Leveraging Volunteers: An Experimental Evaluation of a Tutoring Program for Struggling Readers

Robin Jacob, Catherine Armstrong, A. Brooks Bowden & Yilin Pan

Leveraging Volunteers: An Experimental Evaluation of a Tutoring Program for Struggling Readers

Robin Jacob[a], Catherine Armstrong[b], A. Brooks Bowden[c], and Yilin Pan[c]

ABSTRACT

This study evaluates the impacts and costs of the Reading Partners program, which uses community volunteers to provide one-on-one tutoring to struggling readers in under-resourced elementary schools. The evaluation uses an experimental design. Students were randomly assigned within 19 different Reading Partners sites to a program or control condition to answer questions about the impact of the program on student reading proficiency. A cost study, using a subsample of six of the 19 study sites, explores the resources needed to implement the Reading Partners program as described in the evaluation. Findings indicate that the Reading Partners program has a positive and statistically significant impact on all three measures of reading proficiency assessed with an effect size equal to around 0.10. The cost study findings illustrate the potential value of the Reading Partners program from the schools' perspective because the financial and other resources required by the schools to implement the program are low. Additionally, the study serves as an example of how evaluations can rigorously examine both the impacts and costs of a program to provide evidence regarding effectiveness.

Reading skills are the key building blocks of a child's formal education. Yet, the national statistics on literacy attainment are profoundly distressing: two out of three American fourth graders are reading below grade level and almost one third of children nationwide lack even basic reading skills (National Center for Education Statistics [NCES], 2013). For children in low-income families, the numbers are even more troubling, with 80% reading below grade level (NCES, 2013). Despite several decades of educational reform efforts, only incremental progress has been made in addressing this reading crisis. From 1998 to 2013, the number of low-income fourth graders reading at a proficient level increased by only seven percentage points (NCES, 2013). Although there are a variety of interventions to help struggling elementary school readers, many of the programs with the strongest evidence base are both time- and resource-intensive, and as such may not always be viable options for already under-resourced schools (Hollands et al., 2016; Simon, 2011). Since the 2007 recession, school and district policymakers have been faced with tight budget constraints (Leachman & Mai, 2014). Therefore, it is important that rigorous research provide not only evidence of

CONTACT Robin Jacob ✉ rjacob@umich.edu ▣ 330 Packard Street, Ann Arbor, MI 48104-2910, USA.
[a]University of Michigan, Ann Arbor, Michigan, USA
[b]MDRC, New York, New York, USA
[c]Teachers College, Columbia University, New York, New York, USA
*Catherine Armstrong is now a graduate student at the Harvard Graduate School of Education.

effects, but also information about costs, so that decision makers can enact reform as efficiently as possible. From the perspective of a school, tutoring by volunteers has the potential to be a low-cost option for addressing children's literacy needs because most of the costs are borne by volunteers. To date, only limited evidence exists regarding the efficacy of this approach on a wide scale.

This study evaluates the impacts and costs of the Reading Partners program, which uses community volunteers to provide one-on-one tutoring to struggling readers in under-resourced elementary schools. Reading Partners currently serves more than 8,500 students in over 160 schools throughout California, Colorado, Maryland, New York, Oklahoma, South Carolina, Texas, Washington, and Washington, DC. The evaluation uses an experimental design, in which students were randomly assigned within 19 different Reading Partners sites to a program or control condition, to answer questions about the impact of the program on student reading proficiency. In addition, data were collected to provide information about the context in which the program was implemented and about the degree to which the program was implemented with fidelity. The evaluation also includes a cost study, using a subsample of six of the 19 study sites, to explore the resources needed to implement the Reading Partners program as described in the evaluation.

Background

Several national initiatives (including the Reading Excellence Act of 1997 and Reading First, which was established as part of the No Child Left Behind Act of 2001) have been instituted in recent years in an attempt to provide schools with better resources to support struggling readers in elementary school. In addition, a number of states have either enacted or are considering enacting legislation regarding third-grade reading proficiency, with at least 14 states and Washington, DC, having adopted laws that require students to be retained if they are not reading at grade level by the end of third grade (Rose, 2012).

The Elementary and Secondary Education Act (ESEA) of 2002 requires that schools provide low-income families with extra academic assistance or supplemental education services if their children are attending a Title I school that is in Program Improvement.[1] These supplemental education services come at no cost to families and can include tutoring or remedial help in subjects such as reading, language arts, and math, as long as they are aligned with state content standards and grounded in high-quality research that provides evidence of their effectiveness (Heinrich, Meyer, & Whitten, 2010). Often, external programs and vendors provide this extra help (U.S. Department of Education, 2012).

Many supplemental education services use Response to Intervention (RtI) models, which provide a means of identifying and supporting struggling students. Funding for RtI was authorized under the Individuals with Disabilities Education Improvement Act of 2004. Under this model, all students are monitored and assessed for placement into a multitiered system of intervention. Tier I, usually classroom instruction for all students, is considered the first level of intervention, with Tier II and Tier III being progressively more intense (Fuchs & Fuchs, 2006). Daily one-on-one work with a specially trained reading

---

[1] Title I schools are those that receive federal funds based on the number of low-income children in attendance. A school in Program Improvement has been designated by the state to be in need of improvement for more than one year (U.S. Department of Education, 2012).

interventionist, for example, is considered a Tier III intervention. Reading Partners is typically considered a Tier II intervention for students needing some additional support and is usually one among multiple Tier II interventions available for struggling readers at any given school.[2]

Prior Research on Supplemental Literacy Programs

A variety of supplemental interventions exist to help elementary school students who are struggling to learn to read, including small-group work inside or outside the classroom, instructional technology, and one-on-one tutoring led by teachers, paraprofessionals, or volunteers. Among these, one-on-one tutoring by trained teachers has been shown to be among the most effective. In their meta-analysis of early reading interventions, Slavin, Lake, Davis, and Madden (2011) report a weighted mean effect size among 20 qualifying studies of 0.39, and a recent large-scale evaluation of Reading Recovery, one of the most widely known of these tutoring programs, found effects sizes between 0.42 and 0.68 standard deviations for the first graders participating in the program (May et al., 2013; May et al., 2015). Slavin et al. also find that tutoring by paraprofessionals and small-group tutorials, in which struggling readers are provided with supplementary instruction in small groups daily for 30–45 minutes, are both effective approaches to reading intervention as well, with weighted effect sizes of 0.38 and 0.31, respectively.

Despite their effectiveness, such interventions are typically expensive for schools to implement. For example, reports on Reading Recovery indicate that the total cost of that program is about $4,360 to $9,180 per student, depending on the implementation.[3] Small-group tutorials such as Corrective Reading and Wilson Reading System have been shown to cost $10,640 and $7,050 per student, respectively (Hollands et al., 2016; Simon, 2011). Tutoring by paraprofessionals is less expensive. For example, a cost study of the Sound Partners program, which served kindergarten students once a day for 30 minutes a day, four days a week was estimated to cost $791 per student. However, in this study the paraprofessionals were only able to serve a handful of students in each school because they have other duties at the school to which they have to attend. Hiring enough paraprofessionals to serve large numbers of students would require increased coordination and likely increased costs. Almost all the costs of these supplemental programs were borne entirely by the schools (Simon, 2011).

Research on Volunteers as Tutors

This study focuses on the use of volunteers, as opposed to trained teachers or paraprofessionals, to provide one-on-one tutoring to struggling readers. Tutoring by volunteers has the potential to be a less costly method for addressing students' literacy needs because the resources are provided in large part by the volunteers, rather than being financed entirely by

---

[2] A 2011 Reading Partners survey asked principals to indicate how Reading Partners was used in their schools. Of the 32 principals who answered the question about whether they used Reading Partners as an RtI intervention in their schools, 81% indicated that they did. Of the 26 who answered a question about what tier they used Reading Partners tutoring services for, 50% said they used Reading Partners as a Tier II intervention, and another 35% said they used Reading Partners as a Tier II/III intervention.

[3] The cost estimates are obtained from Hollands et al. (2016) and Simon (2011). They are adjusted to 2012 dollars in order to compare with the cost of Reading Partners reported in this article.

the school or district. Furthermore, using volunteers provides an opportunity to serve large numbers of students, since the number of students that can be served is only limited by the space and the number of volunteers that can be recruited. Although researchers have been evaluating the effectiveness of volunteer tutoring programs for several decades (Elbaum, Vaughn, Hughes, & Moody, 2000; Shanahan, 1998; Topping & Hill, 1995; Wasik, 1998; Wasik & Slavin, 1993), a number of gaps remain in the literature.

First, very few rigorous studies of volunteer tutoring programs have utilized large samples (i.e., over 100 students). A review by Ritter, Barnett, Denny, and Albin (2009) identified 24 randomized trials of volunteer tutoring programs that targeted reading, of which only one involved a sample of more than 100 children; most included fewer than 30. Similarly, Slavin et al. (2011) report on seven studies of volunteer tutoring programs of which only two had both an experimental design and an intervention group larger than 100 students, and among these studies, the results were mixed. One found a negative effect, while the other, a study of the Experience Corps program, found positive effects equal to 0.11 standard deviations.

Prior research on the implementation and effectiveness of programs that rely largely on volunteers generally suggests that bringing them to scale might be quite difficult (Grossman & Furano, 1999; Hager & Brudney, 2004). Thus, research on the scalability of these programs is needed to better understand whether volunteers can be an effective option for helping struggling readers.

The research on volunteer tutoring programs also offers little insight into the potential effectiveness of these programs for students in the upper elementary grades (i.e., Grades 4–6), where the approach to reading instruction is often different than in the lower elementary grades. In Grades 2 and 3, reading instruction typically focuses on learning to read, with more attention given to basic reading skills such as decoding and phonics. Starting in Grade 4, the focus is on reading to learn, with greater emphasis on vocabulary and comprehension (Chall, 1983). Of the studies reviewed by Ritter et al. (2009), only two included students beyond Grade 3. They both found no statistically significant positive impacts on student reading proficiency. Among the seven studies of volunteer tutoring programs reviewed by Slavin et al. (2011), only two included students in Grades 4–5; neither found statistically significant positive effects. More research is needed to understand whether volunteer tutors can be used effectively with students in the upper grades.

Finally, the costs of volunteer tutoring programs have not been examined in detail. Yet, it is their potential for providing services at a low cost that may make them an appealing option for schools.

Research Questions

To address some of these gaps, this study was designed to answer the following research questions:

- Can a volunteer tutoring program, such as Reading Partners, be implemented with fidelity when delivered at scale?
- On average, did the Reading Partners program have a positive impact on students' reading proficiency across three components of early reading ability: sight-word efficiency, reading fluency, and comprehension? Did the impacts differ by grade level, gender, incoming ability, English language learner status or prior receipt of Reading Partners' tutoring?

- What resources were needed to implement the Reading Partners program as described in this evaluation and what proportion of the costs of implementing the program were borne by the school? How did these resource requirements compare with the resources required to implement other supplemental reading services in these schools?

The Reading Partners Program

Reading Partners is a "pull-out" program, meaning that students are pulled out of their regular classrooms or after-school programming for a limited time period to meet with their tutors. The program operates both during the school day and after school in designated reading centers at each partner school. The program aims to serve students in kindergarten through Grade 5 who are half a year to two-and-a-half years behind grade level in reading, who are conversationally fluent in English, and who generally do not have special needs (i.e., do not have a documented disability through an Individualized Education Program). This evaluation includes only students in Grades 2–5. The Reading Partners program consists of twice-weekly, one-on-one tutoring sessions that last 45 minutes each. Students are tutored by community volunteers who do not need to have any prior experience working in education or with children. Tutoring sessions are overseen by site coordinators—full-time Reading Partners team members who are usually AmeriCorps members and who are also responsible for managing the day-to-day operations of the program.[4] Tutors are recruited by volunteer coordinators who serve multiple sites and whose primary role is to recruit volunteers on a year-round basis to ensure that sites have an adequate supply of volunteers.

The Reading Partners model is designed to include six core components: (a) regular, one-on-one tutoring, (b) dedicated school space and materials, (c) a structured and individualized curriculum, (d) data-driven instruction, (e) rigorous ongoing training, and (f) instructional supervision and support.

Regular, One-on-One Tutoring

The Reading Partners model is designed to provide students with individualized reading instruction. Tutors meet with students one-on-one, not in small groups.

Dedicated School Space and Materials

Reading Partners tutoring is designed to take place on school grounds in a dedicated space, or "reading center." The reading centers typically contain a "read-aloud" library where students and tutors can select books at different levels of difficulty, a corner of the room with couches or bean bags for read-aloud sessions, work stations for tutor–student pairs, a "take-home reading" area where students can select books to read at home, and reading-themed materials on the walls. The components and materials are designed to create a learning environment in which multiple tutor–student pairs can be working at the same time.

---

[4] AmeriCorps is a program of the Corporation for National and Community Service that places young adults in service positions at nonprofit organizations, schools, public agencies, and faith-based entities. In return for their service, AmeriCorps members receive a living stipend, health insurance, and, when they complete their program, an education grant (Corporation for National and Community Service, 2014). Reading Partners has been an AmeriCorps program since 2010.

Structured and Individualized Curriculum

At the time the study took place, the curriculum consisted of two modules: a beginning readers' module (consisting of one level) and a comprehension readers' module (consisting of four levels). The beginning readers' module (Level 1) contained 50 lessons, which were primarily phonics-based and covered various letter-sound combinations and practice with high-frequency words. The comprehension readers' module comprised Levels 2, 3, 4, and 5, each of which consisted of 24 lessons. The lessons in this module focused on topics such as cause and effect, making inferences, fact and opinion, summarizing, and predicting.

Each lesson begins with the tutor reading aloud from a text of the student's choice. During that time, the tutor models fluent reading, paying close attention to tone, pronunciation, and punctuation. Tutors pause to ask the student comprehension questions or to check the student's knowledge of key vocabulary. After the read-aloud, the tutor works hand-in-hand with the student to complete the Reading Partners curriculum materials (e.g., a worksheet). These materials either introduce a new skill or concept or reinforce a skill or concept that has been taught previously. Finally, the student applies the new skill or concept while reading aloud from text provided to support the lesson with support from the tutor.

Data-Driven Instruction

Reading Partners uses data to match implementation to student needs and to support the model. After the teacher or principal refers the student to Reading Partners, a staff member administers an assessment of reading skills. At the time the study took place, Reading Partners was using the Rigby PM Benchmark assessment. This assessment tool measures oral fluency, accuracy, and comprehension using a dedicated set of reading materials at different levels of difficulty. A trained administrator asks the student to read aloud and measures accuracy by recording any errors or corrections that the student makes. The student is then asked comprehension questions about the main idea of the story, primary characters, setting, and other salient features of the text. Oral fluency is assessed by asking the student to participate in a second read-aloud that is timed for one minute. Students' performance on this assessment is used to place them into the appropriate lesson in the Reading Partners curriculum.

The same assessment is administered again midway through the school year and once more at the end of the school year. Each student's progress is documented in an Individualized Reading Plan (IRP), which contains assessment scores from the beginning, middle, and end of the school year; reading goals; and areas for improvement. After the midyear assessment, Reading Partners staff use the data to update the IRP, identify areas for growth, and adjust student placement where needed. For students who are struggling or are not making gains during the first half of the school year, a targeted action plan is developed. Students' assessment data and progress are also shared with teachers and principals; some teachers also share these data with parents when report cards are distributed.

Rigorous and Ongoing Training

Reading Partners school-based staff and AmeriCorps members participate in two to three weeks of organized training sessions before the school year begins as well as ongoing (usually monthly) sessions throughout the school year. Topics covered in these training sessions include the Reading Partners model, a detailed review of the curriculum, how to train volunteer tutors, and how to use data and the Reading Partners data system.

For volunteer tutors, the primary training is a "shadow session," which is a combination of orientation and tutoring observation led by the site coordinator. The entire training takes less than one hour, and then tutors begin working with their own students. There are additional optional training sessions offered approximately twice a month by site coordinators on topics such as how to better support English language learners or managing students' behavior and attention.

Instructional Supervision and Support

Instructional supervision and support are provided on an ongoing basis by and for Reading Partners staff. Site coordinators supervise volunteers during tutoring sessions and provide guidance and suggestions to tutors who need additional support. Program managers—more experienced staff members who generally have a background in teaching—work with site coordinators to troubleshoot a range of issues, including those related to communications with school staff, managing tutors, or identifying best practices to better support the progress of a specific student.

Methods

To assess program impacts, the study team used an experimental design, in which students were randomly assigned within schools to either a program group or control group. At the start of the school year, schools and Reading Partners staff developed a list of students who were deemed eligible for Reading Partners based on a combination of test score review and teacher and/or staff recommendation. Reading Partners staff then sought parental consent for those students to participate in the program and the study. A total of 1,657 students were recruited, and 1,265 consented to participate. Students with consent were randomly assigned within schools and grade groups (Grades 2 through 3 and Grades 4 through 5) to either the program group, which was eligible for Reading Partners services, or to an "as is" control group, which was eligible to receive only business-as-usual supplemental reading services provided by the school. Random assignment was stratified at the school and grade-group level to account for the specific number of program spots available at each site and to ensure that equal numbers of upper and lower elementary students were represented in the sample. Random assignment was conducted by using a random number generator to assign a random number to each child in each school and grade-group level. Students were assigned to the treatment group in order from highest to lowest random number, until all eligible spots within a grade were filled.

To assess the costs of the Reading Partners program, the study team used a method developed by Levin (1975) in which all the ingredients that were used in implementing the Reading Partners program were identified, described qualitatively, and assigned a market price. Ingredients were specified based on the requirements for replication, regardless of who paid for or contributed them (Levin & McEwan, 2001). The ingredients included personnel and the amount of time they devoted to providing services, training requirements, facilities used to provide services and training, materials, and any transportation or food or other services that were provided. The same method was used to assess the costs of implementing the other supplemental reading services offered in the cost study sample of schools. The cost study adheres to recent guidance on cost-effectiveness analysis by (a) examining the costs of various sites, (b) conducting the cost study concurrently with the impact evaluation, (c) utilizing

national prices to control for differences among locations, and (d) examining the costs of the "business as usual" control condition (Levin & Belfield, 2015).

School Sample

The sample for this study consists of 19 schools with 17 distinct Reading Partners reading centers that were recruited to participate in the evaluation (two pairs of schools shared a

Table 1. Characteristics of reading partners study schools and other school samples (2011–2012).

Characteristic	Study schools	Other Reading Partners schools[a]			U.S. Title I schools[b]			Average U.S. schools[c]		
		Mean	Difference	P value	Mean	Difference	P value	Mean	Difference	P value
Eligible for Title I program (%)	88.89	96.20	−7.31	0.209	100.00	−11.11***	<0.001	75.39	13.49	0.184
Students eligible for free/reduced-price lunch[d] (%)	81.70	84.52	−2.82	0.411	69.06	12.65***	0.008	53.47	28.23***	<0.001
Race/ethnicity (%)										
Black	20.74	28.44	−7.71	0.304	20.74	0.00	1.000	15.92	4.82	0.404
Hispanic	61.75	52.95	8.80	0.204	29.83	31.92***	<0.001	23.34	38.40***	<0.001
Asian	9.60	8.30	1.29	0.639	3.43	6.16***	0.002	4.67	4.92**	0.025
White	5.70	7.28	−1.58	0.496	41.28	−35.58***	<0.001	51.60	−45.90***	<0.001
Other	2.22	3.03	−0.81	0.336	4.73	−2.51	0.263	4.47	−2.25	0.234
Male (%)	51.89	51.82	0.07	0.970	53.54	−1.64	0.462	53.00	−1.1	0.564
Average number of students	494	451	43	0.318	468	26	0.622	472	22	0.677
Grade 2	79	69	10	0.169	68	10	0.217	70	8	0.331
Grade 3	76	65	11	0.142	68	8	0.384	70	5	0.554
Grade 4	74	65	9	0.227	67	7	0.396	70	5	0.591
Grade 5	76	64	12	0.131	67	9	0.290	69	6	0.488
School setting[e] (%)										
Urban	73.68	91.67	−17.98**	0.027	39.77	33.91***	0.003	31.77	41.91***	<0.001
Suburban	26.32	8.33	17.98**	0.027	21.28	5.03	0.592	29.50	−3.19	0.761
Town	0.00	0.00	0.00	NA	10.07	−10.07	0.145	8.79	−8.79	0.176
Rural area	0.00	0.00	0.00	NA	28.87	−28.87***	0.005	29.94	−29.94***	0.004
Sample size	19	84			24,914			42,202		

Source. Calculations from 2011 and 2012 National Center for Education Statistics Common Core of Data (CCD).

Notes. Sample sizes for individual outcomes may fall short of the reported sample sizes because of missing or unusable data. A two-tailed *t* test is used for all statistical tests presented in this table. Statistical significance levels are indicated as follows:***= 1%;**= 5%;*= 10%.

Rounding may cause slight discrepancies in sums and differences.

[a]"Other Reading Partners Schools" include all other nonstudy Reading Partners schools that meet the "Average U.S. Schools" criteria below.

[b]"U.S. Title I Schools" include all non-Reading Partners schools that meet the "Average U.S. Schools" criteria below and all that were designated Title I schoolwide schools.

[c]"Average U.S. Schools" include non-Reading Partners schools that offer Grade 2 through Grade 5, are defined as "regular" schools by the CCD, and are located within the 50 U.S. states and the District of Columbia.

[d]The value given for students eligible for free/reduced-price lunch is calculated from the 2011 CCD due to missing data in the 2012 CCD. Data for all other variables are from the 2012 CCD.

[e]"Urban" is defined as territory inside an urbanized area and inside a principal city having a population greater than 100,000. "Suburban" is defined as territory outside a principal city and inside an urbanized area with a population of less than 250,000. "Town" is defined as territory inside an urban cluster that is (a) less than or equal to 10 miles from an urbanized area ("fringe"), or (b) more than 10 miles and less than or equal to 35 miles from an urbanized area ("distant"), or (c) more than 35 miles from an urbanized area ("remote"). "Rural" is defined as territory that is (a) less than or equal to 5 miles from an urbanized area and less than or equal to 2.5 miles from an urban cluster ("fringe"), or (b) more than 5 miles but less than or equal to 25 miles from an urbanized area and more than 2.5 miles but less than or equal to 10 miles from an urban cluster ("distant"), or (c) more than 25 miles from an urbanized area and more than 10 miles from an urban cluster ("remote").

reading center). To be eligible to participate in the study, a school's Reading Partners program had to have been operating for at least one year prior to the start of the study, and staff had to anticipate that the program would not have the capacity to serve all the school's struggling readers, ensuring that random assignment would not deny services to students who would otherwise have been served by the school. As shown in Table 1, the sites are largely representative of Reading Partners programs more broadly, though slightly less urban and more suburban, and represent six different programmatic regions served by the organization (New York City; Washington, DC; and, in California, East Bay, South Bay, Sacramento, and Los Angeles). Table 1 also compares the study sites with all federally designated Title I schools and all elementary schools in the United States. Consistent with the Reading Partners model, the Reading Partners schools in the study sample are poorer, more urban, and have higher percentages of minority students than the population of schoolwide Title I schools and U.S. schools as a whole.

For the cost study, a subsample of six schools was chosen based on three criteria: geographic representativeness with all study sites, strong program implementation, and availability of reliable data on the resources used in Reading Partners and other supplemental services offered during the 2012–2013 school year. These sites were slightly larger and composed of slightly more White students than the rest of the evaluation sample but were otherwise comparable.

Student Sample

A total of 1,265 students in the 19 participating schools were randomly assigned. The final respondent sample (those who took at least one follow-up assessment) includes 1,166 students (594 students in the program group and 572 in the control group). These totals represent response rates of 92.0% for the program group and 92.4% for the control group, which were not statistically significantly different from one another. Descriptive statistics for the final sample at baseline are shown in Table 2. More than half of the sample was Hispanic (65%), and one-fifth of the students were Black (19%). About half were officially designated as English language learners (55%), and over 90% of the students were eligible for free or reduced-price lunch. Aside from a slightly larger number of students in the program group who had previously participated in the program (32% compared with 26% of the control group), there were no statistically significant differences between the program and control groups on these variables. After correcting for multiple hypothesis testing, using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995), the difference in program participation is not statistically significant.

Data Collection

To assess the impact of Reading Partners on student reading achievement, independent assessors conducted three separate baseline reading assessments (described in detail below) for all students in the sample as soon as feasibly possible after random assignment.[5] These same assessments were administered at follow-up, as close to the end of

---

[5] In order to retain tutors and ensure that students were able to receive services in a timely manner, tutoring for some students began immediately following random assignment. Baseline test was conducted as soon as feasibly possible after random assignment, but some students were tested after tutoring had begun. However, there were no differences in the baseline test scores of the students in the program and control groups (as shown in Table 2).

Table 2. Baseline characteristics of program and control group students.

Characteristic	Program group	Control group	Difference	P value
Male (%)	54.88	54.50	0.38	0.899
Ethnicity (%)				
Black	19.39	19.03	0.37	0.864
Hispanic	65.09	65.27	−0.18	0.944
Asian	8.43	9.09	−0.66	0.691
White	5.90	5.40	0.50	0.709
Other	1.18	1.02	0.16	0.856
Special education (%)	11.62	10.49	1.13	0.550
English language learner (%)	55.21	55.50	−0.29	0.916
Eligible for free/reduced-price lunch (%)	91.35	90.22	1.13	0.540
Prior Reading Partners participation (%)	31.82	25.75	6.07[*]	0.074
Average age (years)	8.79	8.76	0.03	0.413
Overage for grade[a] (%)	9.83	8.40	1.43	0.408
Average baseline achievement scores[b]				
Comprehension	574.74	573.49	1.26	0.544
Sight word efficiency	91.15	90.38	0.77	0.328
Fluency	0.03	0.00	0.03	0.573
Students by grade level (%)				
Grade 2	23.06	25.29	−2.23	0.256
Grade 3	29.46	27.23	2.23	0.256
Grade 4	24.58	25.72	−1.14	0.572
Grade 5	22.90	21.76	1.14	0.572
Sample size	594	572		

Sources. Calculations using study team-administered baseline and follow-up SAT-10, TOWRE-2, and AIMSweb scores; district-provided demographic data; Reading Partners management information system data.

Notes. Calculations for this table used data for all students in the respondent sample. Sample sizes for individual outcomes may fall short of the reported sample sizes because of missing or unusable data.

The model allows estimates to vary randomly across schools and controls for the random assignment block.

A two-tailed *t* test was applied to differences between research groups. Statistical significance levels are indicated as:
  ***= 1%; **= 5%; *= 10%.

Rounding may cause slight discrepancies in sums and differences.

[a]Students are classified as "overage" for a certain grade if they were the following specified ages on September 1, 2012: 8 or older for second grade, 9 or older for third grade, 10 or older for fourth grade, and 11 or older for fifth grade.

[b]The analyses for student reading achievement were conducted using scaled scores for the SAT-10 (comprehension) and the TOWRE-2 (sight word efficiency) and sample-normed scores for the AIMSweb (fluency). Sample-normed AIMSweb scores range between 0 and 1.

the school year as possible. These assessments were supplemented with teacher surveys, in which teachers provided information on the amount of reading instruction each student received in class and through supplemental services. Teachers were also asked to rate students' academic performance and behavior, which were used as secondary outcomes for the study.

During the winter of the study year, the research team members visited each school in the study to collect implementation data. These site visits consisted of a series of semistructured interviews with school staff (a principal and, if available, reading coordinator or specialist), Reading Partners staff (including the site coordinator, program manager, and outreach coordinator, among others) and volunteer tutors. During visits, the research team also reviewed a random sample of student folders for completeness. These folders are used to document what occurred during each tutoring session. Finally, the research team compared the reading center setup in each site to the requirements in the Reading Partners Operations Manual and observed a sample of tutoring sessions to better understand the flow of the sessions and to observe the dynamics between tutors and students.

During the summer of 2013, additional information was collected for the cost study to supplement information that the team had already learned from program data and site visits. The cost team used structured interviews with school, Reading Partners, and other supplemental services staff to identify a comprehensive list of all resources required to implement the Reading Partners program as well as those resources required to implement the other supplemental reading services that were available to struggling students.

Measures

Reading Assessments

The primary outcome measures for the evaluation include the Stanford Achievement Test, 10th Edition (SAT-10) reading comprehension subtest, The Test of Word Reading Efficiency, 2nd Edition (TOWRE-2) and the AIMSweb oral reading fluency measure. A range of possible reading assessments were reviewed in consultation with reading experts against the following criteria: appropriate for students in Grades 2 through 5, able to detect small improvements among lower-scoring students, and feasible to administer in a relatively short amount of time. In addition, the assessments were selected because they cover multiple domains of reading and are reliable, valid, and well-aligned with the curricular emphasis of Reading Partners. Although the assessments do not include a direct measure of students' decoding skills, this skill set is implicitly measured because decoding skills are prerequisite for both fluency and comprehension.

The SAT-10 subtest is a 30-minute group assessment that measures students' comprehension across four different modes: initial understanding, interpretation, critical analysis, and awareness and usage of reading strategies. Students read short passages and then answer questions about them. Kuder-Richardson 20 reliability on this measure ranges from 0.84 to 0.93.

The TOWRE-2 sight word efficiency subtest is a measure of an individual's ability to pronounce printed words accurately and fluently. The assessment takes approximately five minutes, is administered individually, and assesses the number of real words printed in vertical lists that a student can accurately identify within 45 seconds. Alternate-form reliability on this measure ranged from 0.90 to 0.92.

The AIMSweb oral reading fluency measure is an individually administered one-minute assessment of oral reading fluency that measures accuracy and speed of reading. Students read a passage aloud for one minute, and the assessor records omissions of words, misreading, and substitution of words. The count of the number of words read correctly is used as the outcome. On the forms used during the study, alternate-form and split-half reliability on this measure each ranged from 0.93 to 0.95, and inter-rater reliabilities were 0.99. For the SAT-10 and TOWRE-2, the provided conversions were sufficient for analysis. However, the AIMSweb percentile conversions do not take into account the differing levels of difficulty across the different forms of the test. To standardize this outcome, a z-score was calculated for each student's baseline and outcome scores using the baseline mean and standard deviation of scores for that grade-level and test form.

Surveys

Teacher surveys were used to collect teacher ratings of student academic behavior and performance in the classroom. Teacher ratings of reading performance, math performance, and overall performance were each treated as separate outcomes. For academic behavior, four

individual items (attentiveness, assignment completion, homework completion, and disruptiveness) were combined to create a single composite measure. This was done by reverse-coding each of the items (except disruptiveness) and then summing the scores, with higher scores representing better composite behavior. The possible score ranged from a low of 4 to a high of 16, and Cronbach's alpha for this composite measure was 0.82.

Administrative Data

The research team also collected a variety of extant administrative data to inform the analyses. Reading Partners provided management information system (MIS) data that included information on the number of Reading Partners tutoring sessions the program students received, the tutors with whom they were assigned to work, and the characteristics of those tutors. Additionally, the team collected district records containing demographic information (including gender, age, race/ethnicity, free/reduced-price lunch status, ELL status, IEP status) and student attendance for the 2011–2012 and 2012–2013 school years.

Analysis

The study was designed to assess whether the Reading Partners program was delivered with fidelity to the model and whether, on average, the Reading Partners program had a positive impact on students' reading proficiency across the three components of early reading ability that were measured. It was also designed to assess whether impacts differed by gender, grade, incoming ability, English language learner status, or prior receipt of Reading Partners and to determine both the total cost and the cost to schools of implementing the program. The following section describes the analytic approach to answering these questions.

Assessing Fidelity of Implementation

Implementation fidelity was assessed along the six core components of the Reading Partners model, as described above: (a) regular, one-on-one tutoring; (b) dedicated school space and use of materials; (c) structured and individualized curriculum; (d) data-driven instruction; (e) rigorous and ongoing training; and (f) instructional supervision and support. To help summarize and quantify the overall level of implementation fidelity across all 19 schools, the study team developed a fidelity index based on five of the six core components. The third component listed above—structured and individualized curriculum—was not incorporated into the index because the data collected were not appropriate for quantifying fidelity along this dimension. This component is discussed qualitatively.

The fidelity index contains 17 items that were created using quantitative data from the Reading Partners' MIS; qualitative assessments of fidelity based on interviews with Reading Partners staff, school staff, and volunteer tutors; a review of student folders that are kept on site in reading centers; and a facilities checklist, which assessed whether the reading centers contained the required components. Criteria for each item were developed by the study team in collaboration with Reading Partners. Each study school was scored across all 17 items. Two members of the research team and the principal investigator reviewed all scoring decisions. Fidelity was assessed using data from the entire program year.

Estimating Impacts

For each outcome, the model used in the analysis to estimate impacts was as follows:

$$Y_{ijk} = \gamma_0 Y_{-1,ijk} + \beta_0 T_{ijk} + \sum_k \sum_j \gamma_{1kj} B_{ijk} + \sum_S \gamma_{2s} X_{sijk} + u_k T_{ijk} + \varepsilon_{ijk} \qquad (1)$$

where:

$T_{ijk} =$ 1 if student i from grade j in school k is assigned to Reading Partners and 0 otherwise

$Y_{-1,ijk} =$ the pretest score for student i from grade j in school k before random assignment

$B_{ijk} =$ block dummy, 1 if student i from grade j is in a particular random assignment block, defined by grade block (2/3 or 4/5) and school k, and 0 otherwise

$X_{sijk} =$ the Sth other student-level covariates for student i from grade j in school k; these variables include the time lapse between baseline and follow-up testing as well as the demographic characteristics of age, gender, race, poverty status, special education status, and English language learner status

$u_k T_{ijk} =$ a random error term for the treatment group in school k, assumed to be independently and identically distributed; this term is included in the model to allow for variation in the impact of the program across schools

$\varepsilon_{ijk} =$ a student-level random error for student i from grade j in school k, assumed to be independently and identically distributed

In this model, the treatment effect is measured by β_0, the coefficient on the treatment indicator. The full set of covariates was included in all analyses to control for any unobserved differences between the program and control groups at baseline and to improve the precision of the estimates.

In the case of covariate missing values, imputation was conducted using the mean of the school and grade level to which the student belonged. The overall rate of missing data was quite small. However, if more than 5% of observations were missing data for a given covariate, the model includes an indicator variable to indicate whether a student's data were imputed for that covariate. Missing data were not imputed for any outcome measures.

Cost Study Methods

The cost study examines the total social cost of implementing the Reading Partners program and other supplemental reading services based on a comprehensive list of the resources (or "ingredients") used in implementation and analyzes who bears the burden of financing those costs. Total cost, or the cost to replicate the implementation, was calculated by summing the cost of all the resources needed to implement the program, regardless of who financed them or if the ingredient was provided in-kind.[6]

One important ingredient that was provided in-kind was volunteer time to provide tutoring to the students. To properly document the full resource richness needed to replicate the

---

[6] To calculate the annual value of ingredients that last longer than one year, such as books, computers, or desks, the item was amortized using a 3% discount rate.

program, these donated hours were valued at the price that the services would cost if they had to be purchased.[7] This full-cost valuation allows an appropriate comparison of the resource richness of programs that receive donated time with those that do not. Average national prices were used to control for market variation among site locations.

The total social cost of Reading Partners is presented as the cost per program group student. To calculate the cost per program group student at each school AC_k, the cost of each tutoring session (which is the total cost of serving all the in-study and out-of-study RP students at that school TC_k, divided by the total number of tutoring sessions received by these students TS_k) was multiplied by the average number of sessions received by the students participating in the study AS_k, as shown in Equation (2). The average cost per program student across the six sites in the cost study sample AC is the weighted sum of cost per program student at each site, with weights w_k proportional to site-level scale N_k. The analyses focus on the portion of the total social costs borne by the school, including the financial fee paid for the program as well as the opportunity costs of staff time, space, and materials.

$$
\begin{aligned}
AC &= \sum_{k=1}^{K} w_k AC_k \\
&= \sum_{k=1}^{K} w_k \frac{TC_k}{TS_k} AS_k \\
where \ w_k &= \frac{N_k}{\sum_{i=1}^{K} N_k}
\end{aligned}
\tag{2}
$$

The costs of the other supplemental reading services were calculated in the same way as Reading Partners, with a similar focus on the portion of costs borne by the school. To protect confidentiality, the costs of the other services were averaged at the site level. The cost per student represents the weighted mean of the average cost per student of all supplemental programs at each school, with the weights proportional to the number of students served by each program. Some of these averages combine relatively resource-rich interventions, such as a learning specialist who works intensively with small groups of children, with interventions that require fewer resources, such as computer-based programs overseen by reading specialists.

Results

Fidelity of Implementation

Overall, the fidelity index reveals that implementation fidelity across all 19 schools was relatively high. As shown in Figure 1, the fidelity scores for the 19 schools range from a low of 12 to a high of 22; the maximum possible score on the index is 23. Two sites had scores below 15.

Further detail on fidelity of implementation is shown in Table 3. The table shows that, on average, students in the study received approximately 1.5 tutoring sessions per week and spent 28 weeks in the Reading Partners program. The reduction from 2 sessions to 1.5

---

[7] All prices were expressed in 2012 U.S. dollars. Volunteer time was valued at the national average price of a part-time paraeducator. The results were robust to sensitivity testing that reflected different wages related to the education level of the tutors or valuing all at minimum wage. All prices are available from the authors.
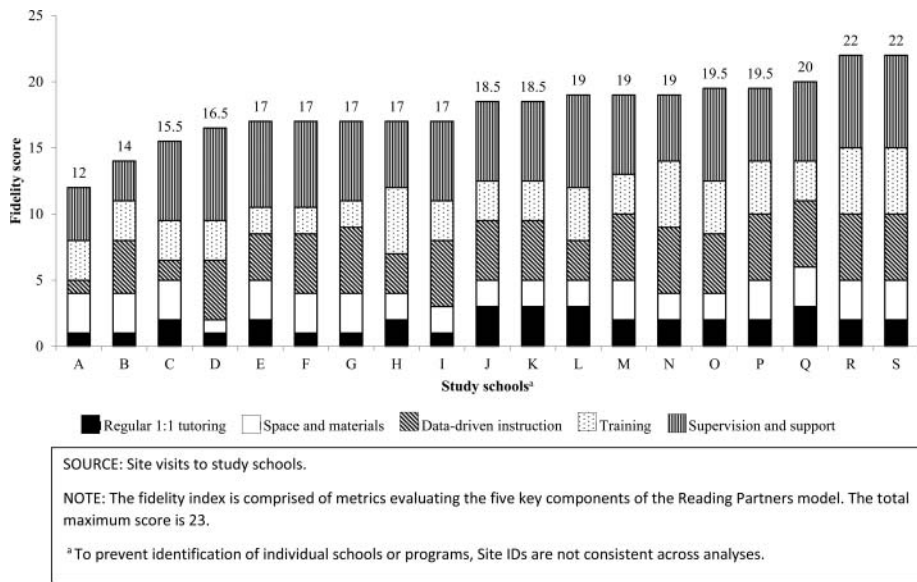
Figure 1. Fidelity scores of study schools.

sessions can largely be accounted for by student absences, school closures, field trips, etc. Although the level of intensity measured is slightly less than the program model recommends, on average it means that students received three tutoring sessions every two weeks. Overall attendance at tutoring sessions was around 80%. All sites in the study had a designated reading center where tutoring took place and where selected materials and resources were made available to program staff, tutors, and students. All students were tutored individually.

MIS data from Reading Partners indicate that 95% of the program group students who participated in the Reading Partners program for the entire year were assessed by Reading Partners staff,

Table 3. Average student's experience in Reading Partners.

		School-level averages	
Program characteristic	Program group	Minimum	Maximum
Number of sessions per week	1.55	1.11	1.76
Length of participation in program (weeks)	28.13	24.24	32.01
Student attendance rate[a] (%)	78.76	55.75	88.98
Number of tutors assigned	2.52	1.67	3.60
Duration of each tutoring relationship (weeks)	19.81	11.20	26.01
Scheduled sessions per week with primary tutor[a]			
Scheduled once per week (%)	76.38	39.58	91.24
Scheduled twice per week (%)	23.62	8.76	60.42
Sample size	594		

Source. Calculations from Reading Partners management information system data.
Notes. Calculations for this table used data for all program group students in the respondent sample. Sample sizes for individual outcomes may fall short of the report sample size because of missing or unusable data.
[a]Student attendance is calculated as the number of sessions attended divided by the number of possible sessions, assuming two sessions every week that school was in session after the student entered Reading Partners.
[b]Among students assigned tutors, primary tutors were those who had the longest assigned relationship with the student.

using the Rigby PM assessment, at the three prescribed points that the Reading Partners model dictates. A review of student folders also indicated that tutors were generally consistent in maintaining records of student progress through the Reading Partners curriculum.

Staff and volunteers reported in interviews that they had the training and support they needed to perform their jobs successfully. For Reading Partners staff and AmeriCorps members, preparation included a comprehensive, multi-day training covering all the key components of the Reading Partners program and staff responsibilities. Although the volunteers received limited training before they began tutoring (approximately an hour of orientation and a tutoring observation), and most tutors did not attend the bi-monthly optional training sessions that were offered, most tutors indicated that they felt adequately trained for the role. They also felt well supported by the site coordinators, who provided monitoring and assistance during tutoring sessions as well as additional feedback on how to address specific issues with students outside of the tutoring session. The full-time Reading Partners staff and AmeriCorps members (including program managers and site coordinators) consistently indicated that they had access to ongoing support from their supervisors.

The biggest challenge that Reading Partners faced in implementing the program was ensuring tutor attendance and retention. Reading Partners requests that tutors make a one-semester commitment, and site coordinators at the study sites reported that many volunteers, particularly high school and college students, did not stay beyond that period of time. As shown in Table 3, the average duration of the tutoring relationship was around 19 weeks, and in some schools the average tutor stayed only 11 weeks. As a result, new tutors had to be brought on throughout the year. This meant that students were often assigned several different tutors over the course of the school year (on average 2.5 tutors over the course of the study year). In some schools, the average was as high as 3.6 tutors per student (see Table 3).

Furthermore, volunteers varied in their consistency and commitment. Site coordinators reported that tutors sometimes failed to arrive at their scheduled time and at times did not notify the site coordinator beforehand. However, Reading Partners put structures in place to address these problems, including the use of substitute tutors, make-up days, and tutoring sessions conducted by site coordinators. This meant that a student did not necessarily miss tutoring sessions as a result of tutor inconsistency, and the overall tutoring rate was relatively high.

Service Contrast

Because reading is a critical component of elementary education, it was unlikely that schools would allow struggling readers who were not assigned to the Reading Partners program to continue to struggle without providing some additional support. Moreover, given the limited resources available to schools, it is possible that once students were involved with Reading Partners, the school would determine them less in need of other supplemental reading services. Therefore, the first step in our analyses was to determine whether and in what ways the program group's reading services differed from those of the control group.

As shown in Figure 2, program group students received 48 minutes of additional instruction in reading each week relative to the control group. Both groups received about five hours of in-class group instruction in reading and a little less than one hour of in-class one-on-one reading instruction each week. Instruction outside of the classroom, however, differed between the program and control groups. Control group students received, on average, two hours of non-Reading Partners supplemental services, whereas program group students
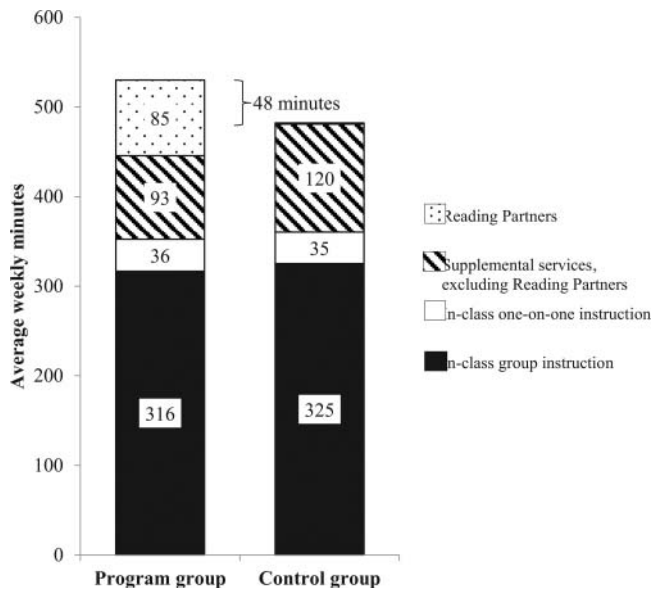
Figure 2. Time spent in reading instruction and supplemental services. Sources: Calculations using study team-administered teacher survey; district-provided demographic data; study team-administered baseline SAT-10 scores; Reading Partners management information system data. Notes: Calculations for this table used data for all students in the respondent sample with teacher survey data. Sample sizes for individual outcomes may fall short of the reported sample sizes because of missing or unusable data. The analysis model allows estimates to vary randomly across schools and also controls for the random assignment block, the student-level pretest measure, the time lapse between baseline and follow-up testing, and student-level demographic covariates. The sample size for the program group is 580; the sample size for the control group is 554.

received closer to three hours of supplemental services, including both Reading Partners and other services provided to both the program and control groups.

The other supplemental services students received included several types of programs: unstructured before- or after-school assistance and support focused on the completion of teacher-assigned work; one-on-one tutoring programs focused on reading (Reading Partners and similar programs); small-group, usually pull-out services for struggling readers conducted by a school reading specialist or other educator; technology-based programs in which the student worked on literacy skills using a computer or tablet; and other reading programs that did not fall into the preceding categories.

As shown in the figure, both groups were receiving a substantial amount of supplemental reading instruction. As shown in Table 4 among the program group, one-on-one tutoring was the type of service received most frequently (95% received Reading Partners and/or other one-on-one tutoring); this percentage was significantly different from the 21% of control group students who received one-on-one tutoring services. However, control group students were significantly more likely to receive small-group intervention support than the program group (32% compared with 23%). Around 10% of the program and control groups received each of the other types of services offered to students, at rates that were not significantly different from each other. Given the menu of programs offered to both groups, the impacts reported

Table 4. Reading instruction received.

Outcome	Program group	Control group	Difference	P value
Students receiving any supplemental service (%)	97.24	64.75	32.49[***]	<0.001
Homework help	11.21	10.86	0.35	0.844
One-on-one tutoring[a]	95.17	20.81	74.36[***]	<0.001
Small-group intervention support	22.59	31.53	−8.94[**]	0.014
Technology-based programs	8.28	9.95	−1.68	0.279
Other programs	7.24	7.48	−0.23	0.647
Sample size	580	554		

Sources. Calculations using study team-administered teacher survey; district-provided demographic data; study team-administered baseline SAT-10 scores; Reading Partners management information system data.

Notes. Calculations for this table used data for all students in the respondent sample with teacher survey data. Sample sizes for individual outcomes may fall short of the reported sample sizes because of missing or unusable data.

The analysis model allows estimates to vary randomly across schools and also controls for the random assignment block, the student-level pretest measure, the time lapse between baseline and follow-up testing, and student-level demographic covariates.

A two-tailed *t* test was applied to differences between research groups. Statistical significance levels are indicated as:
[***] = 1%; [**] = 5%; [*] = 10%.

Rounding may cause slight discrepancies in sums and differences.

[a]One-on-one tutoring includes the receipt of Reading Partners and other one-on-one supplemental services.

below should be interpreted as the impact of Reading Partners relative to other supplemental service receipt, not the impact of Reading Partners compared with no intervention.

Program Impacts

The results shown in Table 5 represent the primary outcomes of interest in this study—namely, the impact of the program on reading comprehension, sight word efficiency, and fluency. They show that overall the program had a positive and statistically significant impact on all three measures of students' reading proficiency, with effect sizes equal to 0.10 for reading comprehension (measured by the SAT-10), 0.11 for sight word efficiency (measured by the TOWRE-2), and 0.09 for fluency (measured by the AIMSweb). These impacts remain statistically significant after controlling for multiple hypothesis testing using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). The impact in reading comprehension is equivalent to approximately one-and-a-half to two months of additional progress in reading relative to the control group over the course of a nine-month school year.

To explore whether or not Reading Partners was more effective for some groups of students than for others, the study team conducted subgroup analyses based on five prespecified characteristics of interest: baseline achievement (based on the fall reading comprehension assessment), prior receipt of Reading Partners services, gender, English language learner status, and grade level. The results are shown in Table 6. These analyses suggest that the Reading Partners had positive impacts for a wide variety of students because impacts did not differ significantly for students from different grade levels or baseline achievement levels, for male or female students, for those who were still learning English, or for those who had previously been served by the Reading Partners program.

There were also positive and statistically significant impacts for both upper and lower elementary students. The pattern was consistent with the nature of reading instruction in those grades—there were statistically significant positive impacts on sight word efficiency for second and third graders (effect size equal to 0.17) and statistically significant positive impacts on reading comprehension for fourth and fifth graders (effect size equal to 0.13). To put these

Table 5. Primary impacts of Reading Partners on reading proficiency.

Outcome	Program group	Control group	Difference	Standard error	Impact effect size	P value
Comprehension	592.42	588.94	3.48**	1.71	0.10	0.043
Sight word efficiency	92.78	91.37	1.42***	0.50	0.11	0.004
Fluency	0.06	−0.03	0.09**	0.04	0.09	0.031
Sample size	594	572				

Sources. Calculations using study team-administered baseline and follow-up SAT-10, TOWRE-2, and AIMSweb scores; district-provided demographic data; Reading Partners management information system data.

Notes. Calculations for this table used data for all students in the respondent sample. Sample sizes for individual outcomes may fall short of the reported sample sizes because of missing or unusable data.

The impact analyses for student reading achievement were conducted using scaled scores for the SAT-10 (comprehension) and the TOWRE-2 (sight word efficiency) and sample-normed scores for the AIMSweb (fluency). Sample-normed AIMSweb scores range between 0 and 1. The model allows estimates to vary randomly across schools and controls for the random assignment block, the student-level pretest measure, the time lapse between baseline and follow-up testing, and student-level demographic covariates.

A two-tailed *t* test was applied to differences between research groups. Statistical significance levels are indicated as: *** = 1%; ** = 5%; * = 10%.

Rounding may cause slight discrepancies in sums and differences.

effect sizes in perspective, Hill, Bloom, Black, and Lipsey (2008) report that students in second grade typically grow 0.60 of a standard deviation over the course of an academic year, while third through fifth graders show an average annual reading gain of somewhere between 0.32 and 0.40 standard deviations over the course of a single school year, depending on the grade.

One other pattern that is worth mentioning is the subgroup analysis based on incoming reading achievement. This analysis indicates that the sight word efficiency and oral reading fluency impacts were largest for students who performed in the lowest quartile of the study sample on the baseline reading comprehension test, with effect sizes equal to 0.22 and 0.19, respectively. In other words, the program group students who were lowest-performing at the beginning of the year performed the best on the sight word efficiency and fluency measures compared with their counterparts in the control group. Reading comprehension impacts were largest for those in the second-lowest quartile, with an effect size equal to 0.16. Although not statistically significantly different from one another, these findings suggest that the program may have been more effective for the students who were lowest-performing at baseline.

Although not shown here, the study team also investigated the impact of the program on students' academic behavior and their teacher-reported performance. Although the impacts were mostly positive, none were statistically significant.

Program Costs

The total cost of Reading Partners is approximately $3,610 per program group student (see Table 7). However, the schools themselves paid only a small proportion of these costs through their contributions of both in-kind resources and a fee to finance the program. The in-kind resources included a classroom for the reading center, space for program events, and a small amount of time from the principal, teachers, administrative staff, and the school's reading specialist or coordinator to help identify eligible students, valued at approximately $390 per program group student. In addition to space and time, the schools also covered a portion of the resources provided by Reading Partners by paying a fee of $320 per program group student to Reading Partners. Thus, the school's total resource commitment was $710 per student, or 20% of the total resources required to implement the program.

Table 6. Subgroup analysis of primary impacts.

	Sample size	Comprehension		Sight word efficiency		Fluency	
		Impact effect size	*P* value	Impact effect size	*P* value	Impact effect size	*P* value
Subgroup							
Prior achievement							
1st quartile	310	0.02	0.839	0.22**	0.030	0.19*	0.080
2nd quartile	305	0.16*	0.056	0.08	0.295	0.05	0.538
3rd quartile	291	0.12	0.258	0.11	0.276	0.08	0.348
4th quartile	260	0.00	0.975	0.14	0.106	0.10	0.300
Bottom 3 quartiles	906	0.11**	0.043	0.10**	0.022	0.09**	0.040
Prior receipt of							
Reading Partners							
Prior receipt	316	0.11	0.178	0.05	0.629	0.03	0.722
No prior receipt	812	0.07	0.142	0.12***	0.010	0.11*	0.060
Gender							
Male	633	0.08	0.179	0.13**	0.014	0.07	0.212
Female	522	0.11*	0.070	0.09	0.137	0.14**	0.023
English language							
learner (ELL) status							
ELL	634	0.11**	0.048	0.11**	0.041	0.14**	0.015
Non-ELL	503	0.01	0.845	0.11*	0.061	0.04	0.577
Grade level							
Grades 2 and 3	605	0.09	0.132	0.17***	0.002	0.09	0.182
Grades 4 and 5	561	0.13*	0.096	0.05	0.479	0.10	0.101

Sources. Calculations using study team-administered baseline and follow-up SAT-10, TOWRE-2, and AIMSweb scores; district-provided demographic data; Reading Partners management information system data.

Notes. Calculations for this table used data for all students in the respective subgroups of the respondent sample. Sample sizes for individual outcomes may fall short of the reported sample sizes because of missing or unusable data.

The impact analyses for student reading achievement were conducted using scaled scores for the SAT-10 (comprehension) and the TOWRE-2 (sight word efficiency) and sample-normed scores for the AIMSweb (fluency). The model allows program impact estimates to vary randomly across schools and controls for the random assignment block, the student-level pretest measure, the time lapse between baseline and follow-up testing, and student-level demographic covariates.

A two-tailed *t* test is used for all statistical tests presented in this table. Statistical significance levels between the program and control groups are indicated as:

*** = 1%; ** = 5%; * = 10%. Differences between subgroups are not statistically significant.

Over half of the remaining $2,900 in resources are provided by in-kind donations from the volunteer tutors. Community volunteers contributed, on average, the equivalent of $1,520 per program group student, which included both their time and transportation costs, and amounts to 42% of the total value of the program.

The Reading Partners organization financed about 31% of the required resources ($1,110 per program group student), which included both staff and materials. AmeriCorps provided a grant to Reading Partners to finance some of the costs of the AmeriCorps members (the site coordinator and outreach coordinator positions). AmeriCorps funding finances about 7% of the total cost of Reading Partners.

Costs of Other Supplemental Services

The average cost per student of the other supplemental reading services provided to both program and control students was $1,780 per student. Among school sites, this cost ranged from $1,050 to $4,890 per student (see Table 8). Unlike Reading Partners, most of the costs of other supplemental reading services were funded solely by the schools or school districts. As a result, the average portion of the costs of the other supplemental reading programs financed by the school was 96% (66% to 100% at each site). By contrast, the portion of the

Table 7. Cost of Reading Partners per program group student.

Cost	Distribution of cost per student				
	Cost per student	Cost to school	Cost to volunteers	Cost to Reading Partners	Cost to AmeriCorps
Ingredients					
Reading Partners staff	690			690	
AmeriCorps members	930			930	
School staff	90	90			
Volunteer time and transportation	1,520		1,520		
Facilities	300	300			
Materials and equipment	80				80
Total cost of ingredients per student ($)	**3,610**	**390**	**1,520**	**1,700**	**0**
Fee for service[a] ($)		320		−320	
AmeriCorps grant[b] ($)				−270	270
Net cost per student		**710**	**1,520**	**1,110**	**270**
(Total cost + fee for service + AmeriCorps grant) ($)					
Portion of net cost per student (%)		20	42	31	7

Source. Calculations from cost data.
Notes. Costs are measured in 2012 U.S. dollars and are rounded to the nearest 10 in the tables and text.
Averages are weighted to account for the size of the program group at each site.
All costs are those for the respondents in the program group and are adjusted for the larger dosage they received.
Rounding may cause slight discrepancies in sums and differences.
[a]Each school pays a fee to Reading Partners to finance a portion of resources provided by the organization, such as staff and materials. The fee presented here is the national average fee per student.
[b]AmeriCorps provides a grant to Reading Partners to finance approximately 30 percent of the costs of the AmeriCorps members employed by the program. Reading Partners pays the remainder of the costs.

costs borne by the school for Reading Partners was 20% on average (15% to 26% at each site). Thus, the average school contribution for Reading Partners was $710 per student while the average contribution by the school for the other supplemental services was $1,700.

It is important to note that the range in the costs of the other supplemental services depended on both the number (between one and four different programs in each school) and type of services provided. The services varied considerably in the resources they required. For example, computer-based programs tend to use staff time, lab or classroom space, curriculum materials, and equipment to provide the services. The costs of such programs were relatively low (on average, around $940 per student in the study sites) because the computers were also used for other purposes, and one staff member could monitor a large group of students at once. Alternatively, a full-time school support staff person (such as a reading specialist) who worked with small groups of students several times each week in a dedicated classroom tended to be more expensive (around $3,200 per student on average).

Costs of Reading Partners at Scale

The cost of Reading Partners and the variability in costs among schools is largely tied to scale—or the number of students served by the program. A sensitivity test was included to adjust for scale by utilizing the operational maximum number of students based on the space and equipment recommended by Reading Partners for each reading center. The program recommends eight spaces for tutor–student pairs to read and work together during each session. The reading centers tended to serve students for five class periods for four days a week. Accordingly, each site's costs were adjusted to ascertain what costs would be if 80 students

Table 8. Site-level costs of Reading Partners and other supplemental services.

Site[a]	School contribution per student[b] ($)		Total resources per student ($)		Contribution provided by school (%)	
	Reading Partners	Other supplemental services[c]	Reading Partners	Other supplemental services	Reading Partners	Other supplemental services
Site A	690	1,840	3,450	1,840	20	100
Site B	520	1,850	3,420	2,230	15	83
Site C	940	2,680	3,570	2,680	26	100
Site D	1,270	1,040	5,190	1,050	25	99
Site E	660	1,310	4,210	1,980	16	66
Site F	480	4,890	2,740	4,890	17	100
Pooled	710	1,700	3,610	1,780	20	96

Sources. Calculations from cost data.
Notes. Costs are measured in 2012 U.S. dollars and are rounded to the nearest 10 in the tables and text.
Averages have been weighted to account for the size of the program group (for Reading Partners) and students served at each site (for other supplemental services).
All the Reading Partners costs in this table are those for the respondents in the program group and are adjusted for the larger dosage they received. The costs for supplemental services are those for all the students served at each site.
[a]To prevent identification of individual schools or programs, site letter identifiers are not consistent across analyses.
[b]Costs borne by school include the fee for services paid by schools.
[c]Other supplemental reading services include small-group support from reading specialists, tutoring programs, homework help, and supplemental curricula.

were served. The results showed that the cost per student drops to $2,390 on average ($2,230 to $2,700). Because the costs borne by the school were assumed to be fixed (up to 80 students), the portion of the total resources used to implement Reading Partners borne by the school dropped from about 20% to 14%. This suggests that some schools could reduce costs by serving more students.

Discussion

The evaluation found positive impacts across all three different domains of reading skills that were assessed: fluency, sight word efficiency, and comprehension. The magnitude of the impacts (around 0.10 standard deviations) is smaller than the average impacts of tutoring by teachers, paraprofessionals, and small group tutorials reported by Slavin et al. (2011), all of which were in the range of 0.35 standard deviations. However, they are consistent with the findings from other studies evaluating tutoring by volunteers. Slavin et al. (2011) report an average weighted effect size for volunteer tutoring programs of 0.16. The one other large-scale randomized evaluation of a volunteer tutoring program, an evaluation of the Experience Corps program, also found impacts of around 0.10 standard deviations.

The evaluation also finds that the program has positive impacts for a wide range of students: males and females, English language learners, and students with varying levels of incoming abilities. The positive and statistically significant finding on the reading comprehension scores for fourth- and fifth-grade students is worth underscoring because there have been no randomized trials of volunteer tutoring programs that have found a statistically significant impact on students in the upper grades.

The findings of the cost study point to the potential value of the Reading Partners program from the perspective of the schools. On average, the program costs schools $710 per student, of which almost half represents the contribution of classroom space. The out-of-pocket cost for the schools is $320 per student. The program itself requires a very limited investment of time from

school personnel—only the identification of students to be served and periodic meetings to share student progress—because the recruitment, scheduling, and oversight of the volunteers are coordinated by Reading Partners staff. When examined in the context of the costs of other types of supplemental reading programs schools typically provide, which in our sample averaged $1,700 per student, the results suggest that Reading Partners could be added to the suite of existing reading intervention services available at a school at a relatively low cost. This could enable schools both to serve substantially more students (Reading Partners served, on average, around 30 students per school in this study and has the potential to serve up to 80 students per school) and to redirect other, more intensive interventions to students most in need.

To implement its model, Reading Partners relies on both fund-raising by the Reading Partners organization and on funding from AmeriCorps. This funding helps to ensure that the model can be offered to schools at a very low cost per student. Others trying to implement a similar program would need to take these costs into account when assessing the cost of the program to schools. If these additional costs are included (the costs that are subsidized by Reading Partners and Ameri-Corps), the cost per student increases to $2,090 per student, which is higher than the current cost per student to schools for Reading Partners ($710 per student), but lower than the average cost of small-group intervention services in our sample of schools ($3,200 per student).

Although Reading Partners increased students' reading skills by about a month-and-a-half to two months relative to the control group using its current structure, strengthening parts of the model could result in a stronger tutoring program and, as a result, larger impacts for students. These improvements could include, for example, increasing the amount of formal training for volunteer tutors, working to reduce volunteer turnover and absenteeism, making improvements to the curricular materials, or increasing the intensity of the program by tutoring students more frequently. The program could also consider hiring and training paraprofessionals, who might be more consistent tutors and might remain in schools for many years. As described below, more research is needed to identify which of these improvements would be most likely to help increase the impacts of the program.

Limitations

There are limitations to the conclusions that can be drawn from the findings presented here and some questions that remain about the quality and importance of particular components of the Reading Partners program. First, the implementation data collected during the study emphasized a structural understanding of fidelity and how closely Reading Partners sites adhered to the national model. Tutor observations were conducted to gain a more thorough understanding of the structural components of a tutoring session. Although the research team observed differences in how tutors explained concepts and modeled reading practices, no attempt was made to evaluate the quality of particular tutoring approaches or styles. Therefore, the question remains as to whether higher quality tutors and tutoring would result in larger impacts and what changes would need to be implemented in order to see higher quality tutoring. Similarly, assessing the quality of the Reading Partners curriculum was beyond the scope of this project. Aside from knowing that the curriculum is highly structured and (according to Reading Partners) increasingly aligned with the Common Core State Standards, the evaluation did not assess whether the individual lessons are effective at teaching students specific literacy

skills and concepts. The only conclusion that can be drawn is that it is designed in a way that allows tutors with minimal experience or training to successfully complete progressive sessions.

Second, the evaluation may have affected the findings by changing the equilibrium of services being provided to struggling readers in the study schools. In the absence of Reading Partners, which brings additional resources to the school in the form of volunteer tutors, supplemental services would have had a different distribution across the student population. Small-group intervention services or other one-on-one tutoring programs represent the most common types of supplemental service provided to students in the control group. These services are typically costly with regard to staff time and place a burden on school resources (the cost study indicates on average they cost schools $3,200 per student). On the other hand, the cost of Reading Partners to schools is relatively low ($710 per student of which half is the cost of the classroom space), meaning that more students can be served for the same amount of money. Thus, it is possible that absent Reading Partners, the "as-is" condition at study schools would not have been able to serve as many students, thus decreasing the overall number of students receiving supplemental instruction at those schools. As a result, our estimates may underestimate the true impact of the program.

Third, the evaluation only assessed students at the end of one year of services. Reading intervention programs with long-lasting effects are more useful to schools trying to improve reading skills among their students. Follow-up studies on these students in subsequent grades would indicate whether Reading Partners is one such program.

Finally, the cost analysis introduced some assumptions. First, the national average wage of a paraeducator was used to value volunteer time. This wage was selected because it best matched the duties of a Reading Partners volunteer and would likely be the wage paid to someone to provide the service if necessary. This approach, however, does not reflect the true wages lost by the volunteers based on their qualifications. If the volunteers were especially qualified or held qualifications that contributed to the impact, those qualities are not reflected in the cost estimate. Second, the cost study relied upon a subsample of sites due to budgetary and time constraints. The sites included in the cost study were chosen partly based on fidelity of implementation, and thus the costs may not generalize to the other sites in the study. Lower fidelity sites may have lower implementation costs.

Conclusions

Overall, the evaluation finds that the Reading Partners program produced measurable impacts on reading skills among students who had a fairly broad range of baseline reading abilities and among students from a wide range of grades (Grades 2 through 5). The cost study findings illustrate the potential value of the Reading Partners program from the schools' perspective because the financial and other resources required by the schools to implement the program are low. Thus, Reading Partners could be added to the suite of intervention services available at a school at a relatively low cost, allowing schools to better target their resources.

This evaluation suggests several areas for future research. First, more information regarding tutor consistency and absences would help the program providers evaluate whether or not the lack of consistency reported by site coordinators is impeding student progress. To the extent that it is problematic, the Reading Partners program could consider putting structures in place to help mitigate this problem.

Second, more information about the quality of the tutor–student relationship would help Reading Partners staff identify program strengths and weaknesses and might surface specific areas in which more training is warranted. Relatedly, research on factors that increase attendance at tutor training might help Reading Partners devise more creative ways to deliver training to their community volunteers. Gathering more information about tutor background and its relationship to student outcomes could also help Reading Partners target its recruitment activities more effectively.

In addition, more information about the quality and effectiveness of the curricular materials themselves is warranted, since the curriculum may be related to observed outcomes. Finally, future studies should examine whether or not impacts are sustained beyond a single year.

Funding

ARTICLE HISTORY

EDITORS

This article was reviewed and accepted under the editorship of Carol McDonald Connor and Spyros Konstantopoulos.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A new and powerful approach to multiple testing. *Journal of the Royal Statistical Society (Series B, Methodological)*, *57*(1), 289–300. doi:10.2307/2346101

Chall, J. S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.

Corporation for National and Community Service. (2014). *Our programs: AmeriCorps*. Retrieved from http://www.nationalservice.gov/programs/americorps

Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (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. doi:10.1037/0022-0663.92.4.605

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. doi:10.1598/RRQ.41.1.4

Grossman, J. B., & Furano, K. (1999). Making the most of volunteers. *Law and Contemporary Problems*, *62*(4), 199–218.

Hager, M. A., & Brudney, J. L. (2004). *Balancing act: The challenges and benefits of volunteers*. Washington, DC: The Urban Institute.

Heinrich, C. J., Meyer, R. H., & Whitten, G. (2010). Supplemental education services under No Child Left Behind: Who signs up, and what do they gain? *Educational Evaluation and Policy Analysis*, *32*(2), 273–398. doi:10.3102/0162373710361640

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*, 172–177. doi:10.1111/j.1750-8606.2008.00061.x

Hollands, F. M., Kieffer, M., Shand, R., Pan, Y., Shand, R., Cheng, H., & Levin, H. M. (2016). Cost-effectiveness analysis of effective reading programs: A demonstration with recommendations for future research. *Journal for Research on Educational Effectiveness*, *9*, 30–53. doi:10.1080/19345747.2015.1055639

Leachman, M., & Mai, C. (2014). *Most states funding schools less than before the recession*. Retrieved from http://www.cbpp.org/research/most-states-still-funding-schools-less-than-before-the-recession

Levin, H. M. (1975). Cost-effectiveness analysis in evaluation research. In M. Guttentag & E. L. Struening (Eds.), *Handbook of evaluation research* (Vol. 2, pp. 89–122). Beverly Hills, CA: Sage Publications.

Levin, H. M., & Belfield, C. (2015). Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness*, *8*, 400–418. doi:10.1080/19345747.2014.915604

Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.

May, H., Goldsworthy, H., Armijo, M., Gray, A., Sirinides, P., Blalock, T. J., & Sam, C. (2015). *Evaluation of the i3 Scale-Up of Reading Recovery: Year two report, 2012–13* (CPRE Research Report #RR-79). Retrieved from http://cpre.org/rryr2

May, H., Gray, A., Gillespie, J., Sirinides, P., Sam, C., Goldsworthy, H., & Tognotta, N. (2013). *Evaluation of the i3 Scale-Up of Reading Recovery: Year one report, 2011–12* (CPRE Research Report #RR-76). doi:10.12698/cpre.2013.rr76

National Center for Education Statistics. (2013). *National assessment of educational progress reading: Grade 4 national results*. Retrieved from http://www.nationsreportcard.gov/ltt_2012/age9r.aspx#1-0

Ritter, G., Barnett, J., Denny, G., & Albin, G. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research*, *79*(1), 3–38. doi:10.4073/csr.2006.7

Rose, S. (2012). *Third grade reading policies*. Denver, CO: Education Commission of the States.

Shanahan, T. (1998). On the effectiveness and limitations of tutoring in reading. In P. Pearson & A. Iran-Nejad (Eds.), *Review of research in education* (pp. 217–234). Washington, DC: American Educational Research Association.

Simon, J. (2011). *A cost-effectiveness analysis of early literacy interventions*. Unpublished manuscript, Columbia University, New York, NY.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, *6*(1), 1–26. doi:10.1016/j.edurev.2010.07.002

Topping, K. J., & Hill, S. V. (1995). University and college students as tutors for schoolchildren: A typology and review of evaluation research. In S. Goodlad (Ed.), *Students as tutors and mentors* (pp. 13–31). London, England: Kogan Page.

U.S. Department of Education. (2012). *Description of supplemental educational services*. Retrieved from http://www2.ed.gov/nclb/choice/help/ses/description.html

Wasik, B. A. (1998). Using volunteers as reading tutors: Guidelines for successful practices. *The Reading Teacher*, *51*(7), 562–570. Retrieved from http://www.jstor.org/stable/20201962

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*, 179–200. doi:10.2307/747888