes et al. (2005) reported effect sizes of .36, .28, .30, and .55, respectively; for the current study, Table 4 reports effect sizes for the same measures as .72, .46, .53, and .63, respectively. Mathes et al. administered the TOWRE four times over the school year to measure growth. Based on growth curve models, they reported effect sizes in Phonemic Decoding Efficiency of .23 for the intercept (i.e., predicted April score) and 1.41 for the Sight Word Efficiency intercept. We administered the TOWRE only at pre- and posttest and found effects sizes of .44 and .47, respectively, on these subtests. Effect sizes for ORF were calculated on the intercept term in both studies. Mathes et al. reported ORF outcomes separately for their two cohorts. Although between-group differences in ORF were statistically significant only for Cohort 2 in that study, large effect sizes of 1.06 and 1.81 were reported for Cohorts 1 and 2, respectively. In the current study, ORF effects were moderate at .45 (for our two cohorts combined). We were unable to directly compare phonological awareness outcomes in the two studies, as Mathes et al. measured this domain using a prepublication version of the CTOPP (Wagner et al., 1999), administered different subtests, and reported phonological awareness outcomes for CTOPP subtests on a single Item Response Theory (IRT)-based scale.

The Nature of Typical School Practice. In both studies, RRI effect sizes represented comparisons with typical school practice groups. One plausible explanation for the stronger effects on several outcomes in the current study is a difference between the content and quality of typical school instruction provided in the two studies. This hypothesis is supported by the fact that 84% of the comparison group in Mathes et al. (2005) demonstrated adequate intervention response based on the decoding benchmark, whereas only 79% of the TSP students in the current study met this goal. In the Mathes et al. study, all six schools, which were in the same school district, were engaged in a comprehensive districtwide reading improvement effort, through which they were provided with substantial PD. Mathes et al. also provided classroom reading teachers with additional PD focused on using assessment results to guide instruction and on peer tutoring. Thus, in Mathes et al. the effect sizes reflect the added value of the supplemental intervention provided along with enhanced classroom instruction. We did not provide support to the classroom reading teachers in the current study, nor were we able to conduct observations that would enable us to describe this instruction in the many first-grade classrooms across the 31 participating schools. However, it is reasonable to assume that there was variation in the quality and content of classroom reading instruction for 40 teachers in different schools across 16 school districts.

Oral Reading Fluency Outcomes

In the current study, RRI had more robust effects than were detected in the efficacy study in several important reading domains, but effect sizes for ORF were considerably smaller than those reported by Mathes et al. (2005). Similarly, only a moderate percentage of RRI students in the current study met a commonly applied ORF passage benchmark at the end of first grade, whereas more RRI students attained this benchmark in the efficacy study. We speculate that RRI teachers in the current study may have progressed through the program at a slower rate than those in the more controlled effectiveness study, so that students in the current study may have had less time to practice reading grade-level text to consolidate skills and develop automaticity. Other researchers have similarly found that reading accuracy is more easily remediated than reading fluency, although this finding is more common for students above Grade 1. For example, Torgesen et al. (2001) provided reading intervention for 2 hr daily over 8 weeks to students served in special education in Grades 3 to 5 who had seriously impaired word recognition. The intervention resulted in remarkable growth in word recognition and comprehension with effects sustained over time, and about 40% of the students exited special education following the study. Surprisingly, students' standard scores in fluency were virtually unchanged. The researchers suggested that this difficulty in remediating fluency is related to the fact that the students had had little exposure to many words in

grade-level reading passages, and so they were unable to recognize these words automatically "at sight."

Alternative Typical Practice Interventions

The results of this study may indicate that RRI was more effective than the alternate school-provided interventions received by some of the TSP students, but it is not possible to verify this using our design. We can only conclude that RRI was more effective than the instruction these schools typically provided their first-grade at-risk readers. Similarly, our research design did not allow for a direct comparison of student outcomes for RRI and the subgroups of TSP students who did and did not receive alternative intervention. At some schools, none of the TSP students received school-provided intervention. At others, some students from this comparison group received an alternative intervention whereas others did not. Still other schools provided all of the TSP students with alternative interventions. As we explored our data, we calculated pre-post descriptive statistics on observed scores for the two TSP subgroups and found that the TSP subgroup that did not receive school-provided intervention (n = 136) had somewhat higher pretest and posttest group means on some measures than those who did receive alternative intervention (n = 104). It appears that schools tended to provide intervention to TSP students with the greatest need. However, the net raw or standard score mean gains made in the two TSP subgroups—those who did and did not receive school-provided intervention—were nearly identical, indicating that these interventions may not have accelerated the reading progress of students who received them.

Implications for Practice

This study has demonstrated that an early reading intervention that had been validated in a previous efficacy study in a limited number of schools is associated with positive student outcomes when implemented in diverse school contexts. It is encouraging that, even in schools varying in setting (urban vs. rural vs. suburban), socioeconomic status, and teachers' levels of experience and training, 91% of at-risk readers whose teachers implemented the intervention were able to adequately read and spell words at the end of first grade. Although this study did not investigate the intervention within whole-school RTI implementations, our findings lend support to the feasibility of RTI reading models. If early reading intervention, even when implementation varies somewhat due to factors such as the inevitable variations in teacher skill and school resources, can produce positive outcomes, there is evidence that RTI models can ultimately be validly implemented in a variety of contexts.

Although the study indicated strong student intervention response for RRI students in terms of word-level skills, response was appreciably weaker when measured with an ORF benchmark. More research is needed to determine how RTI should be validly determined. The development of basic decoding and spelling skills is a primary goal of first-grade reading instruction, and students who lack the ability to read accurately will undoubtedly struggle in second grade. It is clear that adequate decoding skills are necessary, but not sufficient, for extracting meaning from text. However, fluent reading has been shown to be highly predictive of reading comprehension in elementary aged students (e.g., L. Fuchs et al., 2001) and is also an important goal of early reading instruction. The application of differing RTI criteria identifies largely different populations of students as having adequate intervention response (Barth et al., 2008; Denton et al., 2009), and there is insufficient existing research about the implications of adopting these differing criteria. This is an important question that can have far-reaching effects on the application of RTI criteria in the identification of LD.

A final implication of the current study is that simply providing at-risk readers with more time in reading instruction may be insufficient. Rather, it appears to matter what teachers and students actually do during those supplemental reading lessons. The research intervention had more robust outcomes than typical school instruction despite the fact that a large percentage of TSP group students received alternate forms of Tier 2 intervention through their schools. Observations of instructional time use in the RRI intervention groups indicated that RRI students

received explicit phonemic awareness and phonics instruction but also spent proportionally more time engaged in reading connected text than TSP intervention students. The relatively stronger outcomes of the RRI group may be related to this aspect of intervention design.

Study Limitations

Our results should be interpreted in light of several limitations. First, although researcher control over intervention implementation was weaker in this study than is typical of efficacy studies, the fact that we were conducting a randomized field trial meant that we did intrude on typical practices within these schools to a degree. Most notably, we selected the intervention students and requested that they be maintained in intervention for 25 weeks, so that we could compare outcomes in the two research conditions and also provide valid comparisons to the outcomes of the efficacy study, in which students received intervention for 30 weeks. However, our current study does not indicate the likely outcomes if RRI were implemented for briefer periods. In addition, the simple fact that we collected data in these schools may have affected outcomes. For example, the fact that we observed teachers to document fidelity of implementation may have motivated them to attend more closely to the program manual, and the fact that we provided intervention and classroom teachers with graphs of students' ORF progress may have affected their instruction of students who struggled. This effect should have been minimized because classroom teachers received these graphs for students in the TSP group as well as those in the RRI group, but these schools did not typically collect continuous progress monitoring data. The graphs were provided as a service to the schools because we were not serving the TSP students in our intervention.

Second, although the study was implemented in schools with varying contexts, the sample included relatively few urban schools with large percentages of economically disadvantaged students. It may be that these schools will present the greatest challenge to scaling up research-supported interventions. More research specifically examining implementation of interventions in such schools is needed.

Third, our application of less stringent criteria for at-risk status in the first year of the study may have identified some students who were "false positives," students who appeared to be at risk but who would have learned to read adequately without intervention. We accounted for between-cohort differences by including Cohort as a fixed-effect covariate in final model estimates when this improved model fit. However, universally applying the more stringent criteria may have produced somewhat different results.

CONCLUSION

Recent legislative and policy initiatives have placed heavy emphasis on the adoption by schools of approaches and curricula for reading instruction that have evidence of effectiveness from scientific reading research. This study provides evidence that an early reading intervention that had previously been found to have positive effects for at-risk first-grade students in a smaller controlled study can be effectively scaled up in multiple school contexts. If the research-to-practice gap is to be bridged, it will be necessary to conduct more studies of this kind with various interventions that have demonstrated efficacy in controlled research environments.

REFERENCES

Barth, A., Stuebing, K., Anthony, J., Denton, C., Mathes, P., . . . Francis, D. J. (2008). Agreement among response to intervention criteria for identifying responder status. *Learning & Individual Differences, 18*, 296–307.

Batsche, G., Elliott, J., Graden, J. L., Grimes, J., Kovaleski, J. F., . . . Tilly III, W. D. (2006). *Response to intervention: Policy considerations and implementation*. Alexandria, VA: National Association of State Directors of Special Education.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, . . . Shaywitz, S. (2004). Effects of intensive reading remediation for second and third graders. *Journal of Educational Psychology. 96*, 444–461.

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

Clay, M. (2005). *Literacy lessons designed for individuals: Part two: Teaching procedures*. Portsmouth, NH: Heinemann.

Council for Exceptional Children. (October, 2007). *CEC's position on response to intervention (RTI): The unique role of special education and special educators*. Retrieved from www.cec.sped.org

Cox, A. R. (1985). Alphabetic Phonics: An organization and expansion of Orton-Gillingham. *Annals of Dyslexia, 35*, 187–198.

Denton, C. A. (2001). *Responsive reading*. Unpublished curricular materials.

Denton, C. A., Fletcher, J. M., Anthony, J. L., & Francis, D. J. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities, 39*, 447–466.

Denton, C. A., Fletcher, J. M., Stuebing, K., Cirino, P., Barth, A., . . . Francis, D. J. (2009, June). *Characteristics of intervention responders and two groups of impaired readers identified using differing criteria for response to intervention*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Boston, MA.

Denton, C. A., & Hocker, J. L. (2006). *Responsive reading instruction: Flexible intervention for struggling readers in the early grades*. Longmont, CO: Sopris West.

Denton, C. A., & Mathes, P. G. (2003). Intervention for struggling readers: Possibilities and challenges. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 229–251). Timonium, MD: York Press.

Denton, C. A., Mathes, P. M., & Shih, M. (July, 2007). *The effects on student outcomes of coaching support provided to reading intervention teachers*. Poster presented at the annual meeting of the Society for the Scientific Study of Reading, Prague, The Czech Republic.

Edmonds, M., & Briggs, K. L. (2003). The instructional content emphasis instrument: Observations of reading instruction. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for the observation of teaching and learning* (pp. 31–52). Baltimore, MD: Brookes.

Ehri, L. C. (2003). Teaching phonemic awareness and phonics. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 153–186). Baltimore, MD: Brookes.

Engelmann, S., & Bruner, E. C. (1988). *Reading mastery I/II fast cycle: Teacher's guide*. Chicago, IL: SRA.

Engelmann, S., Carnine, L., & Johnson, G. (1988). *Corrective reading: Word attack basics, decoding A*. Chicago, IL: SRA.

Foorman, B. R., Fletcher, J. M., & Francis, D. J. (2004). Early reading assessment. In W. Evers & H. J. Walberg (Eds.), *Testing student learning, evaluating teaching effectiveness* (pp. 81–125). Palo Alto, CA: Hoover Press.

Fuchs, D., Stecker, P. M., & Fuchs, L. S. (2008). Tier 3: Why special education must be the most intensive tier in a standards-driven, No Child Left Behind world. In D. Fuchs, L. S. Fuchs, & S. Vaughn (Eds.), *Response to intervention: A framework for reading educators* (pp. 71–104). Newark, DE: International Reading Association.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239–256.

Good, R. H., Wallin, J., Simmons, D. C., Kame'enui, E. J., & Kaminski, R. A. (2002). *System-wide percentile ranks for DIBELS benchmark assessment* (Technical Report 9). Eugene, OR: University of Oregon.

Haager, D., Klingner, J., & Vaughn, S. (Eds.). (2007). *Evidence-based reading practices for response to intervention*. Baltimore, MD: Brookes.

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *Reading Teacher, 59*, 636–644.

Hunsley, J., & Lee, C. (2007). Research-informed benchmarks for psychological treatments: Efficacy studies, effectiveness studies, and beyond. *Professional Psychology: Research and Practice, 38*, 21–33.

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2007). *The handbook of response to intervention: The science and practice of assessment and intervention*. New York, NY: Springer.

Lee, J., Grigg, W., & Donahue, P. (2007). *The nation's report card: Reading 2007* (NCES 2007-496). Washington, DC: National Center for Education Statistics.

Lovett, M. W., Lacerenza, L., Borden, S. L., Frijters, J. C., Steinbach, K. A., & DePalma, M. (2000). Components of effective remediation for developmental reading disabilities: Combining phonological and strategy-based instruction to improve outcomes. *Journal of Educational Psychology, 92*, 263–283.

Marston, D. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18–78). New York, NY: Guilford.

Mathes, P., & Torgesen, J. (2005). *SRA early interventions in reading* (formerly *Proactive beginning reading*). Columbus, OH: SRA/McGraw-Hill.

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly, 40*, 148–182.

Mathes, P. G., & Torgesen, J. K. (2008). Continuous Monitoring of Early Reading Skills (Version 3.0) [Web-based assessment in reading for grades k-3]. Dallas, TX: Istation.

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

Raudenbush, S.W. (2007). Designing field trials of educational innovations. In B. Schneider & S. McDonald (Eds.), *Scale-up in education: Issues in practice* (pp. 23–40). Lanham, MD: Rowman & Littlefield.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2*, 31–74.

Simmons, D. C., Kame'enui, E. J., Stoolmiller, M., Coyne, M. D., & Harn, B. (2003). Accelerating growth and maintaining proficiency: A two-year intervention study of kindergarten and first-grade children at-risk for reading difficulties. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties* (pp. 197–228). Baltimore: York Press.

Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., . . . Wagner, A. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorders, 37*, 354–366.

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

Spectrum K12 School Solutions. (2009). *Response to intervention (RTI) adoption survey 2009*. Retrieved from www.spectrumk12.com

Swanson, J., Schuck, S., Mann, M., Carlson, C., Hartman, K, . . . McCleary, R. (n.d.) *Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and the SWAN ratings scales*. Retrieved from www.adhd.net

Taylor, B., Pearson, P., Peterson, D., & Rodriguez, M. (2005). The CIERA school change framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly, 40*, 40–69.

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

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: PRO-ED.

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

Tymms, P. (2004). Effect sizes in multilevel models. In I. Schagen & K. Elliot (Eds.), *But what does it mean?* (pp. 55–65). United Kingdom: University of London Institute of Education.

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., . . . Francis, D. J. (2006). Effectiveness of a Spanish intervention and an English intervention for English language learners at risk for reading problems. *American Educational Research Journal, 43*, 449–487.

Vaughn, S., Mathes, P. G., Linan-Thompson, S., Cirino, P. T., . . . Francis, D. (2006). Effectiveness of an English intervention for first-grade English language learners at-risk for reading problems. *Elementary School Journal, 107*, 153–180.

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

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO-ED.

Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review, 36*, 541–561.

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software.* Boca Raton, FL: Chapman & Hall.

Wilson, B. A. (2004). *Wilson reading system* (3rd. ed.). Oxford, MA: Wilson Language Training Corp.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Itasca, IL: Riverside.

ABOUT THE AUTHORS

CAROLYN A. DENTON (CEC TX Federation), Associate Professor, Children's Learning Institute, Department of Pediatrics, University of Texas Health Science Center at Houston. **KIM NIMON,** Assistant Professor, Department of Learning Technologies, University of North Texas, Denton. **PATRICIA G. MATHES** (CEC TX Federation), Professor, Institute for Reading Research, Southern Methodist University, Dallas, Texas. **ELIZABETH A. SWANSON** (CEC TX Federation), Research Associate, Meadows Center for Preventing Educational Risk, University of Texas at Austin. **CAROLINE KETHLEY** (CEC TX Federation), Assistant Professor, Department of Literacy, Language, and Learning, Southern Methodist University, Dallas, Texas. **TERRI B. KURZ,** Research Associate, Meadows Center for Preventing Educational Risk, University of Texas at Austin. **MINYI SHIH** (CEC CA Federation), Assistant Professor, Special Education and Counseling, Charter College of Education, California State University, Los Angeles.

This research was supported by Grant # R305W030257 from the Institute of Education Sciences of the U.S. Department of Education. We acknowledge that the RRI curriculum's primary author was a principal investigator for this study.

RRI was originally developed specifically for the Mathes et al. (2005) study, and the current study grew as a natural extension of that line of research. In the current study, conducted during the 2004–2005 and 2005–2006 school years, RRI teachers were trained in the prepublication version of the intervention (Denton, 2001). The RRI curriculum was subsequently published in 2006. To guard against potential conflicts of interest, the following measures were taken: (a) examiners in pre-post assessment waves were blind to the randomized conditions of the students, although school-based research team members who were more familiar with the students collected the continuous progress monitoring oral reading fluency data, and (b) all data collected at the research site where the principal investigator worked were shipped to the other principal investigator's university immediately after collection, where all data handling and analyses were done by persons with no connections to the RRI program.

Correspondence concerning this article should be addressed to Carolyn A. Denton, Children's Learning Institute, Department of Pediatrics, University of Texas Health Science Center at Houston, 7000 Fannin, UCT 2443, Houston, TX 77030 (e-mail: carolyn.a.denton@uth.tmc.edu).

Manuscript received August 2008; accepted July 2009..