



## Reading Rescue: A follow-up on effectiveness of an intervention for struggling readers

Katharine Pace Miles<sup>a</sup> , Mark D. Lauterbach<sup>a</sup>, Dana M. Murano<sup>b</sup>, and Ginny A. Dembek<sup>c</sup>

<sup>a</sup>Early Childhood, Brooklyn College, Brooklyn, New York, USA; <sup>b</sup>Graduate Center, Educational Psychology, City University of New York, New York, New York, USA; <sup>c</sup>Childhood, Bilingual, and Special Education, Brooklyn College, Brooklyn, New York, USA

### ABSTRACT

The authors examined whether Reading Rescue continues to be an effective literacy intervention and factors that impact its effectiveness. Data were collected on 143 first-grade students, tutored by 104 tutors at 38 schools. There was significant growth on all foundational skills ( $ps < .001$ ) and significant change in proportion of students attaining grade-level reading status pre- and post-intervention ( $ps < .001$ ;  $d = 1.62$  sight words,  $d = 1.68$  oral reading/comprehension). Student-level factors of Individualized Education Program status, program completion, number of session, and invented spelling; tutor-level factors of sessions delivered and years of experience; school level factors of school size and percentage of language minority students all predicted grade-level passage or word reading (all  $ps < .05$ ). Findings indicate Reading Rescue continues to be highly effective, even while expanding substantially as a program since Ehri et al.'s (2007) study. Efforts should focus on ensuring program completion and increasing program exposure in neediest schools.

### ARTICLE HISTORY

Received 23 March 2018  
Revised 16 July 2018  
Accepted 14 August 2018

### KEYWORDS

Tutoring; early literacy instruction; struggling readers; paraprofessionals

### Introduction

Supporting schools to meet the needs of students at risk for academic failure, especially for students who could benefit from early intensive intervention, is both a priority and a challenge. Students not acquiring necessary skills often continue to have academic difficulties and they are at risk for never catching up to their same age peers (Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997; Juel, 1988; NRP, 2000; Stanovich, 1986; U.S. Department of Education, Institute of Education Sciences, and National Center for Education Statistics (NAEP), 2017). Staggering findings show that poor readers in Grade 3 are four times more likely to become high school dropouts (Hernandez, 2011), and that when compared with their average or above average reading peers, students diagnosed with a reading disability at 7 years old were less likely to obtain advanced degrees and more likely to attain lower levels of income (McLaughlin, Speirs, & Shenassa, 2014). It is critical to continually engage in research that can inform schools on how to best support students who have fallen behind grade-level benchmarks.

As Moats (2009, 2011) explained, evidence-based literacy practices have been shown to reduce and/or eliminate the consequences of reading failure, especially in high-risk populations. Kilpatrick (2015) synthesized the literature and explains that when students who have fallen behind in their reading are given the right type of evidence-based intervention, they can develop average reading skills. However, practices lacking in evidence base are continually used in schools

with students at risk for reading difficulties (Moats, 2009, 2011), and in general, for students not reaching grade-level benchmarks reading intervention practices are often a neglected aspect of literacy programs (Allington, 1983; Allington et al., 2015; Chard, Vaughn, & Tyler, 2002; Daly et al., 2015; Kame'enui & Simmons, 2001). There is also confusion and misapplication of non-research-based practices by teachers (Cook & Cook, 2011). Therefore, remedial reading intervention practices warrant further investigation (Wolf & Katzir-Cohen, 2001), and more specifically, clarity is needed regarding the quality and rigor of teacher resources required for practices to be considered evidenced based (Test, Kemp-Inman, Diegelmann, Hitt, & Bethune, 2015). In this study, a Tier 2, evidence-based intervention serving students at risk for reading difficulties was examined.

### Response to intervention

Response to intervention (RTI) is a useful system of supports that can help determine students who may respond to early and intensive intervention. Developed to fit a specific need following the 2004 reauthorization of Individuals with Disabilities Improvement Act (IDEA, 2004), RTI includes supports and services for students at risk for failure in the general education curriculum (Fuchs & Fuchs, 2006). RTI includes three main components: (a) multiple tiers of support, (b) staff and resources for implementation, and (c) collection of evidence-based assessment data to make instructional and intervention decisions. Students move

between tiers using evidence from assessment data to determine if they have met specific grade-level benchmarks. Tier 1 includes high-quality general education instruction, Tier 2 includes evidence-based instruction with more intensity, and Tier 3 includes individualized and intensive support.

The system of RTI supports has been refined (i.e., including four common features: multitiered systems of teaching and caregiving practices; high-quality curriculum for all children; ongoing assessment and continuous progress monitoring for all children; and collaborative problem solving process) and supported as a crucial component of educating all children by stakeholders providing guidance to schools around the country (DEC, NAEYC, NHSA, 2013). High-quality curriculum was added to the components of RTI in the DEC/NAEYC/NHSA position paper, helping to guide educators in a direction of continued educational improvement for all young children.

Fifteen years after RTI was developed, the system of supports has only been evaluated overall once (Balu et al., 2015). Balu et al. determined RTI intervention in Grades 1–3 (i.e., with screening and cutoff percentiles varying from recommendations) had negative or no effects on student performance. It should be noted that many aspects of RTI were lacking in the study, such as the inability to describe the quality of the interventions and the varied margin of identification within the impact schools.

Fuchs and Fuchs (2017) followed with a response citing issues with implementing Response to Intervention. Using a flexible system of supports, the resulting system can be executed in a variety of ways within schools. RTI may be inconsistent on a number of variables such as resources available, state recommendations, and the intensity of teacher training and feedback resulting in variations of the model across the country (Berkeley, Bender, Peaster, & Saunders, 2009; Hudson & McKenzie, 2016). Fuchs and Fuchs found that unlike recommendations in the literature, Tier 2 interventions often wind up being makeshift large group instructional routines without the design required to meet students' needs.

It is not only important for evidence-based approaches to be used throughout the RTI model, but also necessary that there are adequate resources and well-trained staff to deliver the instruction. When viewing RTI from a prevention perspective, different variations of the three tier approach may be encouraged to prevent the need for identification of a disability. Some studies such as O'Connor, Harty, and Fulmer (2005) and Scanlon, Vellutino, Small, Fanuele, and Sweeney (2005) explored how some students may not accelerate with Tier 2 instruction but may benefit and reach grade-level expectations from more intensive one-to-one instruction for a longer amount of time (Scanlon, 2011). Flexibility in approaches may be critical to meet the needs of students who require additional intensity of instruction without a need for special education identification.

## Literacy instruction

In addition to the need for more intensive, individualized instruction, constant examination of current practices and

student performance is necessary. The National Assessment for Educational Progress (NAEP), released as the Nation's Report Card (NAEP, 2017), shows the disheartening data that the majority of students in the United States are struggling to reach grade-level reading proficiency. By fourth grade only 37% of students are reading at a proficient level, while 63% of students are functioning at a basic or below basic level (U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2017). Examining only students who are eligible for the National School Lunch Program, these results increase to 78% performing at a basic or below basic reading level (U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2017). The outcomes for students do not improve as they progress in school, as demonstrated by the finding that 79% of students who qualify for the National School Lunch Program in eighth grade also perform at basic or below basic levels in reading (U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2017). The stagnant number of students functioning at basic or below basic levels suggests that students who fall behind in skills continue to struggle to catch up to their proficient peers.

The National Reading Panel (NRP) recommends evidence-based instruction in five main components to help all students develop proficient reading skills: phonemic awareness, phonics, fluency, vocabulary, and reading comprehension (NRP, 2000). In kindergarten and Grade 1, foundational knowledge in letter knowledge, phonemic awareness, and basic decoding built through phonics instruction are critical skills that enable students to read and spell (Adams, 1990; Ehri, 2014; NRP, 2000; Rack, Hulme, Snowling, & Wightman, 1994; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; Tunmer & Nicholson, 2011). By building strong decoding skills, students are able to free up mental energy to focus on the meaning of text (comprehension). In addition to evidence-based explicit instruction in foundational skills, there is a need for explicitly addressing fluency in reading skills for students with and at risk for disabilities (i.e., with students who have a specific deficit as well as students who have phonological processing deficits; Chard, Vaughn, & Tyler, 2002; Meyer & Felton, 1999). Interventions in early literacy should be systematic in the teaching of each of the five components, which will ultimately support students in reading proficiently.

While students with learning disabilities are most at risk for presenting difficulties in fluency (Meyer & Felton, 1999; Torgesen et al., 2001), second language learners of English are also at risk for reading difficulties (August & Shanahan, 2006). In the literature, *language minority* (LM) refers to students who come from households where a language other than the societal language is spoken (Kieffer, 2011). LM students who have limited English proficiency are at risk for having lower levels of English reading proficiency (August & Shanahan, 2006; Kieffer, 2011, 2008; Lesaux & Kieffer, 2010). Interestingly, Kieffer and Vukovic (2012) demonstrated that LM students and native English speakers in

urban schools with low reading comprehension at the end of third grade had similar profiles of linguistic comprehension and code-related skills, suggesting that these at-risk populations require similar instructional interventions. Early, intensive, and systematic interventions are required to support students with Individualized Education Programs (IEPs) and LM students to prevent further difficulties

There is ample research on the necessary literacy content and instructional approaches to best support struggling readers; however, instruction lacking in an evidence-base continues to be used with students. Rosenshine (2012), Moats (2007), and Sweet (2004) explained the hard truth that ineffective teaching methods limit students' ability to master foundational literacy skills necessary for literacy success. Research also suggests that reading instruction may not be intensive enough in RTI to address more significant reading difficulties (Vaughn, Denton, & Fletcher, 2010). Furthermore, students are not always given systematic and explicit instruction even though research has demonstrated its effectiveness (Archer & Hughes, 2011; Gill & Kozloff, 2004).

Another issue is that tutors from all different educational backgrounds are delivering interventions in reading (Elbaum, Vaughn, Hughes, & Moody, 2000). With a lack of certified reading specialists to work with the large portion of students functioning at or below a basic reading level, interventions are needed that can be implemented successfully by more than just one type of teacher for Tier 2 supports. With sufficient training, the support staff at a school may be a valuable resource for supporting struggling readers. Ehri et al. (2007) found that paraprofessionals tutored as well as reading specialists and other certified teachers for word reading and text comprehension, but there was a significant difference between reading specialist and paraprofessionals, but not other certified teachers, in nonword reading. Causton-Theoharis, Giangreco, Doyle, and Vadasy (2007) review the effective roles paraprofessionals play in supporting student literacy development. More specifically, paraprofessionals have been found to be almost as effective as teachers of older students (Brown, Morris, & Fields, 2005), but less so with first-grade students who are at risk for failure (Wasik & Slavin, 1993). Considering the current state of reading proficiency and the reality of public school funding in high-needs areas, maximizing personnel by training them in a literacy intervention approach may be a wise economic decision. The program examined in this study did just that.

## Reading Rescue

Reading Rescue, developed by Dr. Nora Lee Hoover in 1993, is a reading intervention program geared toward first-grade students. The intervention is provided in a 30-min, one-on-one literacy lesson format that is administered five times a week. The program trains staff such as paraprofessionals, certified teachers, and literacy specialists to conduct the lessons. These individuals are provided with 10 on-site or nearby professional development sessions over the course of two years (Hoover, 2016a, 2016b). In the first year, the

instructors are trained to implement the program. This training includes explicit instruction in the purpose and administration of the assessments and activities conducted in the easing-in sessions, as well as how to conduct the five components of the intervention lesson plan. Also, tutors are trained in how to identify students who have met grade-level benchmarks and are ready to graduate from the program. In the second year of training, tutors are provided with more in-depth training in supporting students through advanced comprehension strategies, evaluating data to drive instruction, and using culturally responsive practices to support students from minority communities, students with reading disabilities, and English language learners (ELLs).

The Reading Rescue program is grounded in evidence-based practices as determined by the National Reading Panel (NRP, 2000; O'Leary & Hoover, 2015), and therefore, the lessons are designed to support students' phonemic awareness, phonics, fluency, vocabulary, and comprehension skills. Each Reading Rescue lesson follows the same standard progression (Hoover, 1998). In the first part of the reading lesson students practice their fluency skills by reading books that are at their independent reading level. In the second part of the lesson tutors take a written record of students' oral reading fluency using a book that was briefly previewed at the end of the previous lesson. The third component of the lesson is a multisensory word attack session designed to enhance phonemic awareness and phonics skills. The fourth part of the lesson involves sentence writing to support syntactic development and mental orthographic images. Last, in the fifth part of the lesson students preview a book that will be used for the written record the next day. Throughout all of the lesson components tutors use specific strategies to support vocabulary and comprehension. Unfamiliar words that the student encounters are defined and both literal and inferential questions are asked for each book the student reads during the session.

Reading Rescue has some similarities to, but also important difference from Reading Recovery. Both use extensive assessments to determine students' baseline skills to inform targeted instruction, and both use a similar lesson plan structure, which focuses on fluency practice and assessment, comprehension check-ins, and writing development. However, Reading Rescue has more flexibility in selecting students for the program, and its lesson plan teaches phonics in an explicit and systematic way, whereas the Reading Recovery lesson plan incorporates word study in-context while students are reading and writing. This key difference comes from the theoretical grounding of each program. Reading Rescue emphasizes the challenge that struggling readers have in decoding print, while Reading Recovery emphasizes the challenge struggling readers may have in accessing three sources of information: meaning, language structure, and print. Other important differences include Reading Rescue serving students beyond Grade 1, and its use of paraprofessionals from the schools as tutors. This last point creates vast differences in the cost of the program for schools. Instead of paying upward of \$100,000 (salary of experienced teacher with benefits package) for master's-level

teachers with classroom experience, which Reading Recovery requires, Reading Rescue trains underutilized staff, such as paraprofessionals, who are already employed by the school.

An independent evaluation using an experimental design provided promising evidence for Reading Rescue. Muller and Davies (2004) conducted a study in which urban, LM, low socioeconomic status (SE), struggling readers were randomly assigned to either Reading Rescue or regular classroom instruction. Results showed that students in the Reading Rescue treatment condition performed significantly higher on the Gates-MacGinitie Reading Test–Fourth Edition (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) than the control group. The effect sizes of the treatment condition was 0.32.

Ehri et al. (2007) also focused on low SES, LM first-grade students with reading difficulties, but they extended the Muller and Davies (2004) study by including multiple comparison groups. Whereas the Muller and Davies control group was comprised of students in the same school as those who were receiving the treatment, Ehri et al. used three control groups: students who received an alternative treatment, students who received regular classroom instruction at the same school as those in the Reading Rescue treatment, and students who received regular classroom instruction at schools that did not have the Reading Rescue program. To further extend the Muller and Davies study, Ehri et al. included an additional standardized test to measure performance, and they did not include any schools who were in their first year of implementing the Reading Rescue program (half of the schools in the Muller and Davies reported were in the first year). Ehri et al. also considered the type of tutor (i.e., paraprofessional vs. certified teacher) and their efficiency of administering the sessions on student performance.

Ehri et al. (2007) found that Reading Rescue–tutored students made significantly greater gains in both reading words and comprehending text. These results were clear across the controls, including a control groups receiving a small group, explicit instruction, commercially prepared program. Students who received Reading Rescue instruction were able to reach average grade-level reading levels, unlike the controls. The other key aspect of this study was the examination of the effectiveness of different backgrounds of tutors. As previously mentioned, the study found that paraprofessionals needed more time to reach the same gains in reading for students but were able to instruct students on reading real words and comprehending texts as well as certified teachers and reading specialists. The one area where the tutors were less successful was in nonword/pseudo word reading.

It has been 11 years since Ehri et al. (2007) published their findings on the Reading Rescue intervention. In the meantime, Reading Rescue, in its current form as a program serving one large metropolitan area, has expanded from serving approximately 20 schools per year starting in 1998 to now serving 90 schools with 836 elementary school staff tutoring 1,031 students. In 11 years, the large metropolitan area has had changes in education leadership, which inevitably brought changes in literacy assessment and curriculum options, and like all metropolitan areas in the United States,

has seen an increase in the number of LM students. Over the same period of time, Reading Rescue has gone from having one individual, its founder, Dr. Hoover, conducting the training sessions or training a handful of staff to conduct the sessions, to having multiple staff members, with varying degrees of experience in literacy instruction conducting the trainings, without the benefit of being trained by the program developer.

While an experimental design was outside the scope of this study, we examined whether Reading Rescue continues to be effective as a literacy intervention program in the early grades. As previously mentioned, research needs to continually investigate the effectiveness of Tier 2 interventions to ensure they are improving reading performance of students who have fallen behind grade-level benchmarks. Similarly to Ehri et al. (2007), we were interested in the effects of Reading Rescue on foundational level skills and grade-level sight word reading and reading comprehension. Though we used a single group design, we were also interested in whether the pre- to postintervention gains in our sample were similar in magnitude to those found in the intervention condition of Ehri et al. Additionally, we explored student-, tutor-, and school-level characteristics hypothesized to impact Reading Rescue's effectiveness. Ehri et al. found that tutor backgrounds impacted students' ability to reach grade-level proficiency; we therefore considered this as well as other tutor- and school-level factors. Our research questions were as follows,

Research Question 1: Do students make significant growth in all foundational level skills?

Research Question 2: Will Reading Rescue tutoring increase students' reading achievement to grade-level proficiency on measures of sight word reading and reading comprehension?

Research Question 3: Will student characteristics (gender, LM status, IEP status, number of session, graduation from program, pretest scores) predict how much first-grade students improve in their reading as a result of tutoring?

Research Question 4: Will tutors with certain characteristics (type of tutor, years of tutoring experience, number of easing-in and easing-out sessions, and regular instructional sessions) be equally effective at improving the reading achievement of their students?

Research Question 5: What school factors (% free/reduced lunch, % IEP, % LM, size of school, number of first-grade students, number of years school has received Reading Rescue tutoring) will predict student performance?

## Method

### Research design

As this project was a retrospective analysis of data collected by Reading Rescue, both the pool of participants and the research design were constrained. The complete student data that Reading Rescue had available was used for the analysis. While this afforded the researchers a rich data set, a control group was not available for analysis. To overcome this, a pretest-posttest design was used to analyze performance in foundational literacy skills, oral reading with comprehension, and sight word reading. Effect sizes from this dataset

were compared with previous research. Logistic regression was employed to identify factors at the student, tutor, or school level that influenced the odds that students would reach grade level in sight word reading and oral reading with comprehension.

### Participants

Data on foundational literacy skills, sight word reading, and reading comprehension ability were collected from 161 first-grade students, who were tutored by 104 tutors at 38 schools during the 2016–2017 school year. Of the 161 first-grade students, 143 had complete data to use in the analysis. Seventy-five participants were females, 13 had IEPs, 20 were ELLs, and 87 completed the program before the end of the school year (see Table 1). All students were included in the analysis, whether they graduated the program or not over the course of the school year. Schools identified students as ELLs based on parental report at school entry of a language other than English spoken at home, and the students' performance on a state-developed English proficiency test. These students will be referred to as LM students in this paper to align with the literature. Teachers determined that LM students had sufficient English language skills to participate in the intervention, and students with IEPs were included in the intervention only if this type of reading intervention was an appropriate fit based on the recommendations of the IEP.

Of the 104 tutors, 65 were paraprofessionals, 29 were teachers, 2 were Read More Corps teachers (AmeriCorps members specifically trained in reading interventions), 2 had a literacy specialty (Reading Specialist or Literacy Coach), 3 were school aides, 1 was a librarian, 1 was a mathematics coach, and 1 was a counselor. Years of tutoring experienced ranged from one to nine years, with tutors having a mean of 1.98 ( $SD = 1.53$ ) years of experience. Tutors were full-time staff or faculty members at the school. They all had other responsibilities during the school day, but they were released from these responsibilities by the administrators at the school to tutor their assigned Reading Rescue students during the day. All tutors remained with their student throughout the school year. Eighty-five tutors provided instruction to only one student over the course of the year, while 19 provided instruction to two or more students (range = 2–7 students).

All tutors received training in the Reading Rescue intervention. There were 10 training days administered over two years. The first five training days administered in year one taught tutors the essentials for administering the program. This included how to conduct the five easing-in sessions, which build rapport between the tutor and student. During the easing-in sessions, fun literacy activities are conducted and preassessments are given. After learning about the easing-in sessions, tutors were trained on conducting the intervention sessions, each of which consists of a five-part lesson plan. Tutors were thoroughly trained on the purpose of each part of the lesson plan and the instructional approach and strategies to use during each part of the lesson. Tutors were

Table 1. Foundational skill scores at pretest by subgroup.

	Group	Combined letter names	Letter sounds	Print concepts	Invented spelling	Blending	Segmenting	Nonword reading	Orthographic memory
Gender	Overall (N = 143)	50.30 (3.10)	22.30 (3.40)	17.65 (3.89)	24.17 (6.52)	7.84 (2.50)	5.57 (3.02)	4.62 (3.93)	12.14 (7.71)
	Boy (n = 68)	50.21 (2.91)	21.99 (3.55)	17.28 (4.58)	23.18 (7.20)	7.65 (2.69)	5.44 (3.24)	4.12 (3.90)	10.75 (7.54)
IEP	Girl (n = 75)	50.39 (3.28)	22.59 (3.26)	17.99 (3.12)	25.07 (5.74)	8.01 (2.32)	5.69 (2.82)	5.11 (3.91)	13.41 (7.50)*
	No (n = 128)	50.32 (2.93)	22.31 (3.50)	17.81 (3.93)	24.09 (6.51)	7.84 (2.61)	5.57 (2.99)	4.57 (3.96)	12.18 (7.37)
	Yes (n = 13)	49.92 (4.71)	22.12 (2.64)	16.62 (3.43)	24.50 (7.22)	7.92 (1.38)	5.61 (3.45)	5.15 (3.82)	12.62 (10.24)
LM	No (n = 123)	50.27 (3.27)	22.20 (3.53)	17.75 (4.00)	24.54 (6.62)	7.83 (2.53)	5.55 (3.04)	4.82 (4.10)	12.17 (7.51)
	Yes (n = 20)	50.50 (1.76)	22.92 (2.49)	17.05 (3.17)	21.93 (5.53)	7.90 (2.38)	5.70 (2.99)	3.50 (2.33)*	12.00 (8.35)
Status	Completed (n = 87)	50.30 (3.30)	22.45 (3.47)	18.24 (3.62)*	24.53 (6.66)	8.06 (2.30)	5.70 (3.08)	4.58 (3.93)	12.52 (6.53)
	Exited (n = 56)	50.31 (2.78)	22.06 (3.31)	16.73 (4.14)	23.62 (6.31)	7.50 (2.76)	5.38 (2.94)	4.71 (3.95)	11.55 (9.07)

Note. Total n for IEP group is 141 due to missing IEP data for two students. IEP = Individualized Education Plan; LM = language minority.

\*  $p < .05$ .

also trained on administering the postassessments in the first year of training. All of the materials, including assessment sheets, series of leveled books, magnetic letters, daily lesson plan logs, phoneme segmentation sheets, and practice writing pages were provided to the tutors, and instructions on how to use all of the materials was incorporated into training received in year one. In the second year of training, tutors were provided with more in-depth training in supporting students through advanced comprehension strategies, evaluating data to drive instruction, and how best to use culturally responsive practices to support students from minority communities, students with reading disabilities, and LM students.

On average, 55.47 regular sessions were provided to students, with 5.04 easing-in and 4.19 easing-out sessions. The easing-in and easing-out sessions provided the time to do pre and posttest for the students. Table 2 provides a breakdown based on the number of students receiving the intervention from the different tutoring categories. As there were statistically significant differences in the number of easing-in and -out sessions, this was included in the analysis of tutor effectiveness (see the following sections).

Of the 38 participating schools in this study, on average, these schools had 571.97 students ( $SD = 248.94$ , range = 182–1,484), there was an average of 3.65 ( $SD = 1.66$ , range = 1–7) Grade 1 classrooms per school, 24.4% of students at these schools had IEPs ( $SD = 0.06$ , range = 9–39%), 16.6% of students at the schools were LM students ( $SD = 12.3\%$ , range = 1–67%), 89.8% of students at the schools qualified for free lunch ( $SD = 14.6\%$ , range = 27–100%), and the schools had the Reading Rescue program for an average of 2.26 years ( $SD = 1.80$ , range 1–9 years). The schools where participants were enrolled used a variety of curriculum as there is no set curriculum for schools in this large metropolitan area. Common curricula are as follows: Teachers College Reading and Writing Project, Wilson's Foundations, iReady, and ReadyGEN.

## Measures

Individual assessments were administered by the tutors to students who were selected for the intervention. The pretests, posttests and intervention were conducted by the trained Reading Rescue tutors. The researchers had no direct contact with the Reading Rescue tutors or the students. All the measures were administered before the start of and at the completion of the intervention (pretest-posttest design).

### Reading Rescue assessments

A battery of nine assessments (seven practitioner-created measures from the Reading Rescue Individual and Classwide

Assessments of Early Literacy and two researcher-created measures) were administered to determine pre and posttest performance of foundational literacy skills. During the first five sessions, referred to by the program as easing-in sessions, tutors administered the assessments. The same assessments were given as posttests during the easing-out sessions (i.e., after the intervention was complete).

1. Letter-Sound Knowledge: Students were asked to identify the name of uppercase letters, provide the sound that the letter makes, and provide a word that starts with this letter. The task was repeated with lowercase letters. If two sounds were available for a consonant, either sound was accepted as correct. For vowel sounds, after the student provided one sound she or he was asked by the tutor if the letter made another sound. Responses were scored as correct or incorrect for each of the 26 total items. Reliability was  $\alpha = .834$  at pretest and  $\alpha = .827$  at posttest.
2. Print Concepts: Tutors presented students with a book that was upside down and backward. Then the tutor moved through a series of prompts to determine if the student could identify different parts or characteristics of the book. Prompts included, "Please show me where to begin reading on a page," and "Please show me a space between two words." Responses were scored as correct or incorrect for each of the 25 total items. Reliability was  $\alpha = .793$  at pretest and  $\alpha = .682$  at posttest.
3. Phoneme Blending: Students were asked to listen to the parts of words said by the tutor and then blend the sounds together to make a word. For example, tutors would say the two syllables in the word *before* as *be-fore*. Students were given one example in the directions and five practice items. Items were separated into syllables (*pup-py*), word families (*m-at*), or phonemes (*d-o-g*). Responses were scored as correct or incorrect on ten total items. Reliability was  $\alpha = .881$  at pretest and  $\alpha = .554$  at posttest.
4. Phoneme Segmentation: Students were asked to segment words into their individual phonemes. Students were given one example in the directions, three practice words which had two phonemes each, and then five test items. Then students received three more practice words with more than two phonemes, and then the remaining five test items. Responses were scored as correct or incorrect for ten total items. Reliability was  $\alpha = .888$  at pretest and  $\alpha = .671$  at posttest.
5. Invented Spelling: Students were asked to spell a list of orthographically complex real words for a first-grade student. Students were encouraged to sound out the words and represent the sounds to the best of their

**Table 2.** Average number of regular, easing in, easing out, and total sessions, based on number of students receiving Reading Rescue by tutoring category.

Tutor type	Teacher ( $n = 36$ students)	Paraprofessional ( $n = 87$ students)	Other ( $n = 20$ students)	Total	$F$
Regular session	59.39 (27.26)	51.56 (24.02)	65.40 (35.75)	55.47 (27.03)	2.70
Easing in	5.31 (0.71)	4.93 (0.70)	5.05 (0.22)	5.04 (0.67)	4.16*
Easing out	3.81 (1.45)	4.21 (1.03)	4.80 (0.70)	4.19 (1.14)	5.17*

Note. Bonferroni post hoc analysis revealed the following differences: for easing-in Teacher > Para, for easing-out Other > Teacher.

\* $p < .05$ .

ability. A rubric was used for scoring. Half a point was given if one phoneme other than the initial phoneme was written, one point if the initial phoneme was accurately represented, two points if the initial phoneme and one other accurate phoneme was represented, three points if the initial phoneme and two or more other phonemes were accurately represented, and four points if it was a phonologically accurate spelling. There were 10 total items. Reliability was  $\alpha = .850$  at pretest and  $\alpha = .742$  at posttest.

6. **Nonword Reading:** Students were asked to read a list of words that are not real words. The tutor acknowledged that the students had never seen the word before so the student should just do their best to read the word. If students did not want to try to read the words the task was discontinued. Tutors stopped administering the task after three incorrect responses within a set of five. There were 15 total items. Reliability was  $\alpha = .882$  at pretest and  $\alpha = .812$  at posttest.
7. **Word Writing:** The purpose of this task was to evaluate the student's orthographic memory. Students were given 10 min to write all the words they knew how to spell. If after a few minutes of writing the student was at a loss for what to write, the tutor could prompt the student by asking if he or she had siblings or pets, or the tutor may ask about other aspects of the student's life to probe for potential known words. The task was score for the total number of words spelled correctly.
8. **Sight Word Reading:** The sight word subtest (i.e., San Diego Quick Assessment) of the informal reading inventory discussed below was administered. Lists of 10 words organized in increasing difficulty (preprimer to Grade 4). Students began reading the preprimer list and moved to the next list only if they scored 8 of 10. The task was discontinued when students read seven or fewer words correct in a list.
9. **Oral Reading and Comprehension Questions (Informal Reading Inventory):** The Ekwall/Shanker Reading Inventory was administered to assess students oral reading fluency and comprehension skills. Student performance on the sight word task determined the starting level for the passage reading. A set of comprehension questions accompanied each passage. Two scores were calculated for each passage: correct words read and comprehension accuracy. The determination of whether a student is at the independent, instructional or frustration reading level is based on a combination of those two scores. For example, if a student read 66 of 71 words correctly on the Grade 1 passage they could miss one comprehension question and still be at the instructional level. Alternate form reliability was reported in test manual as  $\alpha = .82$ .

### Reading Rescue intervention protocol

First-grade students were selected for the program based on their performance on Reading Rescue Classwide Assessments (whole group-administered tasks of letter writing, writing

words they know how to spell, and developmental spelling, which examined accurate letter-sound representations in words) independent from or in conjunction with their performance on school adopted oral reading and comprehension assessments. The students were administered the measures listed above as pretests during a series of easing-in sessions. The pretests were interwoven with activities to support literacy (clapping the syllables in "Twinkle, Twinkle, Little Star"; drawing and then writing a sentence about the picture; using magnetic letters to put the alphabet in order) during the easing-in sessions. The purpose of the easing-in sessions was both to collect baseline data and to develop the tutor-student relationship. After all pretests and activities were completed, the tutor began using the Reading Rescue lesson plan. The lesson plan is comprised of five parts, as detailed below. The lesson plan is differentiated to meet the needs of students as they progress through four phases of reading development, and a series of leveled books was provided to each school to ensure tutors had a sufficient number of books at each reading level. Through a series of informal assessments (written records) tutors gathered evidence that students were reading on grade level, according to the leveled reading system, and at that point the tutor began the easing-out sessions. These easing-out sessions were similar to the easing-in sessions in that the measures listed above were administered as post tests and literacy activities were conducted to support students in transferring their skills back to the classroom.

*Reading Rescue Lesson Plan.* The lesson plan is comprised of five components: fluency building with comprehension support, oral reading assessment, multisensory word work, writing, previewing a new book with comprehension support. There are specific allotments of time for each part, and the lesson takes 30 min in total.

1. **Fluency Building and Comprehension Check-In.** Students begin each lesson by reading several familiar books. These books were previously used for assessment and are now at students' independent reading level. In Phase one the focus is on building sight words, accurate reading, and comprehension as modeled through think alouds and question asking from the tutor. In Phase two the focus is on accurate decoding using phonetic elements taught in previous sessions as well as blending, segmenting, and syllabication skills that have been previously taught. In Phase three and four timed readings are used to further build speed, while also maintaining a focus on accuracy and prosody.
2. **Oral Reading Assessment.** Students are assessed on their ability to accurately read a book that was introduced to them the previous day in part five of the lesson. A written record is taken to document students' ability to accurately read the words. Notes are taken on any miscues or strategy supports that may be beneficial to the students (i.e., using your finger to point at the words, breaking long words into smaller parts). An accuracy score on reading 100 words from the book is used to determine if the student is ready to move up to the next level of books. An accuracy score of 96- 99%

indicated that the book was at a student's independent level. An accuracy score of 90 – 95% indicated a book was at a student's instructional level, and a score less than 90% indicated a book was at a student's frustration level. A set of multiple written records that demonstrated a student's score of reading books at a certain level in the independent range signaled that the student was ready to move up the next level of books. Tutors were advised to move students to the next level when the student received 95% or greater accuracy three days in a row on a given book level. Tutors were also trained to keep student's comprehension accuracy in mind when moving up levels. If the child read with 95%, but could not sufficiently answer the tutor's literal, informational, and evaluative questions about the book, the student was not ready to move up in levels.

3. **Multisensory Phoneme Segmentation and Word Analysis.** Several activities could be used during this part of the lesson to work on phoneme segmentation and phonetic elements in words. For the magnetic letters activity, a series of ten short steps were used to have students say a word, stretch the sounds in the word, segment the sounds, pull down the magnetic letters in the word, remove letters in the word to target a particular sound, and then use their whole arm, finger and magnetic letters to practice the sound in isolation and then rebuild the word. A practice notebook called "Letter and Words I Know" was also available to use during this time. In this workbook, pictures of words that begin with the corresponding letter (and sound) of the alphabet were presented on each page. Students practiced segmenting the phonemes by sliding chips that represent each sound in the word into an Elkonin box. Students began segmenting sounds in two phoneme words and then progressed to three and four phoneme words. There was also a page for each letter of the alphabet where a vocabulary word that begins with the letter could be spelled and defined.
4. **Writing.** In the fluency building part of the lesson the tutor documented a sentence that the student said about the book used in part one of the lesson. The tutors annotated each word in the sentence as a word the student knows, a word that should be segmented into Elkonin boxes, a word that is too difficult so the tutor will simply give the spelling, and a word that is good for high-frequency/sight word practice. In this part of the lesson the tutor reads the sentence and the tutor and student practice saying the sentence together three times. The tutor and student practice writing each word according to the categorizations on the Practice Page. When the spelling of each word has been sufficiently practiced, the student adds the word to the sentence on the Perfect Page.
5. **Comprehension and Vocabulary Building with a New Book.** At the end of each lesson a new book is introduced. In the early phases the tutor walks the student through the book page by page to make predictions, decode difficult words, and define unknown vocabulary words. In

later phases, students are encouraged to make predictions about the book based on the cover and title of the book.

## Data collection

A sample of schools that had adopted Reading Rescue and that had procedures in place for ensuring complete pre-post data were used for this analysis. It should be noted that all schools conducting the Reading Rescue program are required to complete the pre and post assessments for each student, and the instructors and Reading Rescue program managers went over student performance on the assessments to inform instruction and to determine graduation status from the program. While the pre/post assessments were conducted, completion of the data report sheet, an extra step in the process used for external purposes, was not always completed possibly due to a lack of understanding the larger purpose of the sheet as a way to evaluate the program overall.

Completed data report sheets were gathered by the Reading Rescue staff for general program evaluation purposes, and the final spreadsheet of the completed data was turned over to the researchers for evaluation in de-identified form. The tutors were not told about the study at the time of data collection because this study had yet to be conceived. Instead, tutors collected data as they normally would and entered it into the Reading Rescue data management system. The researchers did not interact with students or instructors. After the data collection process was complete and after the researcher were granted IRB approval through their university to conduct a secondary data analysis, the data spreadsheet was turned over to the research team. It should be noted that the lead researcher provides academic consulting to the organization so a research team (coauthors) was convened to ensure that biases did not affect the results. The researchers who were unaffiliated with the program lead the data analysis of this study to ensure that the lead author's involvement with the organization did not bias any of the results.

Student performance on pre to post assessments was analyzed. Of interest was whether there was substantial growth in foundational literacy scores and whether students reached grade-level benchmarks as indicated by the oral reading and comprehension measures. Also of interest was the student-level characteristics that predict success in the program, as well as the school and tutor characteristics that impact student performance.

## Analyses

To determine the effectiveness of Reading Rescue, the researchers first examined students' baseline scores on all foundational skill reading measures. Additionally, we computed the number of students who attained grade-level oral reading with comprehension and sight word reading scores before receiving the intervention. To determine how much students improved after receiving the intervention, we conducted paired-samples t-tests on each of the foundational skills reading measures and computed effect sizes. We also calculated the percentage of students who reached grade



level on the oral reading with comprehension and sight word reading measure after receiving the intervention and computed an effect size.

In addition to growth on foundational reading measures, oral reading with comprehension, and sight word reading, we were also interested in various student-level factors that could affect whether or not students reached grade-level oral reading with comprehension and sight word reading scores by the end of the intervention. Student-level variables of gender, LM status, IEP status, number of sessions attended, whether students graduated (i.e., completed) or exited from the program, and pre test scores on all foundational reading skills measures were entered as predictors in logistic regression models. We were interested if any of these factors significantly predict students being at or above grade level on posttest administrations of the oral reading with comprehension composite score and sight word reading measures.

We also hypothesized that tutor-level characteristics would predict whether or not students reached grade-level reading proficiency, as measured by the oral reading with comprehension and sight word reading tasks, by the end of the intervention period. Specifically, we hypothesized if type of tutor and years of tutoring experience would be equally effective at improving the reading achievement of their students. Additionally, we looked at the number of easing-in sessions, regular instructional sessions, and easing-out sessions. Each of these factors was entered into a logistic regression model to determine if tutor-level factors also predict students reaching grade level by the end of the intervention.

Last, we tested whether school-level factors predict students being at or above grade level at the end of the intervention period. School-level variables of percent free lunch, percent of students with IEPs, percent of LM students, size of school, number of first-grade students, and the number of years the school has received Reading Rescue training were all entered as predictor variables in a logistic regression model to determine if any of these factors predicted students reaching grade-level reading proficiency in oral reading with comprehension and sight word reading at the end of the intervention.

## Results

### Pre-intervention baseline scores

Table 1 shows various subgroups of participants' pretest scores on all baseline measures before receiving the Reading Rescue intervention. In general, student scores were low

compared with the maximum scores students could have received on each of the measures. At baseline, girls performed significantly better than boys on the orthographic memory measure,  $t(141) = -2.111, p = .037$ . LM students scored significantly lower on nonword reading than non-LM students did,  $t(41.45) = 2.064, p = .045$ . Students who graduated from the program (i.e., fully completed it) scored significantly higher on print concepts than students who exited from the program,  $t(141) = 2.301, p = .023$ . No other group differences were statistically significant at baseline, but the small number of students with IEPs and LM students ( $n = 13$  and  $20$ , respectively) likely contributed to an insufficient amount of power. Pretest measures were moderately correlated with one another (Table 3).

Due to the administration procedures for the oral reading and comprehension measure (Ekwall/Shanker), in which testing is discontinued when the student's performance falls below the criteria for grade-level reading according to the administration manual, we did not have sufficient continuous data to compute pre- and posttest means for this measure. This was a limitation. Instead, we calculated the percentage of students who were at or above grade level using the scoring criteria from the test manual. Based on this criterion, only 9.1% of all students were at or above grade level before receiving the intervention, whereas 67.8% of all students were at or above grade level at the end of the intervention. From pre- to posttest, the difference in proportions of students who achieved grade-level proficiency in oral reading/comprehension showed a large effect size (Cohen's  $d = 1.68$ ). This shows that many students who received the intervention were able to achieve grade-level proficiency in oral reading/comprehension by the end of the intervention.

A similar limitation stemmed from the administration pattern of the San Diego sight word measure. The test was not administered to students who experienced frustration at levels before the Grade 1 measure. Similarly, we computed whether or not students were at grade level at the start of the intervention and then at the end of the intervention year using criteria from the scoring manual. Only 14.5% of students were at grade level at the beginning of the school year, which increased to 76.1% by the end of the intervention year. From pre- to posttest, the difference in proportions of students who achieved grade-level proficiency in word reading showed a large effect size (Cohen's  $d = 1.62$ ), similarly demonstrating that students receiving the intervention were able to achieve grade-level performance on sight word reading by the end of the intervention year.

Table 3. Correlation matrix of foundational reading measures at pretest.

	Combined letter names	Letter sounds	Print concepts	Invented spelling	Blending	Segmenting	Nonword reading
Combined letter names	—						
Letter sounds	.550**	—					
Print concepts	.271**	.345**	—				
Invented spelling	.332**	.455**	.293**	—			
Blending	.201*	.267**	.348**	.446**	—		
Segmenting	.231**	.283**	.383**	.395**	.524**	—	
Nonword reading	.200*	.341**	.280**	.590**	.392**	.441**	—
Orthographic memory	.227**	.193*	.362**	.248**	.219**	.183*	.216**

\* $p < .05$ ; \*\*  $p < .01$ .

### Growth on foundational skills, sight words, and oral reading with comprehension

From pre- to posttest, students made significant growth on all foundational reading skills. Table 4 shows the means and standard deviations for pre- to posttest on each measure, as well as the paired samples *t* value and computed effect size (Cohen's *d*) for each foundational skills measure. Effect sizes ranged from Cohen's *d* = 0.70 (combined letter names) to Cohen's *d* = 1.68 (print concepts), with a mean of Cohen's *d* = 1.29. These can be interpreted as large pre to post difference effect sizes, particularly in the domain of education. Students who participated in the Reading Rescue intervention group in Ehri et al.'s (2007) study showed improvements similar to this scale in reading comprehension (Cohen's *d* = 1.33). Participants in Ehri et al.'s prior study also showed pre to post effect size gains of Cohen's *d* = 1.30 in letter sound recognition and Cohen's *d* = 2.24 in word writing. These magnitudes in similar foundational skills are comparable to the magnitudes of the effects in the current study. Despite the lack of a control group in this study, students who received Reading Rescue made just as large of gains in both Ehri et al. and the current study, suggesting the program continues to remain effective in practice.

### Student-level factors on grade-level performance

To determine which student-level factors most significantly affect student oral reading and comprehension at the end of the intervention year, stepwise binary logistic regression models were run. The dependent outcome of interest was whether students reached grade-level proficiency in a composite score of oral reading and comprehension. With the constant entered as the only variable in the baseline model, the classification table showed that there was a 66.9% chance of predicting the correct outcome (whether the student was at grade level or not). First, student-level demographic characteristics of gender, IEP status, and LM status were entered into the model. The percentage correct did not increase and the model showed a Cox and Snell *R*<sup>2</sup> value of .037, showing that adding these variables into the model did not improve

**Table 4.** Pre- to posttest differences on foundational skill measures.

Measure	<i>M</i> ( <i>SD</i> )	<i>t</i>	Cohen's <i>d</i>
Combined letter names - pre ( <i>n</i> = 143)	50.30 (3.10)		
Combined letter names - post ( <i>n</i> = 143)	51.85 (.56)	-6.093**	0.6958
Letter sounds - pre ( <i>n</i> = 142)	22.30 (3.40)		
Letter sounds - post ( <i>n</i> = 142)	25.53 (1.40)	-12.201**	1.2423
Print concepts - pre ( <i>n</i> = 143)	17.65 (3.89)		
Print concepts - post ( <i>n</i> = 143)	22.87 (2.02)	-18.988**	1.6842
Invented spelling - pre ( <i>n</i> = 143)	24.17 (6.52)		
Invented spelling - post ( <i>n</i> = 143)	33.02 (4.32)	-17.601**	1.6002
Blending - pre ( <i>n</i> = 143)	7.84 (2.50)		
Blending - post ( <i>n</i> = 143)	9.80 (0.56)	-9.686**	1.0819
Segmenting - pre ( <i>n</i> = 143)	5.57 (3.02)		
Segmenting - post ( <i>n</i> = 143)	9.47 (5.00)	-8.403**	0.9442
Nonword reading - pre ( <i>n</i> = 141)	4.62 (3.93)		
Nonword reading - post ( <i>n</i> = 141)	11.51 (3.61)	-21.029**	1.8259
Orthographic memory - pre ( <i>n</i> = 143)	12.14 (7.61)		
Orthographic memory - post ( <i>n</i> = 143)	27.69 (16.06)	-13.137**	1.2374

\*\**p* < .01.

the model much, though IEP status had a strong odds ratio at this first step.

In the second step, student-level characteristics associated with program exposure were entered (whether students completed the program or ended the school year without completing the program, and the number of sessions they received). This improved the classification table proportion to 77% and *R*<sup>2</sup> increased to .20. Last, we added pretest scores on foundational skill measures in as step 3. This model yielded the highest percentage correct (79.9%) and increased *R*<sup>2</sup> to .307, suggesting that this is the best model. In the final model IEP status (*B* = 1.817, *SE* = 0.924), reason exited (*B* = -1.701, *SE* = 0.449) and the number of sessions received (*B* = 0.042, *SE* = 0.012) were significant (*p* < .05 for IEP status and *p* < .001 for the other two) predictors in the model. Students with IEPs (though we do not know the diagnosis of the IEP) did remarkably well with an odds ratio of 6.2. Students who did not complete the program due to time constraints, had an odds ratio of 0.194, meaning they were over 80% less likely to reach grade level than were those who did. Students who received more sessions increased odds of achieving grade-level comprehension scores by 4% for each session (odds ratio = 1.04). Students with IEPs were far more likely to reach grade level, exiting early from the program, rather than completing it, which reduces the odds that a student will attain grade-level proficiency, whereas attending more sessions increases the likelihood that a student will reach grade level.

In looking at the strong IEP effect and the noneffect for LM students, the actual percentage of students reaching grade level was further examined. For students without IEPs or LM status the percentage was 66.7%, for students with LM status it was 65% and for students with IEPs it was 84.6%. It is important to note while the success rate was high for students with IEPs, the small sample size makes it difficult to draw too strong of a conclusion from it—if two more students did not achieve grade level the percentage would have been 69.2% very close to the other groups.

The same blocks were entered to predict grade-level status on the San Diego sight words measure. The full model showed a classification table percentage of 82.1% and *R*<sup>2</sup> of .225. Both reason exited (*B* = -0.872, *SE* = 0.453) and the number of sessions received (*B* = 0.032, *SE* = 0.011) were significant (*p* < .05) predictors in the model as well, in addition to the pretest score on the invented spelling measure (*B* = 0.108, *SE* = 0.046, *p* < .05). Exiting the program early decreased the odds of being at grade level by almost 60% (odds ratio = 0.418), more sessions received increased the odds of being at grade level by 32% for each session (odds ratio = 1.032), and higher scores on the invented spelling pretest increased the odds of being at grade level by 11% for every point (odds ratio = 1.114).

### Tutor-level factors on grade-level performance

The effects of tutor-level factors on grade-level oral reading and comprehension were also explored using binary logistic regression models. Dummy variables were created using

paraprofessionals as the baseline category. The dummy variables for teacher and other school-professionals were entered into the first block of the model, along with the number of years the tutor had spent delivering the Reading Rescue intervention. In the second block, the number of regular sessions and the number of easing-in and -out sessions were included as predictors. The overall model only increased the predicted percentage from 67.8% (constant only) to 72% with an  $R^2$  of .089, suggesting that student-level characteristics were more salient in predicting student grade-level outcomes postintervention. Of the variables in this model, the number of regular sessions ( $B=0.016$ ,  $SE=0.008$ ) was the only significant predictor ( $p<.05$ ). Number of years as a Reading Rescue tutor ( $B=0.277$ ,  $SE=0.163$ ) and the dummy teacher variable ( $B=0.807$ ,  $SE=0.453$ ) were both marginally significant ( $p<.09$ ). The effect of session was small, an almost 2% increase in the odds for every session (odds ratio = 1.016). However, while not significant at the .05 level, the effects of the number of years of being a Reading Rescue teacher was a 32% increase in the odds for every year of experience and if a teacher was delivering the Reading Rescue intervention compared with a paraprofessional, the student was 112% times more likely to attain grade-level oral reading and comprehension at the end of the intervention year (odds ratio = 2.12). Given the small number of teachers ( $n=29$ ) and that only 11 tutors in total had more than two years of experience, it is not surprising that these factors did not meet the .05 significance threshold. It is important to look at the underlying numbers when understanding the relationship; of the 87 students tutored by paraprofessionals 61% attained grade level and of the 36 students tutored by teachers 75% reached grade level. Additionally, of the 11 tutors with more than two years of experience, nine of them are paraprofessionals, suggesting these experienced paraprofessionals drove the tutor experience effect.

The same analysis was done for the San Diego sight words measure. The constant model showed a percentage correct ratio of 76.6%, a very high percentage considering most students were at grade level on the sight word measure by the end of the intervention year. The full model only increased the percentage to 77.8 with an  $R^2$  of .062, again suggesting that the student variables were more salient. Of the variables in this model, the number of years as a Reading Rescue tutor ( $B=0.446$ ,  $SE=0.212$ ) was the only significant predictor ( $p<.05$ ). Every additional year of experience increases the odds of the student reaching grade level by 56% (odds ratio = 1.56).

### **School-level factors on grade-level performance**

Last, logistic regression with school-level characteristics were run to determine if any school-level factors affected attainment of grade-level proficiency. For oral reading and comprehension, the percent of free lunch, IEP, and LM students within the school were entered in step 1, followed by the school's total enrollment number, number of first grade classes, and number of years the school has been

implementing Reading Rescue in block 2. With all predictors in the model, the percentage of predicting the correct outcome only changed from 67.8% (constant only) to 71.3%, with a  $R^2$  value of .065. Though the overall model was not significant, the total school enrollment size reached statistical significance ( $B=-0.002$ ,  $SE=0.001$ ,  $p<.05$ ), with a 1% decrease in the odds for each student enrolled (odds ratio = 0.99). The percent of students receiving free lunch was marginally significant ( $p<.12$ ); there was a 91% decrease (odds ratio = 0.09) for every point on the logged percentage of students receiving free lunch (it is important to note that the standard deviation on the log transformation was 0.19 so the differences between schools, in a numerical sense, is small). The more free-lunch students within a school, the less likely students were to attain grade-level comprehension. While the effect is pronounced, its significance is attenuated due to the high proportion of schools with a high percentage of students receiving free lunch; eight of the 38 have 100%, 21 are above 90% and only four are below 80%.

The same analysis was done to determine effects of school-level factors on sight word reading. The predictive ability of the full model did not improve beyond the initial 76.1% with a  $R^2$  value of .077, suggesting these predictors do not add much information into the prediction model. The one significant predictor in this model was the log transformation of the percentage of LM students in the school ( $B=-0.871$ ,  $SE=0.337$ ,  $p<.05$ ). Students from schools with higher percentages of LM students were less likely to be at grade level on the sight word measure, with an 58% decrease for every point on the log scale (odds ratio = 0.42,  $SD$  of log scale for LM percentage was 0.92).

### **Discussion**

Taken together, the results provide support that Reading Rescue continues to be effective in helping struggling readers succeed. Students who received the intervention made significant gains in foundational reading skills from pre- to posttest. The percentage of students at grade level for both oral reading and comprehension and sight word reading also improved dramatically. Large effect sizes for both foundational skill growth and growth in higher-order reading skills show that students improved substantially throughout the course of the intervention year. These results suggest that this intervention is an effective evidence-based approach to be used within an RTI framework. The success of this program supports Scanlon's (2011) claim that one-on-one instruction is not always an indication of a reading disability, but instead exactly the type of instruction needed to provide dramatic results in supporting students in reaching grade level in a short amount of time.

This study also enabled the investigation of various mechanisms that could have impacted the growth of students receiving the intervention. Using grade-level oral reading with comprehension scores and grade-level sight word reading status as outcomes of interest, we were able to determine which factors at the individual, tutor, and school

levels were associated with students achieving grade-level scores. The factors associated with the highest oral reading with comprehension proportions were the number of sessions received, completion of the entire intervention program (rather than exiting early), and IEP status, with number of session, program completion, and invented spelling posttest scores being associated with the highest sight word reading proportions. This suggests that the greater the dosage of the intervention, the more positive the outcomes for participating students. Further investigation is needed to explore why students with IEPs performed so well in this intervention.

Tutor- and school-level characteristics also had small impacts on the likelihood of students reaching grade-level proficiency. Students working with teachers who have delivered more Reading Rescue sessions were more likely to score at grade level on the oral reading with comprehension and sight word measures. While there were differences in the number of easing-in and -out sessions for the different types of tutors, unlike in Ehri et al. (2007), where they found a negative correlation between number of easing-in sessions and the posttest, we found no impact of the number of easing-in or out sessions on student performance. It should be noted that in the Ehri et al. study the average number of easing-in sessions (2.3) and easing-out sessions (1.3) are much lower than in the current study (5.04 and 4.19, respectively). This seems to represent a change in Reading Rescue in that they may have standardized the intake and exit procedures to make them more detailed and uniform.

Tutor training in the intervention also shows to be salient; tutors with more expertise in delivering the intervention are more effective in increasing student growth in oral reading and comprehension. This is somewhat similar to Ehri et al.'s (2007) finding that paraprofessionals took more time to ease students in and out of the program. It is different in that in the 2007 study paraprofessionals and reading specialists equally supported students in making gains in skills other than nonword decoding. While paraprofessionals as literacy tutors is not a well-studied group, there is support that they are effective in supporting literacy development (Causton-Theoharis et al., 2007; Granger & Grek, 2005), and at times they are as effective as certified teachers (Brown, Morris & Fields, 2005; Rashotte, MacPhee, & Torgesen, 2001).

In this study, there was large, albeit nonsignificant, odds ratio in favor of teachers being more effective than paraprofessionals in boosting oral reading and comprehension. However, looking more closely at the experience level of tutors, of the 11 with more than two years of experience nine of them are paraprofessionals, indicating that these paraprofessionals are driving the "tutor experience" effect. In other words, experienced paraprofessionals are effective, but in this sample most paraprofessionals were not experienced. To clarify the importance of experience, of the 16 students who were tutored by a tutor with more than two years of experience, 14 achieved grade level, and of the 11 of those 16 tutored by paraprofessionals, 100% achieved grade level. This is important, given that it is economically much more

difficult to have certified teachers doing one-on-one tutoring. Unfortunately, waiting for tutors to develop that experience is not necessarily ideal. Reading Rescue should explore additional training or prioritize the current training for paraprofessional tutors to try to close the gap in effectiveness without waiting a year or more for them to develop experience. It may be that certain topics addressed in the second year of tutor training are instrumental to their success in supporting oral reading or comprehension and therefore, that topic should be provided in the first year of training.

Though school-level factors did not contribute as much to the predictive models as individual factors did, school size, percentage of LM students, and percentage of free/reduced lunch slightly improved model prediction, though it did not reach statistical significance. Students from larger schools and students from schools with higher percentages of free/reduced lunch were less likely to reach grade-level status on reading comprehension. Students from schools with higher percentages of LM students were also less likely than were students from schools with lower percentage of LM students to reach grade level on the sight word measure. This is consistent with Kieffer and Vukovic (2012) that it is not the individual language status but the shared SES context that is influencing growth since individual LM status had no effect in the analysis. It should be noted that in the Ehri et al. (2007) study, when groups of low socioeconomic status LM students were compared against one another the LM students who received Reading Rescue training made significantly greater gains in overall reading than LM students in the two control groups, and the Reading Rescue students reached average reading levels.

The present findings suggest that increased exposure to the intervention is the optimal solution for enabling students to reach grade-level proficiency. In particular, students in large schools and schools with high percentages of free/reduced lunch and LM students would likely benefit from increased exposure to the intervention as to counteract the risk factors associated with lower success rates. Enhancing tutor training by including techniques specific to support LM literacy and providing more practice sessions with program managers before first year tutors begin instruction in these high-need schools may be beneficial as well. This may enhance first year tutor effectiveness to combat the finding that students who work with experienced tutors are more likely to reach grade-level oral reading and comprehension benchmarks. Also, it may be that students in the most vulnerable schools should remain in the program for an extended amount of time. Reading Rescue should closely examine why students exit instead of graduate the program, and the organization should consider how to maximize student reach while reducing the number of exited students.

## Limitations

There were several limitations of the study. First, the data used for the analysis were the completed data from the 2016–2017 school year. It was through this rigorous data collection process that the organization discovered the

amount of incomplete data from the easing-out sessions (i.e., posttests). As a result, the data used for the analysis were not randomly selected from the total population, which creates a dataset that is bias toward schools with thorough data collection processes in place. Also, it was outside the scope of this study to include a control group, which is instrumental in showing that student growth is attributed to the intervention and not to other factors. Future studies should continue to investigate the effects of this intervention model by randomly selecting schools who have adopted the intervention and including a control group of students who are both attending a Reading Rescue school but are not receiving the intervention and attending a non-Reading Rescue school.

Another limitation of the study is the validity and reliability of the assessments. The majority of the assessments were practitioner-created assessments that the organization had used since its inception. The assessments were used by the organization to determine pre-post growth and to inform instruction. A close investigation of the reliability scores suggests that the organization should consider adopting measures with research-based reliability and validity support. These measures will strengthen their claims about student growth in foundational and overall reading skills from pre- to postintervention. Also, the benchmarks typically provided with research-based assessments will allow the organization to compare student growth within the Reading Rescue program to national norms of performance on the literacy tasks.

### Strengths and future directions

This study shows the viability of a low-cost, data-driven, research-based intervention. Since the last evaluation, the program has drastically increased in size. Whenever programs expand there is the potential for dilution of results due to lack of fidelity. While in the present study we did not collect information on the fidelity to treatment, results indicate that the program continues to be effective in supporting students in reaching grade-level benchmarks based on the available measures. The current evaluation was based on raw data collected from the field without researcher oversight. While this can be seen as a limitation, it is also a strength in that the results are not impacted by a Hawthorne effect. This study serves as an evaluation of the authentic implementation of the program by the staff members of the Reading Rescue organization. This was a secondary data analysis so neither the students, tutors, nor the program members knew they would eventually be evaluated. While this caused a lack of complete data on the posttest measures, it allowed the researchers to evaluate the program in its most authentic state. The reality is that after researchers leave carefully designed studies, interventions are conducted in the dynamic school environment where multiple variables impact implementation and results. This study was able to capture the effectiveness of such a program in its natural state. Furthermore, pre to post effect sizes computed for oral reading and reading comprehension, word writing,

and letter knowledge were similar in magnitude as those in Ehri et al. (2007) study. This suggests that the intervention remained effective, even outside of the context of a rigorously controlled intervention study.

The Reading Rescue intervention is a valuable intervention not only because this study and others have demonstrated its effectiveness, but also because it is low cost. The program leverages a large body of school staff members, paraprofessionals and others, who are able to set aside 30 min a day to tutor a student. This is different from other much costlier programs that train outside tutors to come to the schools to provide the intervention. The Reading Rescue training provided to the school staff members builds capacity within the schools. These tutors' increased knowledge of literacy development, instruction, and assessments may impact other aspects of a school's instructional approach via tutors' conversations with classroom teachers and the way the tutors interact with other struggling students in need of literacy supports.

Reading Rescue's use of assessments to track student growth in foundational and overall reading skills is promising. The pre- and postassessments are outside of the assessment battery used in the school and therefore they provide a finer grain profile of the students' reading abilities. Daily written records of oral reading fluency are also used to assess students' decoding, fluency and comprehension ability. Performance on the preassessments and daily written records are used to inform instruction during the Reading Rescue lessons. This form of data-driven instruction is exactly the type of feedback loop necessary to pinpoint instruction for students who have fallen behind grade-level benchmarks.

Even though the previous recommendation is to transition to more valid and reliable measures, it is important to note the use of practitioner-created measures served an essential role of using data to inform instruction. While improvements to the measures may be made in the future, this does not discredit the organization's commitment to data-driven instruction over the years. Other literacy interventions across the RTI spectrum should consider how they are using assessments to inform instruction on a pre-post and daily basis. The students in these interventions do not have any time to waste on instruction that is not supporting them in making large gains in their literacy skills.

Finally, this study shows the impact of providing explicit instruction in the five components of literacy as laid out by the National Reading Panel. With an alarming number of students reading below proficiency levels in the United States (NAEP, 2017), an intervention like this with such large effect sizes in student growth from pre- to posttest, should serve as a model for other intervention models. By targeting each of the five pillars throughout the lesson plan, Reading Rescue has demonstrated impressive levels of student growth. It is only through dramatic changes in student scores over a short period of time, that students who have fallen behind in their reading scores will be able to catch-up with their peers (Foorman et al., 1997; Juel, 1988; NRP, 2000; U.S. Department of Education, Institute of Education

Sciences, National Center for Education Statistics, 2017). As the Matthew Effect states, the rich get richer and the poor get poorer (Stanovich, 1986). This is the trap that most struggling readers fall into; however, this simple 30-min intervention has proven to be effective in closing the gap between the rich and poor readers within less than one school year. Simply stated, more programs like this are needed.

## Acknowledgments

M. D. Lauterbach and D. M. Murano contributed equally to this work.

## ORCID

Katharine Pace Miles  <https://orcid.org/0000-0002-5039-8921>

## References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: The MIT Press.
- Allington, R. E. (1983). The reading instruction provided readers of differing ability. *Elementary School Journal*, 83(5), 548–559.
- Allington, R. L., McCuiston, K., & Billen, M. (2015). What research says about text complexity and learning to read. *The Reading Teacher*, 68(7), 491–501.
- Archer, A., & Hughes, C. (2011). *Explicit Instruction: Efficient and effective teaching*. New York: Guilford Publications.
- August, D. E., & Shanahan, T. (2006). *Developing literacy in second language learners: Report of the National Literacy Panel on Language Minority Children and Youth*. Mahwah, NJ: Erlbaum.
- Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., & Gersten, R. (2015). *Evaluation of response to intervention practices for elementary school reading*. Retrieved from <https://ies.ed.gov/ncee/pubs/20164000/pdf/20164000.pdf>
- Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*, 42(1), 85–95.
- Brown, K., Morris, D., & Fields, M. (2005). Intervention after grade 1: Serving increased numbers of struggling readers effectively. *Journal of Literacy Research*, 37(1), 61–94.
- Causton-Theoharis, J., Giangreco, M. F., Doyle, M. B., & Vadasy, P. F. (2007). The "sous-chefs" of literacy instruction. *Teaching Exceptional Children*, 40(1), 56–62. Retrieved from <https://login.ez-proxy.brooklyn.cuny.edu/login?url=https://search-proquest-com.ez-proxy.brooklyn.cuny.edu/docview/201107341?accountid=7286>
- Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, 35(5), 386–406.
- Cook, B. G., & Cook, S. C. (2011). Unraveling evidence-based practices in special education. *The Journal of Special Education*, 47(2), 71–82.
- Daly, E. J., Neugebauer, S., Chafouleas, S. M., & Skinner, C. H. (2015). *Interventions for reading problems: Designing and evaluating effective strategies* (2nd ed). New York: The Guilford Press.
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18(1), 5–21.
- Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, 44(2), 414–448.
- 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(4), 605–619.
- 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.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41(1), 93–99.
- Fuchs, D., & Fuchs, L. S. (2017). Critique of the national evaluation of response to intervention: A case for simpler frameworks. *Exceptional Children*, 83(3), 255–268.
- Gill, D., & Kozloff, M. (2004). *Introduction to reading first*. Wilmington: University of North Carolina, Wilmington. Retrieved from [http://people.uncw.edu/kozloffm/Introduction\\_to\\_Reading\\_First.htm](http://people.uncw.edu/kozloffm/Introduction_to_Reading_First.htm)
- Granger, J. D., & Grek, M. (2005). Struggling readers stretch their skills. *Journal of Staff Development*, 26(3), 32–36. Retrieved from <https://login.ezproxy.brooklyn.cuny.edu/login?url=https://search-proquest-com.ezproxy.brooklyn.cuny.edu/docview/211517306?accountid=7286>
- Hernandez, D. J. (2011). *Double jeopardy. How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: The Annie E. Casey Foundation.
- Hoover, N. L. (1998). *The Reading Rescue Tutor's Handbook*. New York: The Literacy Trust, Inc. (10th ed.).
- Hoover, N. L. (2016a). *Reading Rescue Training Manual: Days 1–5*. New York: The Literacy Trust, Inc.
- Hoover, N. L. (2016b). *Reading Rescue Training Manual: Days 6–10*. New York: The Literacy Trust, Inc.
- Hudson, T. M., & McKenzie, R. G. (2016). Eliminating the use of RTI to identify SLD: A survey of state, policy, procedures, data collection, and administrator perceptions. *Contemporary School Psychology*, 20(1), 31–45.
- IDEA. (2004). *IDEA (Public Law 108-446) Individuals with Disabilities Education Act (IDEA) of 2004* (Public Law 108-446).
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80(4), 437–447.
- Kame'enui, E. J., & Simmons, D. C. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading*, 5(3), 203–210.
- Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, 100(4), 851.
- Kieffer, M. J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to grade 8. *American Educational Research Journal*, 48(5), 1187–1225.
- Kieffer, M. J., & Vukovic, R. K. (2012). Components and context: Exploring sources of reading difficulties for language minority learners and native English speakers in urban schools. *Journal of Learning Disabilities*, 45(5), 433–452.
- Kilpatrick, D. A. (2015). *Essentials of assessing, preventing, and overcoming reading difficulties*. New Jersey: John Wiley & Sons.
- Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, 47(3), 596–632.
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading tests* (4th ed.). Itasca, IL: Riverside.
- McLaughlin, M. J., Speirs, K. E., & Shenassa, E. D. (2014). Reading disability and adult attained education and income: Evidence from a 30-year longitudinal study of a population-based sample. *Journal of Learning Disabilities*, 47(4), 374–386.
- Meyer, M. S., & Felton, R. H. (1999). Repeated reading to enhance fluency: Old approaches and new direction. *Annals of Dyslexia*, 49(1), 283–306.
- Moats, L. (2009). Knowledge foundations for teaching reading and spelling. *Reading and Writing*, 22(4), 379–399.
- Moats, L. (2011). Knowledge and practice standards for teachers of reading: A new initiative by the international dyslexia association. *Perspectives on Language and Literacy*, 37(2), 51–52.

- Moats, L. C. (2007). Whole-language high jinks: How to tell when "scientifically-based reading instruction" *isn't*. Washington, DC: Thomas B. Fordham Institute. <http://edexcellence.net/foundation/publication/publication.cfm?id=376>
- 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 (NRP). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Report of the subgroups*. Rockville, MD: National Institute of Child Health and Human Development Clearinghouse.
- O'Connor, R. E., Harty, K. R., & Fulmer, D. (2005). Tiers of intervention in kindergarten through third grade. *Journal of Learning Disabilities, 38*(6), 532–538.
- O'Leary, R., & Hoover, N. L. (2015). Reading Rescue®'s research base: A review of scientific studies supporting assessments and lesson components and research documenting the model's effectiveness in inner city, high poverty schools. Retrieved from <http://www.reading-rescue.org>
- Rack, J., Hulme, C., Snowling, M., & Wightman, J. (1994). The role of phonology in young children's learning of sight words: The direct-mapping hypothesis. *Journal of Experimental Child Psychology, 57*(1), 42–71.
- Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disabilities Quarterly, 24*(2), 119–134.
- Rayner, K., Foorman, B., Perfetti, C., Pesetsky, D., & Seidenberg, M. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2*(2), 31–74.
- Rosenshine, B. (2012). Principals of instruction: Research based strategies that all teachers should know. *American Educator, 36*(1), 12–19.
- Scanlon, D. M. (2011). Response to intervention as an assessment approach. In A. McGill-Franzen & R.L. Allington (Eds.), *Handbook of reading disability research* (pp. 139–148). New York, NY: Routledge.
- Scanlon, D. M., Vellutino, F. R., Small, S. G., Fanuele, D. P., & Sweeney, J. M. (2005). Severe reading difficulties: Can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality, 13*(4), 209–227.
- Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–406.
- Sweet, R. W. Jr. (2004). The big picture: Where we are nationally on the reading front and how we got here. In P. McCardle and V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 13–44). Baltimore, MD: Paul H. Brooks.
- Test, D. W., Kemp-Inman, A., Diegelmann, K., Hitt, S. B., & Bethune, L. (2015). Are online sources for identifying evidenced-based practices trustworthy? An evaluation. *Exceptional Children, 82*(1), 58–80.
- The Division for Early Childhood of the Council for Exceptional Children, National Association for the Education of Young Children, & National Head Start Association (DEC, NAEYC, NHSA). (2013). Frameworks for response to intervention in early childhood: Description and implications. Retrieved from <https://www.decdocs.org/position-statement-frameworks>
- 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*(1), 33–58.
- Tunmer, W. E., & Nicholson, T. (2011). The development and teaching of word recognition skill. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research (vol. IV)*, pp. 405–431. New York, NY: Routledge.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (NAEP). (2017). NAEP data explorer. Retrieved from [https://www.nationsreportcard.gov/reading\\_2017/#/nation/scores](https://www.nationsreportcard.gov/reading_2017/#/nation/scores)
- Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in the Schools, 47*(5), 432–444.
- Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28*(2), 178–200.
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading, 5*(3), 211–238.