Title: Exploring the Implementation, Effectiveness and Costs of the Reading Partners Program

Authors and Affiliations: Robin Jacob, University of Michigan
Dean Elson, Reading Partners
Brooks Bowden, Teachers College
Catherine Armstrong, MDRC

Abstract

Background / Context:
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. For children in low-income families, the numbers are even more troubling, with 80 percent reading below grade level. 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 (U.S. Department of Education, National Center for Education Statistics, 2013). While there are a variety of interventions to help struggling readers in elementary school, many of the programs with the strongest evidence base are both time- and resource-intensive approaches, and as such may not always be viable options for already under-resourced schools. Tutoring by volunteers or paraprofessionals has the potential to be a more cost-effective method for addressing the problem of low literacy, but to date, little rigorous evidence exists regarding the efficacy of this approach on a wide scale.

Purpose / Objective / Research Question / Focus of Study:
This study reports on an evaluation of the Reading Partners program, which uses community volunteers to provide one-on-one tutoring to struggling readers in under-resourced elementary schools. In the years since its inception, Reading Partners has grown to serve more than 7,000 students in over 150 schools throughout California, Colorado, Maryland, New York, Oklahoma, South Carolina, Texas, Washington, and Washington, DC. To answer questions about the effectiveness of the Reading Partners program and the potential for volunteer tutoring more generally to help improve the reading proficiency of struggling readers, the evaluation included an implementation study, an impact study (in which students were randomly assigned within schools to receive Reading Partners or to a control condition) and a cost study. The implementation and impact studies included 19 Reading Partners sites in three states and the cost study included a sub-sample of six of these sites. Together, these three facets of the evaluation are designed to address the following broad research questions:

1. In what context was the Reading Partners program implemented and was it implemented as intended or with fidelity? How much variability in fidelity of implementation was observed across the sites? What factors contributed to any observed variability?

2. On average, did the Reading Partners program have a positive impact on students' reading proficiency across three key components of early reading ability: sight word reading efficiency, reading fluency and comprehension?

3. What resources are needed to implement the Reading Partners program as described in the evaluation and what proportion of the costs of implementing the program are borne by the school?

Setting:
The sample for the study consists of 19 Reading Partners schools that were recruited to participate in the evaluation. Only schools in which Reading Partners had been in operation for a

least one year prior to the start of the study were eligible to participate, although the staff in these schools included both new and returning staff members. The 19 participating schools were spread across 12 school districts. There were 16 schools in California, two in New York, and one in Washington D.C. All of the schools were established Reading Partners sites; eight of the 19 schools were in their second year of operating, while the rest had been operating a Reading Partners center for 3 or more years. The schools in the Reading Partners study sample had a high number of students receiving free or reduced price lunch, high numbers of minority students, and high numbers of English Language Learners. The majority of the participating schools were school-wide Title I schools. Eight of the study schools were in varying stages of school improvement and two were in the final year of a three-year School Improvement Grant (SIG).[1] Two of the participating schools underwent complete restructuring several years ago, meaning that almost all school staff, including the principal, were hired at that time. Table 1 shows the characteristics of the study sample compared to other Reading Partners schools that were not included in the sample, to all school-wide Title I schools in the US, and to all elementary schools in the US. The sample of sites included in the evaluation is broadly representative of Reading Partners schools across the country.

Population / Participants / Subjects:
A total of 1265 2-5th grade students in the 19 schools participated in the evaluation. The final respondent sample (those that took at least one follow-up assessment) included 1166 students. This represents a response rate of 91.9 percent for the program group and 92.4 percent for the control group. The response rates of the treatment and control groups were not statistically different from one another. Descriptive statistics for the final sample are shown in Table 2. On average, more than half of the sample was Hispanic (~65 percent) and one fifth of the students were black (~19 percent). About half were officially designated as English Language Learners (~56 percent) and over 90 percent of the students were eligible for free or reduced price lunch. Aside from slightly larger number of students in the program group who had previously participated in the program (32 percent of the program group compared to only 25 percent of the control group), there were no statistically significant differences between the program and control groups on these variables. Correcting for multiple hypothesis testing, this difference is not statistically significant. Regardless, a full set of covariates were included in all analyses to control for any differences in the two groups at baseline.

Intervention / Program / Practice:
At each school, Reading Partners transforms a dedicated space into a reading center, places an full-time AmeriCorps member on site, and recruits a corps of 40 to 100 community volunteers to work one-on-one with struggling readers in grades K-5 (this evaluation included only students in grades 2-5). Reading Partners is a "pull-out" program that operates both during the school day and afterschool. The highly structured, modular curriculum is delivered by volunteer tutors on a one-on-one basis in 45-minute sessions, twice a week. Students in the evaluation sample received approximately 28 weeks of tutoring over the 2012-2013 school year.

Research Design:
The implementation study included site visits to all the schools participating in the evaluation in

[1] The goal of SIGs are to turn around the nation's lowest performing schools by awarding money to state education agencies to be used to support school improvement goals.

the winter of the study year, interviews with key program staff and volunteers, and the collection of programmatic data from the sites and from Reading Partners own management information system. It was designed to enable the study team to describe the program and the context in which it was implemented in detail and to assess whether the program was delivered as intended. In addition, it was designed to explore whether there were factors that helped facilitate the implementation of the program or things that posed barriers to effective implementation, and how sites addressed those challenges.

The impact study used a student-level randomized controlled trial (RCT) design, in which students were randomly assigned within school to either participate in Reading Partners during the 2012-2013 school year or to an "as is" control condition. The students in the control group were free to receive whatever other supplemental reading services were available to students in their schools. The availability of these services varied considerably across the 19 sites.

The cost study identified the resources utilized to implement Reading Partners during the evaluation period in a sub-sample of six schools and estimated the cost of the program with a focus on the school's contributions (in-kind and total costs). The cost analysis also examined the costs of the other supplemental reading services provided in these same six schools. Sites with strong implementation, that were geographically representative of the other sites in the study and where reliable data could be collected on the resources used in Reading Partners and other supplemental reading services offered during the 2012-2013 school year were selected for the cost study. The study utilized "the ingredients method" to estimate the costs of Reading Partners and the supplemental reading services. The ingredients method works by calculating an intervention's total cost by summing the cost of all of the resources (or ingredients) necessary to implement the program regardless of who financed them (Levin, 1975; Levin & McEwan, 2001). For example, in the case of Reading Partners, ingredients or resources include books, personnel, volunteer tutors, transportation, space, etc.

Data Collection and Analysis:

To study implementation, the research team analyzed interview transcripts, notes and observation write-ups using a systematic coding process. A coding scheme was developed to align with the overall framework of the implementation study and organized to capture data in the three main categories: the context in which the program was implemented, the fidelity of implementation and implementation lessons. Team members were assigned a set of codes to apply across all of the qualitative data. Once all data had been coded, team members then provided a written synthesis that described emerging themes.

The primary outcome of interest in the impact study was student performance in reading proficiency, as measured by test scores. To measure reading proficiency three assessments were administered in the fall and spring of the study year: a group administered reading comprehension test (the SAT-10), an individually administered sight word efficiency assessment (TOWRE2), and an individually administered oral reading fluency assessment (AIMSweb). To test the impact of the program on the outcomes of interest, the impact analysis pools together the sample of students across schools and compares outcomes for students receiving Reading Partners with those not receiving those services.[2] Indicators for each of the blocks used in the random assignment process (defined by the school and the grade level of the student) were

[2] Intent-to-treat impact estimates are used to characterize the program's impact. The impact estimate is allowed to vary randomly across schools – this is considered a "random-effect" estimate of the program's impact that can be generalized to the broader population of schools and students served by Reading Partners.

included in the model to account for the study design and for differential rates of treatment assignment, by block. These block indicators also controlled for variation in mean outcome levels across blocks (which can be due to different characteristics of school settings or other factors). The model controls for the individual-level pretest measure and the time lapse between baseline testing and follow-up testing. Other baseline covariates are added to the model to improve precision. These covariates include student's gender, race/ethnicity, free/reduced-price lunch status, age, English language learner (ELL) status, special education status, and whether the student is overage for grade.

For the cost study component of the evaluation, all of the resources (or "ingredients") used to implement Reading Partners were identified from discussions with program administrators, interviews and surveys of program and school staff, and data collected during the observational school site visits. After a comprehensive list of ingredients was compiled, each ingredient was matched with its national price or cost in 2012 dollars. National prices were obtained from publicly available sources such as the Bureau of Labor Statistics, the Department of Labor, the National Center for Education Statistics, and AmeriCorps and then a cost per student was calculated. The costs of the other supplemental reading services provided at each school were calculated in the same detailed way and followed the same pricing approach using national prices. The Reading Partners costs are compared to the cost of the other supplemental reading services provided in the school.

Findings / Results:

In general, the Reading Partners program was implemented with a high degree of fidelity to the program model. The biggest challenge that Reading Partners faced in implementing the program was tutor attendance and retention. However, there were structures in place to address these challenges and as a result most students were tutored twice a week on a regular basis throughout the year, even if that tutoring was not consistently delivered by the same person.

As shown in Table 3, Reading Partners had a positive and statistically significant impact on all three measures of student reading proficiency, with effect size impacts of 0.10 on reading comprehension scores, 0.09 on reading fluency and 0.11 on sight word reading. The impacts were equivalent to what a student learns in approximately 1.5 to 2 months of additional school. Impacts did not differ significantly for students from different grade levels or baseline achievement levels, for male or female students, for those for whom English was a second language or for those who had previously been served by the Reading Partners program.

Analysis of the cost study findings have not been finalized but will be available by the time of the conference.

Conclusions:

The findings all suggest that the Reading Partners model "works"--i.e., it produces measurable impacts in reading skills among participants and that the program was effective in helping students with a fairly broad range of reading abilities and from a variety of different grade levels and backgrounds. The cost data will shed light on how much the program costs relative to the reading impacts it generates.

Appendices

Appendix A. References

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

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

U.S. Department of Education, National Center for Education Statistics. (2013). *National assessment of educational progress reading: Grade 4 national results*. Retrieved from http://www.nationsreportcard.gov/reading_math_2013/#/.

Appendix B. Tables and Figures

Table 1

Characteristics of Reading Partners Study Schools and Other School Samples (2011-2012)

		Other Reading Partners Schools[a]			U.S. Title I Schools[b]			Average U.S. Schools[c]		
Characteristic	Study Schools	Mean	Difference	P-Value	Mean	Difference	P-Value	Mean	Difference	P-Value
Eligible for Title I program (%)	88.9	96.2	-7.3	0.209	100.0	-11.1	<0.001	75.4	13.5	0.184
Students eligible for free/reduced-price lunch[d] (%)	81.7	84.5	-2.8	0.411	69.1 ***	12.6	0.008	53.5 ***	28.2	<0.001
Race/ethnicity (%)										
Black	20.7	28.4	-7.7	0.304	20.7	0.0	1.000	15.9	4.8	0.404
Hispanic	61.7	53.0	8.8	0.204	29.8 ***	31.9	<0.001	23.3 ***	38.4	<0.001
Asian	9.6	8.3	1.3	0.639	3.4 ***	6.2	0.002	4.7 **	4.9	0.025
White	5.7	7.3	-1.6	0.496	41.3 ***	-35.6	<0.001	51.6 ***	-45.9	<0.001
Other	2.2	3.0	-0.8	0.336	4.7	-2.5	0.263	4.5	-2.2	0.234
Male (%)	51.9	51.8	0.1	0.970	53.5	-1.6	0.462	53.00	-1.1	0.564
Average number of students	494	451	43.1	0.318	468	26.2	0.622	472	22.2	0.677
Grade 2	79	69	10.2	0.169	68	10.4	0.217	70	8.4	0.331
Grade 3	76	65	10.6	0.142	68	7.5	0.384	70	5.3	0.554
Grade 4	74	65	9.2	0.227	67	7.2	0.396	70	4.7	0.591
Grade 5	76	64	12.0	0.131	67	9.1	0.290	69	6.2	0.488
School setting[e] (%)										
Urban	73.7	91.7 **	-18.0	0.027	39.8 ***	33.9	0.003	31.8 ***	41.9	<0.001
Suburban	26.3	8.3 **	18.0	0.027	21.3	5.0	0.592	29.5	-3.2	0.761
Town	0.0	0.0	0.0	NA	10.1	-10.1	0.145	8.8	-8.8	0.176
Rural area	0.0	0.0	0.0	NA	28.9 ***	-28.9	0.005	29.9 ***	-29.9	0.004
Sample size	19	84			24,914			42,202		

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 percent, ** = 5 percent, * = 10 percent.

[a]"Other Reading Partners Schools" include all other non-study Reading Partners schools that met the "Average U.S. Schools" criteria below.

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

[c]"Average U.S. Schools" include non-Reading Partners schools that offer grade 2 through grade 5, defined as "regular" schools by the Common Core of Data, and 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 a territory inside an urbanized area and inside a principal city having a population greater than 100,000. "Suburb" is defined as a territory outside of a principal city and inside an urbanized area with a population of less than 250,000.

Table 2

Baseline Characteristics of Reading Partners and Control 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 by grade (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		

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 percent; ** = 5 percent; * = 10 percent.

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: eight or older for second-grade, nine 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.

Table 3

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				

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 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 percent; ** = 5 percent; * = 10 percent.

Rounding may cause slight discrepancies in sums and differences.