



WWC Single Study Review

A review of the design and summary of findings for an individual study



June 2016

WWC Review of the Report “Mobilizing Volunteer Tutors to Improve Student Literacy: Implementation, Impacts, and Costs of the Reading Partners Program”¹

The findings from this review do not reflect the full body of research evidence on *Reading Partners*.

What is this study about?

The study authors examined whether participation in *Reading Partners*, a supplementary program that provides elementary school students with individualized tutoring from community volunteers, led to improvements in reading achievement, mathematics achievement, and student attendance. The study included over 1,200 students in 19 elementary schools in California, New York, and Washington, DC.

The authors randomly assigned eligible students either to be invited to participate in the *Reading Partners* program during the 2012–13 school year or to be in a comparison group that did not participate. Students in both groups received a regular reading curriculum and could participate in other supplementary reading programs offered by the schools. A total of 1,166 students who completed at least one follow-up assessment were included in the analysis.

To measure program impacts, the authors compared the reading achievement, mathematics achievement, and student attendance of the intervention and comparison groups in spring 2013. The authors identified the primary outcomes as results from standardized assessments in reading comprehension, reading fluency, and sight word efficiency. Secondary outcomes included state test scores in reading and mathematics provided by school districts, as well as student attendance rates.² The authors also examined impacts on the three standardized reading assessments for subgroups

defined by grade level, gender, English learner status, baseline reading achievement, and prior receipt of *Reading Partners* services.³

WWC Rating

The research described in this report meets WWC group design standards without reservations

This study is a randomized controlled trial with low attrition.

Features of *Reading Partners*

Reading Partners is a supplementary program that aims to serve students who are 0.5 to 2.5 years behind grade level in reading. The program has six core components: (1) one-on-one tutoring by a volunteer twice a week for 45 minutes; (2) dedicated school space and use of resource materials; (3) a structured curriculum that tutors use with students; (4) regularly timed assessments that inform students’ Individualized Reading Plans; (5) ongoing training for tutors; and (6) instructional supervision and support for tutors. Each program site is managed by a full-time coordinator who oversees the instruction and training of volunteers.

What did the study find?

The study authors reported, and the WWC confirmed, that the *Reading Partners* program led to statistically significant improvements in reading comprehension,

reading fluency, and sight word efficiency (i.e., pronouncing printed words) for the full sample. The study authors also reported, and the WWC confirmed, that there was a statistically significant improvement in sight word efficiency for students in grades 2 and 3. The study reported that the following subgroup findings were statistically significant: improvements in reading comprehension, reading fluency, and sight word efficiency among English language learners who were in the bottom three quartiles of reading achievement at baseline; improvements in reading fluency for female students; and improvements in

sight word efficiency for male students, students in the bottom quartile of reading achievement at baseline, and students who had not received *Reading Partners* services prior to the study. The WWC did not confirm that these subgroup findings were statistically significant, after the WWC adjusted for multiple comparisons.⁴

The study found no statistically significant differences in reading or mathematics achievement scores on state standardized tests or on attendance.

Appendix A: Study details

Jacob, R. T., Armstrong, C., & Willard, J. A. (2015). *Mobilizing volunteer tutors to improve student literacy: Implementation, impacts, and costs of the Reading Partners program*. New York: MDRC. <http://files.eric.ed.gov/fulltext/ED558508.pdf>.

- Setting** The study took place in 19 schools: 16 schools in California, two in New York, and one in Washington, DC. The majority of the schools were Title I schools. Eight of the 19 schools were “in varying stages of federal School Improvement status.”
- Study sample** Study authors randomly assigned students in grades 2–5 either to receive the supplemental *Reading Partners* one-to-one tutoring intervention or to serve in a business-as-usual comparison group. Students in both conditions were eligible to receive other supplemental reading services. Students were randomized to *Reading Partners* within school and grade level blocks (grades 2 and 3 were combined, as were grades 4 and 5) to ensure that equal numbers of upper and lower elementary school students were represented in the sample. There were 1,265 students from 19 schools that were randomly assigned to the two study conditions. Of this baseline sample, 1,166 students completed at least one assessment at follow-up. The analytic sample was 55% male, 65% Hispanic, 19% African American, 9% Asian, 6% White, and 1% other race or ethnicity. Over 90% of students in the analytic sample were eligible for free or reduced-price lunch, and 55% of students were English learners. The intervention and comparison groups had similar characteristics.
- Intervention group** The *Reading Partners* intervention was a pull-out program that was offered during school hours or afterschool programs. Volunteer tutors and students met in a dedicated space, and the program was designed to provide two 45-minute sessions per week. The average number of *Reading Partners* tutors assigned to each student was 2.5. Across schools, the average number of tutors assigned to each student ranged from 1.7 to 3.6 tutors. On average, students received about 1.5 tutoring sessions per week for 28 weeks. In eight of the 19 schools, the *Reading Partners* center was in its second year of operation, while the rest had been operating a *Reading Partners* center for at least 3 years.
- Comparison group** Students assigned to the comparison group did not receive one-to-one tutoring through the *Reading Partners* program during the study period, but they were eligible for other supplemental reading services at school. The supplemental programs were usually offered in small group settings. The researchers found that 65% of comparison group students received supplemental reading services at school, and 21% of comparison group students received one-to-one tutoring. On average, the comparison students received 57 fewer minutes of supplemental reading instruction time per week than the intervention group. The researchers also found that 26% of comparison group students had received *Reading Partners* services prior to the study period.

Outcomes and measurement

Scores from three reading assessments were identified as the primary outcomes by the study: (1) the Stanford Achievement Test, 10th Edition (SAT-10); (2) the AIMSweb 1-minute Oral Reading Fluency measure; and (3) the Test of Word Reading Efficiency, 2nd Edition (TOWRE-2). The SAT-10, AIMSweb, and TOWRE-2 measure reading comprehension, reading fluency, and alphabets, respectively. All three assessments are standardized tests. Assessments were administered by independent staff hired by the study team who participated in a 2-day training program. Baseline data collection took place as soon as possible after random assignment, and follow-up data collection was conducted in the spring of 2013, as close to the end of the school year as possible. The authors also identified two secondary outcomes that are eligible for review: school attendance and state standardized test results obtained from school districts. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

In addition to volunteer tutors, other core components of the *Reading Partners* program that supported the implementation of the individualized tutoring included the following: dedicated space and materials (including work stations, a library, and resource materials); a structured curriculum that tutors use with students; availability and use of assessment tools to be administered to students during the academic year; 2 weeks of training for new program managers; ongoing training provided to volunteers at the beginning of the year and on a monthly basis thereafter; and instructional supervision and support provided by staff to volunteer tutors.

Reason for review

This study was identified for review by receiving media attention.

Appendix B: Outcome measures for each domain

Alphabetics	
<i>Test of Word Reading Efficiency, 2nd Edition (TOWRE-2)</i>	The TOWRE-2 is a standardized test that measures sight word efficiency, or the ability to pronounce printed words. Scores are based on the number of words that a student can identify accurately within 45 seconds. This assessment takes about 5 minutes to complete.
Attendance	
<i>School attendance rate</i>	Attendance rates were measured using data from individual school districts.
General reading achievement	
<i>State reading achievement test</i>	Students' scaled scores on their state's reading achievement test were converted to z-scores. The z-scores were calculated by subtracting the mean score of the comparison group from the intervention group mean and dividing by the standard deviation of the comparison group in the same state and grade level.
<i>Achieved proficiency according to the state reading achievement test</i>	This is a binary measure equal to "1" if a student met the proficiency criteria for the state reading achievement test established in his or her state. The measure is equal to "0" if a student did not meet the proficiency criteria established by the state.
Mathematics achievement	
<i>State mathematics achievement test</i>	Students' scaled scores on their state's mathematics achievement test were converted to z-scores. The z-scores were calculated by subtracting the mean score of the comparison group from the intervention group and dividing by the standard deviation of the comparison group in the same state and grade level.
<i>Achieved proficiency according to the state mathematics achievement test</i>	This is a binary measure equal to "1" if a student met the proficiency criteria for the state mathematics achievement test established in his or her state. The measure is equal to "0" if a student did not meet the proficiency criteria established by the state.
Reading comprehension	
<i>Stanford Achievement Test, 10th Edition (SAT-10)</i>	The SAT-10 is a group-administered test in which students read short passages and answer questions about the passages.
<i>Moved out of lowest national quartile on SAT-10 reading test</i>	This is a binary measure that is defined for students whose baseline scores on the SAT-10 were among the lowest 25% of scores nationally (lowest quartile). About 60% of the study sample scored in the lowest 25% of scores nationally on the SAT-10. This measure is equal to "1" if the student's score at the end of the study period was higher than 25th percentile nationally, and "0" if the student's score at the end of the study period was not higher than 25th percentile.
<i>Moved up a national quartile on SAT-10 reading test</i>	This is a binary measure that is defined for students whose baseline scores on the SAT-10 were among the lowest 75% of scores nationally (bottom three quartiles). This measure is equal to "1" if the student's score at the end of the study period had improved sufficiently enough that it was in a higher quartile than the baseline score. The measure is equal to "0" if the student's scores at the end of the study period were in the same or lower quartile of scores nationally than at baseline.
Reading fluency	
<i>AIMSweb</i>	The AIMSweb is an individually administered oral reading fluency assessment in which students read a passage aloud for 1 minute. The score is the number of words that are read correctly.

Table Notes: Teacher reports of students' academic behavior (satisfactory homework completion, satisfactory assignment completion, satisfactory level of attentiveness, unsatisfactory level of disruptiveness, and composite behavior measure) and teacher ratings of academic performance (satisfactory performance in reading, satisfactory performance in mathematics, and satisfactory performance overall) were assessed by the study but are not included in this review. These outcomes are ineligible for review because the criteria used to measure the outcomes may have varied by teacher.

Appendix C: Study findings for each domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Alphabetics								
<i>Test of Word Reading Efficiency, 2nd Edition</i>	All students	19 schools/ 1,147 students	92.78 (nr)	91.37 (nr)	1.42	0.11	+4	< .01
Domain average for alphabetics						0.11	+4	Statistically significant
Reading comprehension								
<i>Stanford Achievement Test, 10th Edition (SAT-10)</i>	All students	19 schools/ 1,146 students	592.42 (nr)	588.94 (nr)	3.48	0.10	+4	.04
Domain average for reading comprehension						0.10	+4	Statistically significant
Reading fluency								
<i>AIMSweb</i>	All students	19 schools/ 1,151 students	0.06	-0.03	0.09	0.09	+4	.03
Domain average for reading fluency						0.09	+4	Statistically significant

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. nr = not reported.

Study Notes: The study characterized these three outcomes as the primary outcomes. The WWC did not need to make corrections for clustering, multiple comparisons, or to adjust for baseline differences. The mean differences, effect sizes, and p-values presented here were reported in the original study. The authors confirmed the sample sizes for these analyses. This study is characterized as having a statistically significant positive effect on alphabetics, reading comprehension, and reading fluency because the estimated effect in each domain is positive and statistically significant after any necessary adjustments. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0), p. 26.

Appendix D: Supplemental findings by domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Alphabetics								
<i>Test of Word Reading Efficiency, 2nd Edition (TOWRE-2)</i>	Students in grades 2 and 3	594 students	nr (nr)	nr (nr)	nr	0.17	+7	< .01
<i>TOWRE-2</i>	Students in grades 4 and 5	557 students	nr (nr)	nr (nr)	nr	0.05	+2	.48
<i>TOWRE-2</i>	Male students	627 students	nr (nr)	nr (nr)	nr	0.13	+5	.01
<i>TOWRE-2</i>	Female students	514 students	nr (nr)	nr (nr)	nr	0.09	+4	.14
<i>TOWRE-2</i>	Students who are English learners (ELs)	625 students	nr (nr)	nr (nr)	nr	0.11	+4	.04
<i>TOWRE-2</i>	Students who are not ELs	498 students	nr (nr)	nr (nr)	nr	0.11	+4	.06
<i>TOWRE-2</i>	Students in three lowest quartiles	892 students	nr (nr)	nr (nr)	nr	0.10	+4	.02
<i>TOWRE-2</i>	Students in lowest quartile	306 students	nr (nr)	nr (nr)	nr	0.22	+9	.03
<i>TOWRE-2</i>	Students in second quartile	299 students	nr (nr)	nr (nr)	nr	0.08	+3	.30
<i>TOWRE-2</i>	Students in third quartile	287 students	nr (nr)	nr (nr)	nr	0.11	+4	.28
<i>TOWRE-2</i>	Students in highest quartile	259 students	nr (nr)	nr (nr)	nr	0.14	+6	.11
<i>TOWRE-2</i>	Students with prior receipt of <i>Reading Partners</i> services	311 students	nr (nr)	nr (nr)	nr	0.05	+2	.63
<i>TOWRE-2</i>	Students without prior receipt of <i>Reading Partners</i> services	802 students	nr (nr)	nr (nr)	nr	0.12	+5	.01
Attendance								
<i>School attendance rate</i>	All students	19 schools/ 1,214 students	0.96 (nr)	0.96 (nr)	0.00	-0.04	-2	.47
General reading achievement								
<i>State reading achievement test</i>	All students	19 schools/ 1,117 students	0.08 (nr)	0.02 (nr)	0.06	0.06	+2	.23
<i>Achieved proficiency according to the state reading achievement test</i>	All students	19 schools/ 1,117 students	0.15 (nr)	0.15 (nr)	0.01	0.02	+1	.67

Mathematics achievement								
<i>State mathematics achievement test</i>	All students	19 schools/ 1,117 students	0.01 (nr)	0.00 (nr)	0.01	0.01	0	.91
<i>Achieved proficiency according to the state mathematics achievement test</i>	All students	19 schools/ 1,117 students	0.29 (nr)	0.32 (nr)	-0.03	-0.06	-2	.35
Reading comprehension								
<i>Moved out of lowest national quartile on SAT-10 reading test</i>	All students	19 schools/ 1,146 students	0.19 (nr)	0.12 (nr)	0.08	0.22	+9	.01
<i>Moved up a national quartile on SAT-10 reading test</i>	All students	19 schools/ 1,146 students	0.14 (nr)	0.11 (nr)	0.03	0.11	+4	.18
<i>SAT-10</i>	Students in grades 2 and 3	602 students	nr (nr)	nr (nr)	nr	0.09	+4	.13
<i>SAT-10</i>	Students in grades 2 and 3	544 students	nr (nr)	nr (nr)	nr	0.13	+5	.10
<i>SAT-10</i>	Male students	619 students	nr (nr)	nr (nr)	nr	0.08	+3	.18
<i>SAT-10</i>	Female students	516 students	nr (nr)	nr (nr)	nr	0.11	+4	.07
<i>SAT-10</i>	Students who are ELs	625 students	nr (nr)	nr (nr)	nr	0.11	+4	.05
<i>SAT-10</i>	Students who are not ELs	492 students	nr (nr)	nr (nr)	nr	0.01	0	.85
<i>SAT-10</i>	Students in three lowest quartiles	894 students	nr (nr)	nr (nr)	nr	0.11	+4	.04
<i>SAT-10</i>	Students in lowest quartile	304 students	nr (nr)	nr (nr)	nr	0.02	+1	.84
<i>SAT-10</i>	Students in second quartile	302 students	nr (nr)	nr (nr)	nr	0.16	+6	.06
<i>SAT-10</i>	Students in third quartile	288 students	nr (nr)	nr (nr)	nr	0.12	+5	.26
<i>SAT-10</i>	Students in highest quartile	252 students	nr (nr)	nr (nr)	nr	0.00	0	.98
<i>SAT-10</i>	Students with prior receipt of <i>Reading Partners</i> services	310 students	nr (nr)	nr (nr)	nr	0.11	+4	.18
<i>SAT-10</i>	Students without prior receipt of <i>Reading Partners</i> services	798 students	nr (nr)	nr (nr)	nr	0.07	+3	.14
Reading fluency								
<i>AIMSweb</i>	Students in grades 2 and 3	594 students	nr (nr)	nr (nr)	nr	0.09	+4	.18
<i>AIMSweb</i>	Students in grades 4 and 5	557 students	nr (nr)	nr (nr)	nr	0.10	+4	.10

<i>AIMSweb</i>	Male students	627 students	nr (nr)	nr (nr)	nr	0.07	+3	.21
<i>AIMSweb</i>	Female students	514 students	nr (nr)	nr (nr)	nr	0.14	+6	.02
<i>AIMSweb</i>	Students who are ELs	625 students	nr (nr)	nr (nr)	nr	0.14	+6	.02
<i>AIMSweb</i>	Students who are not ELs	498 students	nr (nr)	nr (nr)	nr	0.04	+2	.58
<i>AIMSweb</i>	Students in three lowest quartiles	892 students	nr (nr)	nr (nr)	nr	0.09	+4	.04
<i>AIMSweb</i>	Students in lowest quartile	306 students	nr (nr)	nr (nr)	nr	0.19	+8	.08
<i>AIMSweb</i>	Students in second quartile	299 students	nr (nr)	nr (nr)	nr	0.05	+2	.54
<i>AIMSweb</i>	Students in third quartile	287 students	nr (nr)	nr (nr)	nr	0.08	+3	.35
<i>AIMSweb</i>	Students in highest quartile	259 students	nr (nr)	nr (nr)	nr	0.10	+4	.30
<i>AIMSweb</i>	Students with prior receipt of <i>Reading Partners</i> services	311 students	nr (nr)	nr (nr)	nr	0.03	+1	.72
<i>AIMSweb</i>	Students without prior receipt of <i>Reading Partners</i> services	802 students	nr (nr)	nr (nr)	nr	0.11	+4	.06

Table Notes: The supplemental findings presented in this table are additional findings that meet WWC design standards without reservations, but do not factor into the determination of the study rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. nr = not reported.

Study Notes: This table contains subgroup results for the three primary outcomes, and also contains the results for all non-primary outcomes. The effect sizes and *p*-values presented here were reported in the original study. A correction for multiple comparisons was needed within each of the domains in this table except for the Attendance domain. In the Alphabets domain, the correction resulted in a WWC-computed critical *p*-value of .0045 for the TOWRE-2 outcome for students in grades 2 and 3, and this was the only case in which the *p*-value of the impact estimate (.002) was smaller than the WWC-computed critical *p*-value. Therefore, the WWC finds the result for the TOWRE-2 outcome for students in grades 2 and 3 to be statistically significant but does not find the results for the TOWRE-2 outcome for other subgroups to be statistically significant. In the Reading Comprehension and Reading Fluency domains, no outcomes had *p*-values that were smaller than the WWC-computed critical *p*-values. Therefore, the WWC does not find any of the impacts on reading comprehension and reading fluency in this table to be statistically significant. Information on sample sizes of state test score outcomes and subgroup analyses of primary outcomes was provided by the authors. The quartile-based subgroups (e.g., students in second highest quartile) are relative to national quartiles of SAT scores from students' scores prior to the intervention.

Endnotes

¹ Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the authors) to assess whether the study design meets WWC group design standards. The review reports the WWC's assessment of whether the study meets WWC group design standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the single study review protocol (version 2.0). Two related protocols, the Primary Math review protocol (version 3.1) and the Beginning Reading review protocol (version 3.1), were consulted to guide the determination of eligible outcomes. A quick review of this study was released in August 2015, and this report is the follow-up review that replaces that initial assessment. The rating of the study has not changed. The WWC rating applies only to the study outcomes that were eligible for review under this topic area. The WWC rating and analyses reported in this single study review are only for those eligible outcomes that either met WWC group design standards without reservations or met WWC group design standards with reservations, and do not necessarily apply to all results presented in the study.

² As reflected in the title of the report, the study also examined the cost of *Reading Partners* and other supplemental services available in the study sample of the schools. However, cost is not an eligible outcome under the single study review protocol, so these results were not included in this WWC report.

³ There were eight teacher-reported outcomes included in the study that are not described in this WWC report. See the table notes in Appendix B for more information.

⁴ This single study review follows the study's characterization of the reading comprehension, reading fluency, and sight word efficiency outcomes as being the primary outcomes. The study findings for these outcomes with the full sample are presented in Appendix C. Subgroup findings for these outcomes and for all other outcomes are presented in Appendix D. The WWC conducted a multiple comparisons adjustment within domains among the supplementary outcomes in Appendix D and determined that some findings were not statistically significant after the adjustment. See the table notes in Appendix D for more information.

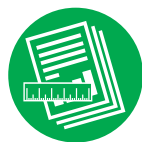
Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2016, June). *WWC review of the report: Mobilizing volunteer tutors to improve student literacy: Implementation, impacts, and costs of the Reading Partners program*. Retrieved from <http://whatworks.ed.gov>

Glossary of Terms

Attrition	Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.
Clustering adjustment	If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.
Confounding factor	A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.
Design	The design of a study is the method by which intervention and comparison groups were assigned.
Domain	A domain is a group of closely related outcomes.
Effect size	The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.
Eligibility	A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.
Equivalence	A demonstration that the analytic sample groups are similar on observed characteristics defined in the review area protocol.
Improvement index	Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual starts at the 50th percentile, the measure ranges from -50 to +50.
Multiple comparison adjustment	When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.
Quasi-experimental design (QED)	A quasi-experimental design (QED) is a research design in which study participants are assigned to intervention and comparison groups through a process that is not random.
Randomized controlled trial (RCT)	A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.
Single-case design (SCD)	A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.
Standard deviation	The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample are spread out over a large range of values.
Statistical significance	Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < .05$).
Substantively important	A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the [WWC Procedures and Standards Handbook \(version 3.0\)](#) for additional details.



Intervention
Report



Practice
Guide



Quick
Review



Single Study
Review

A **single study review** of an individual study includes the WWC's assessment of the quality of the research design and technical details about the study's design and findings.

This single study review was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.