WRC Impact Evaluation

September 6, 2023

Table of Contents

Executive Summary

Washington Reading Corps (WRC) supports literacy development for preschool children and K-12 students across Washington State. WRC designed Reading Power as a program that builds literacy and reading skills for students in kindergarten through 4th grade by developing strong individual relationships between WRC members and students and providing additional reading practice time focused on specific literacy and reading skills. In Reading Power, WRC members provide tutoring for a minimum of three 20-minute sessions weekly for at least 12 weeks. Tutors receive training on the continuum of literacy and reading development and on relevant age- and grade-level literacy and reading skills.

Mission360 Consulting, LLC conducted a randomized controlled trial study to examine whether Kindergarten and 1st-grade students who received skills-development through Reading Power tutoring showed more growth in letter-name and letter-sound identification than did students who received other LAP-ELA interventions but not Reading Power. Information from the evaluation will help guide decision-making about continuous improvement as well as guide future evaluation.

The population of interest for this study consisted of kindergarten and 1st-grade students enrolled in public elementary schools in Washington state who were identified by their school as in need of LAP-ELA intervention to build their literacy/reading skills. The sample consisted of 292 students enrolled in one of the 14 schools that agreed to participate in the study.

The evaluation question this study was designed to answer is: *Do kindergarten and 1st-grade students who receive Reading Power tutoring demonstrate a greater increase in letter-name identification and letter-sound identification from pre-intervention assessment to post-intervention assessment than similar students who receive other LAP-ELA interventions (but not Reading Power)?* The question was answered through multiple linear regression (using the lm function in the base package for the statistical programming language R).

The 12-week study period began 2023-01-30 and ended 2023-05-25. Baseline (pre-intervention) assessments were administered to students in the treatment and control groups at the beginning of the study (between 2023-01-30 and 2023-02-24). Post-intervention assessments were administered to students in the treatment and control groups at the end of the study (between 2023-05-02 and 2023-05-25).

The study found that, for students in both the treatment and control groups, the average number of letter-names and letter-sounds identified correctly increased from the beginning to end of the study period. The study did not find any statistically-significant differences between the treatment and control groups in the size of that increase: both groups had similar average baseline and post-intervention scores on the outcomes of interest.

The Impact of WRC Reading Power on Students' Letter-Name and Letter-Sound Identification

Washington Reading Corps (WRC) supports literacy development for preschool children and K-12 students across Washington State. The first program of its kind in the nation, WRC was created in 1998 through an unprecedented collaboration between four public agencies: the Washington Governor's office; the Washington State Commission for National and Community Service; the Washington Employment Security Department (which houses the Washington Service Corps); and the Office of the Superintendent of Public Instruction (OSPI), the state's education agency. Now jointly led by Washington Service Corps and the OSPI, Washington Reading Corps recruits, trains, places, and supports AmeriCorps members who provide reading tutoring for kindergarten-to-4th-graders.

Study Background

Since its inception, WRC has placed members at partner sites to provide tutoring and other literacy and reading services for children and students who need additional support to become fluent readers. WRC is one of the kindergarten to 4th-grade English Language Arts (ELA) best practices approved by OSPI's Learning Assistance Program (LAP). LAP-ELA interventions target students who do not meet grade-level proficiency standards in reading. While Reading Power tutoring is WRC's core intervention, WRC members also provide other supports as needed at the sites where they are placed.

Reading Power was designed to build literacy and reading skills for students in kindergarten through 4th grade by developing strong individual relationships between WRC members and students and providing additional reading practice time focused on specific literacy and reading skills.

For kindergarten and 1st-grade students the focus of skills-building is on:

- Alphabet knowledge (specifically letter-name and letter-sound identification)
- Print awareness, or the understanding that print carries meaning and that books contain letters and words.
- Phonological awareness or the ability to hear, identify, and manipulate sounds in spoken language.
- Beginning word recognition - the emergent ability to see a word and recognize its pronunciation.

For 2nd- through 4th-grade students, the focus of skills-building is on:

- Phonological awareness – the ability to hear, identify, and manipulate sounds in spoken language.
- Word recognition – the ability to see a word and recognize its pronunciation immediately, without any conscious effort.
- Reading fluency – the ability to read with speed, accuracy, and proper expression.
- Reading comprehension – the ability to understand and interpretate of what is read.

In Reading Power, WRC members provide tutoring for a minimum of three 20-minute sessions weekly for at least 12 weeks. Tutors receive training on the continuum of literacy and reading development and on relevant age- and grade-level literacy and reading skills.

WRC has a strong commitment to evaluating the impact of its services and using evaluation data to guide program design and implementation:

1) An external evaluation conducted in 2015 tested the impact that specific WRC program elements had on student reading outcomes, examining data from a sample of 25 schools and 1,972 children who had received WRC tutoring. Results suggested that, for these children, the only significant predictor of meeting grade-level reading benchmarks was Site Supervisor helpfulness, where higher supervisor helpfulness was associated with a higher likelihood of meeting reading benchmarks. A notable limitation of this study was the lack of a control group.

2) An external evaluation conducted in 2018 was intended to address some of the limitations of previous evaluation. This study measured the difference in growth in letter-name identification and letter-sound identification between two groups of students: a treatment group consisting of kindergarten and 1st-grade students who received WRC's tutoring intervention and a control group of similar students who received other LAP-ELA interventions, but not WRC. This study had multiple limitations, including the lack of a common assessment to measure reading skills growth, inconsistencies in data collection and reporting, and variability in tutoring content and delivery. While the study found that staff at schools where WRC Members were placed perceived a strong positive impact for WRC on the culture of reading at their schools, the study did not detect any statistically-significant differences in reading-skills growth between the treatment and control groups.

   Based on recommendations from the researchers, WRC developed the Reading Power program to address these limitations and began implementing Reading Power as its core tutoring model in the 2020 school year.

Study Purpose and Design

The current study continued to build evidence of effectiveness for WRC tutoring. This randomized controlled trial study was designed to examine whether Kindergarten and 1st-grade students who received skills-development through Reading Power tutoring showed more growth in letter-name and letter-sound identification than did students who received other LAP-ELA interventions but not Reading Power. Information from the evaluation will help guide decision-making about continuous improvement as well as guide future evaluation.

Methodology

Evaluation questions

The evaluation question this study was designed to answer is: *Do kindergarten and 1st-grade students who receive Reading Power tutoring demonstrate a greater increase in letter-name identification and letter-sound identification from pre-intervention assessment to post-intervention assessment than similar students who receive other LAP-ELA interventions (but not Reading Power)?*

Population and Sample

The population of interest for this study consisted of kindergarten and 1st-grade students enrolled in public elementary schools in Washington state who were identified by their school as in need of LAP-ELA intervention to build their literacy/reading skills. The sample consisted of 292 students enrolled in one of the 14 schools that agreed to participate in the study.[1]

Table 1, below, describes the characteristics of the schools that were recruited and agreed to participate (the percentage of enrolled students who were English Learners, eligible for the federal free or reduced-price lunch program, had high growth on standardized tests for English Language Arts (ELA), met or exceeded standards for ELA proficiency, and regularly attend school).

Table 1: School Characteristics

School Characteristics	Mean (N = 14)
English Learners	16%
Free or Reduced Priced Lunch	73%
High ELA Growth	31%
Met ELA Standards	41%
Students Regularly Attend	76%

[1] Fifteen schools originally agreed to participate, but one school dropped out before the study began.

Eligibility for participation in this study was based on standardized testing conducted by schools in fall 2022. During this testing, students must have been identified as being in need of LAP-ELA intervention (because they were behind one or more grade levels in reading), but not in need of an Individualized Educational Plan (IEP) or eligible for special education services. Of the students tested, 292 met the inclusion criteria.

Each of the 292 students who met the criteria were then randomly assigned to either the treatment group (Reading Power tutoring) or the control group (other LAP-ELA intervention, but not Reading Power), or were not included in the study.[2] This random assignment was conducted using the open source statistical programming language R.[3] The final analytic sample consisted of 151 students (80 in kindergarten and 71 in 1st grade).

A power analysis conducted by the evaluation team (using GPower 3.1[4]) indicated that, to detect a difference of at least 0.25 (Cohen's d, with a power of 0.80) on post-intervention assessment scores, and assuming the use of an independent samples t-test, the estimated necessary sample size was approximately 398 students.

Study Period and Baseline Characteristics

The 12-week study period began 2023-01-30 and ended 2023-05-25. Baseline (pre-intervention) assessments were administered to students in the treatment and control groups at the beginning of the study (between 2023-01-30 and 2023-02-24). Post-intervention assessments were administered to students in the treatment and control groups at the end of the study (between 2023-05-02 and 2023-05-25).

Table 2, below, shows the scores on baseline assessments of letter-name identification and letter sound identification for the sample of students in this study. This table shows that:

- the average baseline score for letter-name identification was 22.4 for the treatment group and 23.2 for the control group. This difference of -0.78 was not significant.
- the average baseline score for letter-sound identification was 21.9 for the treatment group and 21.7 for the control group. This difference of 0.24 was not significant.

[2] Students were randomly assigned to the treatment or control group based on a maximum capacity of each member to be able to provide Reading Power to no more than 5 students per grade.

[3] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

[4] Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. Behavior research methods, 41(4), 1149-1160.

Table 2: Baseline Student Characteristics

Student Characteristics	Control Mean (SD), N	Reading Power Mean (SD), N
Baseline Score, Letter-Name Identification	23.2 (4.08), 72	22.4 (5.62), 79
Baseline Score, Letter-Sound Identification	21.7 (7.31), 72	21.9 (7.25), 79

The similarities between the treatment and control groups on baseline scores suggest that the random assignment of students into the treatment and control groups was successful.

Treatment and Control Conditions

For students in the treatment group:

- WRC members delivered Reading Power during tutoring sessions three times a week for at least 20 minutes per session. Members administered the baseline (pre-intervention) assessment during the first tutoring session and the formative (post-intervention) assessment during the last tutoring session.
- During each tutoring session, WRC members provided regular opportunities for students to practice fundamental phonemic awareness, phonics, word identification and/or oral reading fluency skills during scripted, one-to-one pull-out sessions.

For students in the control group:

- WRC administered the pre-intervention and post-intervention assessments during the same weeks they administered those assessments to the treatment group.
- Other adults at the program sites provided other LAP services, but not Reading Power tutoring, during the 12-week study. These students were eligible to receive Reading Power tutoring as soon as the study concluded.

Independent and Dependent Variables

This study included two dependent variables:

1) The score on the assessment of letter-name identification. These scores ranged from 2 to 26 and reflected the number of letters a student was able to identify correctly.
2) The score on the assessment of letter-sound identification. These scores ranged from 2 to 30 and reflected a student's ability to efficiently reproduce the sound commonly associated with a given grapheme (i.e., the discrete sounds that correspond to specific letters in the English-language alphabet).

Copies of the kindergarten and 1st-grade letter-name identification and letter-sound identification assessments are included in Appendix A.

Data on both dependent variables were collected before treatment began (i.e., baseline) and after treatment ended (i.e., post-treatment).

The study included the following independent variables:

1) Assignment: A dichotomous variable indicating the group to which each student was assigned (1 = treatment, 0 = control).
2) Treatment: A dichotomous variable indicating whether or not the student received the treatment (1 = treatment, 0 = no treatment). This information was documented to determine the extent to which treatment assignment was adhered to and to allow for additional sensitivity analyses.
3) SchoolID: A numeric code indicating the school in which the student was enrolled during the timeframe of this study. In addition to this numeric code, the evaluation team recorded the names of each of the schools in the study.

A full list of the variables captured as part of this study is included in Appendix B.

Analysis

The evaluation question – *"Do kindergarten and 1st-grader students who receive Reading Power tutoring demonstrate a greater increase in letter-name and letter-sound identification from pre-intervention assessment to post-intervention assessment than similar students who receive other LAP services (but not Reading Power)?"* – was answered through multiple linear regression[5] (using the lm function in the base package for the statistical programming language R). Baseline differences on the letter names and letter sounds assessments were also assessed using a linear regression. The evaluation team used multiple linear regression to control for baseline scores while estimating the effect of the treatment (Reading Power) on post-intervention scores.

The regression model took the general form:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + e_i$$

[5] Although the results presented in this paper are based on multiple linear regression, the authors used a variety of statistical models to estimate the treatment effect of Reading Power on letter names and letter sounds. These models include independent samples t-test, hierarchical linear models (in which students were nested within schools), and Poisson regression. In all cases, the results were consistent in that the effect of the treatment on either outcome was not significantly different from zero. For simplicity, only the results from the multiple linear regression are presented.

Where Y_i is the *i*th student's post-intervention letter-name identification or letter-sound identification score, β_0 is the average value for the post-intervention letter-name identification or letter-sound identification score for students in the control group; β_1 is a coefficient reflecting the relationship between Reading Power tutoring and the post-intervention assessment score; X_{1i} is a dichotomous variable indicating whether student *i* was in the treatment or control group (i.e., 0 = control group, 1 = treatment group); $X_{2i}$is the baseline assessment score (either letter names or letter sounds) for student *i* that was included to take into account baseline differences between the treatment and control groups; and β_2 is the coefficient reflecting the relationship between the baseline assessment score and the post-intervention score.

In addition to carrying out the analysis above, the evaluation team examined the missing data patterns present in the data and discussed the implications of the missing data on the estimate of the treatment effects.

The evaluators hypothesized that students in Reading Power would have greater gains in letter names and letter sounds than students in the control group. This would indicate that Reading Power results in a significant improvement in letter-name identification and letter-sound identification compared with other LAP-ELA interventions.

Results

The evaluation team performed descriptive and inferential analyses to answer the evaluation question "*Do kindergarteners and 1st graders who receive Reading Power tutoring demonstrate a greater increase in letter-name and letter-sound identification from pre-intervention assessment to post-intervention assessment than similar students who receive other LAP services (but not Reading Power)?*" This section presents these results.

The study found that, for students in both the treatment and control groups, the average number of letter-names and letter-sounds identified correctly increased from the beginning to end of the study period. The study did not find any statistically-significant differences between the treatment and control groups in the size of that increase: both groups had similar average baseline and post-intervention scores on the two outcomes of interest.

The following tables and figures provide the details of those findings.

Table 3, below, presents descriptive statistics by grade level and treatment group for student scores on assessments of letter-name identification at baseline and post-intervention.

Table 3: Comparison of Baseline and Post-Intervention Scores on Assessments of Letter-Name Identification

Group	Baseline Mean (SD), N	Post-treatment Mean (SD), N
Kindergarten Reading Power	21 (6.43), 42	23.74 (4.3), 42
Kindergarten Control	22.05 (4.5), 38	24.39 (3.01), 38
1st Grade Reading Power	24 (4.06), 37	25.32 (2.15), 37
1st Grade Control	24.44 (3.17), 34	25.44 (2.43), 34

The results in Table 3 suggest that, for both kindergarten and 1st-grade students in the sample, students in the treatment and control groups had similar average baseline and post-intervention scores on assessments of letter-name identification.

Table 4, below, presents descriptive statistics by grade level and treatment group for student scores on assessments of letter-sound identification at baseline and post-intervention.

Table 4: Comparison of Baseline and Post-Intervention Scores on Assessments of Letter-Sound Identification

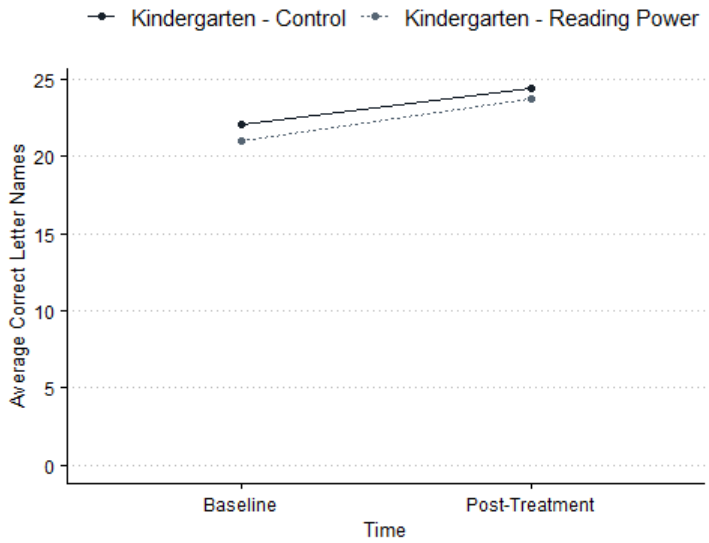
Group	Baseline Mean (SD), N	Post-treatment Mean (SD), N
Kindergarten Reading Power	18.71 (7.55), 42	25.33 (5.14), 42
Kindergarten Control	17.76 (6.98), 38	23.58 (6.05), 38
1st Grade Reading Power	25.54 (4.84), 37	28.16 (2.83), 37
1st Grade Control	26.03 (4.82), 34	27.97 (2.94), 34

As with the results for letter names, the results in Table 4 suggest that, for both kindergarten and 1st-grade students in the sample, students in the treatment and control groups had similar average baseline and post-intervention scores on assessments of letter-sound identification.

From baseline to post-treatment, 15% of kindergarten students and 19% of 1st-grade students dropped out of the study. This pattern was similar between the treatment and control groups.

Figure 1, below, presents the average score on assessments of letter-name identification by group (treatment and control) for kindergarten students in the sample.

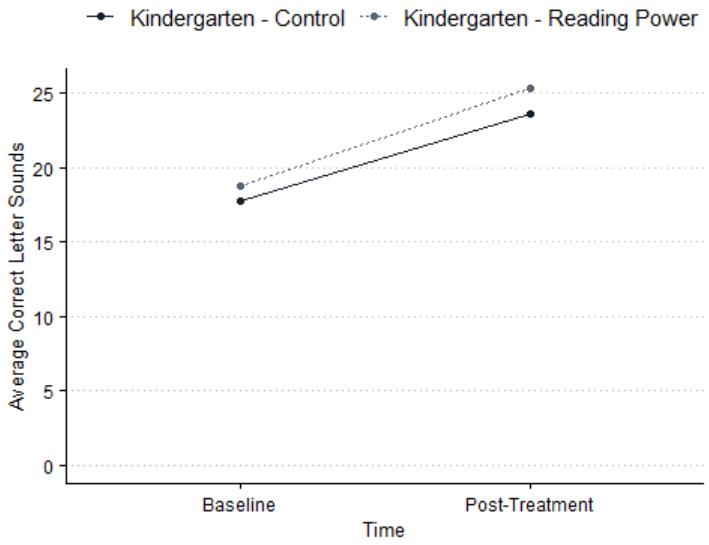
Figure 1: Average Score on Assessments of Letter-Name Identification by Group – Kindergarten Students



As shown in Figure 1, for the kindergarten students in both the treatment and control groups, the average number of letter-names identified correctly increased from the beginning to end of the study period. Additionally, the average number of letter-sounds identified correctly appears to be slightly lower for the treatment group, both at baseline and after intervention (although the results later show that the baseline and post-treatment differences were not statistically significant).

Figure 2, below, presents the average score on assessments of letter-sound identification by group (treatment and control) for kindergarten students in the sample.

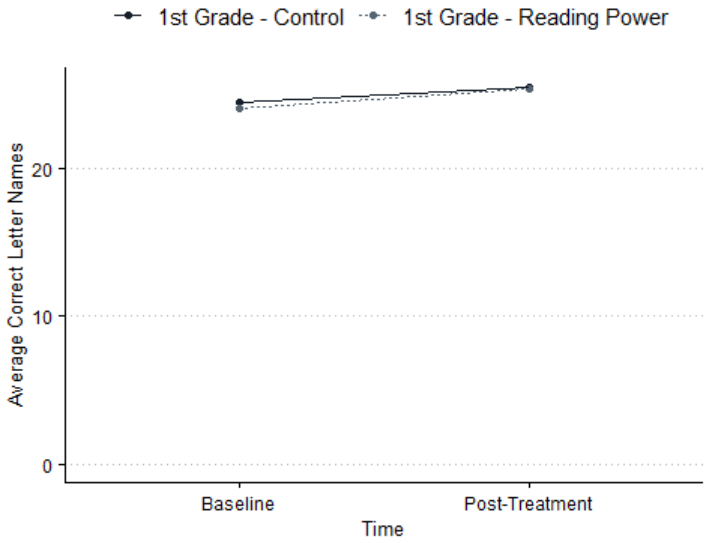
Figure 2: Average Score on Assessments of Letter-Sound Identification by Group – Kindergarten Students



As shown in Figure 2, for the kindergarten students in both the treatment and control groups, the average number of letter-sounds identified correctly also increased from the beginning to end of the study period. Additionally, the average number of letter-sounds identified correctly appears to be slightly lower for the control group, both at baseline and after intervention (although the results later show that the baseline and post-treatment differences were not statistically significant). This result suggests that the treatment and control group had a similar rate of growth in letter sounds over the study period.

Figure 3, below, presents the average score on assessments of letter-name identification by group (treatment and control) for 1st-grade students in the sample.

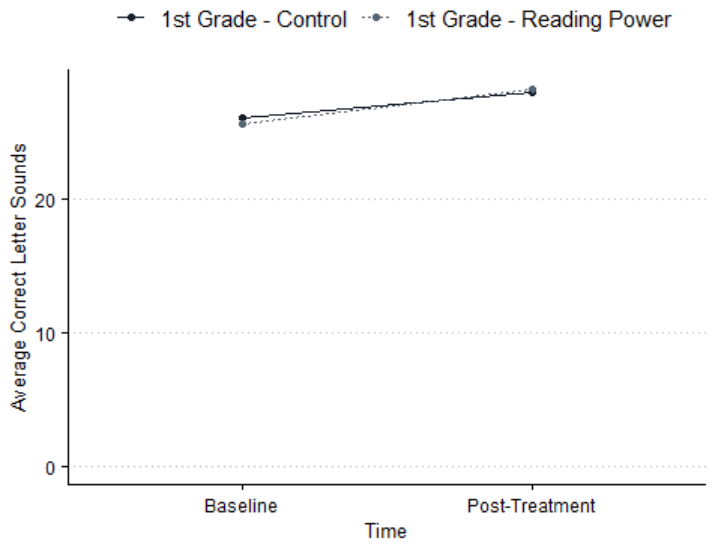
Figure 3: Average Score on Assessments of Letter-Name Identification by Group – 1st-Grade Students



As shown in Figure 3, for the 1st-grade students in both the treatment and control groups, the average number of letter-names identified correctly increased from the beginning to end of the study period. Additionally, the average scores for the treatment and control groups appear to be nearly identical both at baseline and after treatment, suggesting that the average growth for both groups was similar.

Figure 4, below, presents the average score on assessments of letter-sound identification by group (treatment and control) for 1st-grade students in the sample.

Figure 4: Average Score on Assessments of Letter-Sound Identification by Group – 1st-Grade Students



As shown in Figure 4, for the 1st-grade students in both the treatment and control groups, the average number of letter-names identified correctly increased from the beginning to end of the study period. Additionally, the average scores for the treatment and control groups appear to be nearly identical both at baseline and after treatment, suggesting that the average growth for both groups was similar.

Table 5, below, presents the results from the regression models used to answer the evaluation question for each grade and outcome. Cells containing an "*" indicate that a variable was a statistically significant ($p < 0.05$) predictor of that student outcome for that grade level and outcome.

Table 5: Model Results by Grade

Variable	Coefficient (SE) Kindergarten - Correct Letter Names	Coefficient (SE) Kindergarten - Correct Letter Sounds	Coefficient (SE) 1st Grade - Correct Letter Names	Coefficient (SE) 1st Grade - Correct Letter Sounds
Intercept	13.11** (1.13)	12.29** (1.08)	13.36** (1.17)	18.14** (1.49)
Baseline Score	0.51** (0.05)	0.63** (0.05)	0.49** (0.05)	0.38** (0.06)
Reading Power	-0.11 (0.5)	1.18 (0.71)	0.11 (0.33)	0.36 (0.49)

Note: * denotes statistical significance at $p < 0.05$,
** denotes statistical significance at $p < 0.01$.

These results are described below for each grade and outcome:

- *Kindergarten Letter Names*: The baseline score for letter-name identification for kindergartners was predictive of their post-intervention score ($B = 0.51$, $p < 0.01$). However, no difference was found between the treatment and control groups in the post-intervention score ($B = -0.11$, $p > 0.05$).
- *Kindergarten Letter Sounds*: The baseline score for letter-sound identification for kindergartners was predictive of their post-intervention score ($B = 0.63$, $p < 0.01$). However, no difference was found between the treatment and control groups in the post-intervention score ($B = 1.18$, $p > 0.05$).
- *1st Grade Letter Names*: The baseline score for letter-name identification for 1st-graders was predictive of their post-treatment score ($B = 0.49$, $p < 0.01$). However, no difference was found between the treatment and control groups in the post-intervention score ($B = 0.11$, $p > 0.05$).
- *1st Grade Letter Sounds*: The baseline score for letter-sound identification for 1st graders was predictive of their post-intervention score ($B = 0.38$, $p < 0.01$). However, no difference was found between the treatment and control groups in the post-intervention score ($B = 0.36$, $p > 0.05$).

Conclusions, Study Limitations, and Recommendations

Conclusions

The results of this impact evaluation suggest that Reading Power produces growth in students' letter-name identification and letter-sound identification that is similar to the growth seen in response to other LAP-ELA interventions provided to kindergarten and 1st-grade students. The fact that the control and treatment groups had similar baseline scores on assessments of letter-name identification and letter-sound identification suggests that the random assignment used in the study was successful at balancing both groups in terms of their baseline characteristics.

Study Limitations

The researchers have identified several limitations to the study, any one of which could potentially be a reason that no statistically-significant difference between the two groups was found.

1) **Covid-Related Implementation Limitations** – COVID affected the sample size. Because fewer schools partnered with Reading Power during the pandemic, fewer tutors were available to provide tutoring and fewer students were available to include in the study. In addition, the pandemic interrupted the timing of Reading Power rollout. Washington Reading Corps was unable to consistently implement the program before the pandemic changed the education landscape.

2) **Assessment Limitations**: The researchers encountered multiple challenges with the tools Reading Power currently uses to assess student reading growth in for kindergarten and 1st-grade students:

   - While the Reading Power program supports skills building in multiple areas for these two grades – alphabet knowledge (letter-name and letter-sound identification), print awareness, phonological awareness, and beginning word recognition – the Reading Power assessment tool used for kindergarten and 1st-grade students measures only changes in students' ability to correctly identify letter-names and letter-sounds.
   - The Reading Power letter-name assessment measures only upper-case letter names rather than both upper- and lower-case letter names, restricting the range of possible scores to a maximum of 26, rather than 52. Among students assessed, the scores ranged from 2 to 26 at baseline and the median score was 25 for both the treatment group and the control group. Because the range of possible scores was restricted to 26, the median scores at baseline were almost as high as the maximum scores, leaving little room for growth.
   - The letter-sound identification measures only 30 phonemes, rather than the full range of 44 phonemes, also restricting the range of possible scores. Among students assessed, the scores ranged from 2 to 30 at baseline and the median score was 24 for the treatment group

and 23 for the control group. Again, because the range of possible scores was restricted, the median scores at baseline were almost as high as the maximum scores, leaving little room for growth.

In addition, the psychometric properties of the assessment tools used in this study are unknown and it is unclear whether the assessment tools were administered in a way that aligns with the processes for widely-recognized assessments such as Fastbridge or the IGDIs.

3) **Potential Similarity of Assignment Conditions**: Only very limited information was available about what specific LAP-ELA intervention(s) students in the control group received and how those interventions may have differed from Reading Power. It is also unclear whether Reading Power was used as "double dose" of skills-building for students (i.e., treatment students received both Reading Power during pull-out sessions AND other LAP-ELA interventions provided during regular classroom instruction time, while control group students receive ONLY other LAP-ELA interventions provided during regular classroom instruction time). It is also possible, therefore, that Reading Power and the other LAP-ELA interventions provided to the control group students were similar in content, dosage, and approach and, as a result produced similar results in terms of their impact on students' growth in letter-name identification and letter-sound identification.

4) **Insufficient Sample Size**: While the power analysis for this study suggested that, to detect a difference of at least 0.25 (Cohen's d, with a power of 0.80) on post-intervention assessment scores, a sufficient sample size would be 398 students, the study included only 151 students. As a result, it is possible that the reason this study did not show an effect is because the sample size was not large enough to be able to detect the true effect of Reading Power on participating students.

These limitations may have contributed to the fact the study found no statistically-significant differences between the two groups for the outcomes of interest.

Recommendations

The authors recommend several courses of action for WRC to address the limitations listed above and better determine the impact of the Reading Power program:

1) **Make Changes to the Assessment Process and Tools:** The researchers recommend that WRC make the following changes to its assessment process and tools:

   - To assess letter-name identification skills more accurately, include the full range of 52 upper-case and lower-case letters in the assessment tool, rather than only the 26 upper-case letters currently measured and more accurately
   - To assess letter-sound identification skills more accurately, include the full range of 44 phonemes in the assessment tool, rather than only the 30 phonemes currently measured.

- Assess progress for all of the reading-skill areas WRC tutoring supports, not only letter-name and letter-sound identification skills. That is, for kindergarten and 1st-grade students, assess print awareness, phonological awareness, and beginning word recognition as well as alphabet knowledge (letter-name and letter-sound identification), and, if future evaluations include 2nd through 4th-students, assess phonemic awareness, word recognition, oral reading fluency, and comprehension for those students.
- Ensure that tutors use standardized assessment administration processes when measuring student performance for outcomes of interest.

Multiple validated tools measure alphabet knowledge, print awareness, phonological awareness, and beginning word recognition. Using a validated assessment tool that not only includes the full range of possible answers (e.g., 52 letter names and 44 letter sounds) but also measures all skills being tutored for will improve the likelihood of detecting an effect among the target population of students WRC serves.

2) **Ensure that Reading Power is Offered as a "Double Dose" Intervention.** The researchers recommend that, to avoid the problem of similarity of treatment conditions, WRC will ensure that sites that offer Reading Power as a pull-out tutoring program delivered in addition to other LAP-ELA interventions (which are delivered in the classroom), so that the Reading Power intervention represents a "double dose" of skills-building for students. In other words, treatment students will both receive Reading Power during pull-out sessions AND receive other LAP-ELA interventions during regular classroom instruction time, while control group students will receive ONLY other LAP-ELA interventions provided during regular classroom instruction time and will not be pulled out for WRC tutoring.
3) **Increase the Sample Size:** For future evaluation studies, it will be important for WRC to find a way to include more sites and/or more students in order to create a sample pool that is large enough to detect a difference of at least 0.25 (Cohen's d, with a power of 0.80) on post-intervention assessments.

Appendices

Appendix A: Assessments

Grades K & 1 Student Assessment: Letter Identification

A	F	K	P	W
Z	B	H	O	J
U	C	Y	L	Q
M	D	N	S	X
I	E	G	R	V
T				

Grades K & 1 Student Assessment: Letter Sounds

a	f	k	p	w
z	b	h	o	j
u	c	y	l	q
m	d	n	s	x
i	e	g	r	v
t	ch	sh	ck	th

Grades K & 1 Student Assessment: Response Form

WRC Member:	**Date:**
Student Name:	**Grade:**

Upper-Case Letter Identification

Using the "Upper-Case Letter Identification" sheet, ask the student to name each letter. If they ask for help, you can say, "try your best" and "you can skip it if you don't know".

Indicate correct responses with: /√/

Indicate no response with: /x/

If the child responds incorrectly, record what he/she says.

	A	F	K	P	W	Z	B	H	O	J	U	C	Y	L	Q	M	D	N	S	X	I	E	G	R	V	T
Letter Name																										

Total correct:_____of 26 letter names/minute

Lower-Case Letter-Sound Identification

Using the "Lower-Case Letter-Sound Identification" sheet, ask the student to make the sound of each letter. If they ask for help, you can say, "try your best" and/or "you can skip it if you don't know". Before beginning, tell them that you only want one sound for each letter, even if they know more than one. Accept as correct any possible sound the letter makes. If they notice that you are now using lowercase letters instead of uppercase, explain that we use lowercase letters more often to read and write than uppercase letters. If they think they'd be faster at naming them this way, you can re-do the first part with this prompt.

	a	f	k	p	w	Z	b	h	o	j	u	c	y	l	q	m	d	n	s	x	i	e	g	r	v	t
Letter Sound																										

	ch	sh	ck	th
Letter Combination Sound				

Total correct:_____of 30 sounds/minute

Appendix B: Format of Data Files Provided by WRC

Table 6: Assessment Data

Variable Name	Variable Description
SchoolName	The name of the school in which the student was enrolled while taking the assessment.
StudentID	A unique alphanumeric student identifier. This should use the same naming convention that WRC uses to provide students unique ID values (e.g.,3-digit member id, first student initial, last student initial, etc.).
AssessmentDate	The date the intervention was administered for a particular student in YYYY-MM-DD format (e.g., 2022-10-01)
AssessmentName	Name of assessment that was administered for a particular student on the particular assessment date (e.g., Fastbridge- Letter Name, Fastbridge - Letter Sound).
Score	The numeric score achieved by the particular student (e.g., 23, 31)

Table 7: Student Randomization

Variable Name	Variable Description
SchoolName	The name of the school in which the student was enrolled during the study.
StudentID	A unique alphanumeric student identifier. This should use the same naming convention that WRC uses to provide students unique id values e.g.,3-digit member id, first student initial, last student initial, etc.). This studentID should be generated for all students who are eligible for LAP-ELA interventions.
Grade	The grade in which the student is enrolled during the 2022-23 school year (e.g., “K”, “1”)
Assignment	A dichotomous variable indicating the group to which each student was assigned (1 = treatment, 0 = control).
MemberID	The alphanumeric unique member identifier (this is the same member number that is part of studentID).
MemberFullName	First and last name of member, formatted in the following way: “Lastname, Firstname” (e.g., “Medhanie, Amanuel”)

Table 8: Members and Schools Served

Variable Name	Variable Description
SchoolName	The name of the school to which the member was assigned during the study.
MemberID	The alphanumeric unique member identifier (the same member number that is part of studentID).
MemberFullName	First and last name of member, formatted in the following way: “Lastname, Firstname” (e.g., “Medhanie, Amanuel”)

Table 9: Reading Power Sessions

Variable Name	Variable Description
SchoolName	The name of the school in which the student was enrolled during the study.
MemberID	The alphanumeric unique member identifier indicating which member delivered the Reading Power intervention session to the particular student on the particular date.
StudentID	A unique alphanumeric student identifier. This should use the same naming convention that WRC uses to provide students unique id values (e.g., 3-digit member id, -, first student initial, last student initial, etc.).
Date	The date the intervention was administered for a particular student in YYYY-MM-DD format (e.g., 2022-10-01)
SessionMinutes	A numeric variable indicating the amount of time a particular student received the particular intervention (measured in minutes) on the given date.

Table 10: Reading Power Lessons

Variable Name	Variable Description
SchoolName	The name of the school in which the student was enrolled during the study.
MemberID	The alphanumeric unique member identifier indicating which member delivered the relevant intervention to the particular student on the particular date.
StudentID	A unique alphanumeric student identifier. This should use the same naming convention that WRC uses to provide students unique id values (e.g., 3-digit member id, -, first student initial, last student initial, etc.).
Date	The date the intervention was administered for a particular student in YYYY-MM-DD format (e.g., 2022-10-01)
Lesson	A numeric variable indicating the Reading Power lesson that was administered on that day to that student (1-36).

Table 11: Recruited Schools

Variable Name	Variable Description
SchoolID	A numeric code identifying the school in which the student was enrolled while taking the assessment.
SchoolName	The name of the school (e.g., Greenwood Elementary)
District	The district in which the school exists
InSample	A dichotomous variable indicating whether the school participated in the study (1 = Yes, 0 = No).
FRL	Percentage of students who received free or reduced priced lunch during the 2021-22 school year (e.g., 62%).

Variable Name	Variable Description
MetELAStandards	Percentage of K-4 students who met ELA standards during the 2021-22 school year (e.g., 57%).
HighELAGrowth	Percentage of K-4 students who exhibited high ELA growth over the 2021-22 school year (e.g., 45%).
EnglishLanguageLearners	Percentage of English Language Learners during the 2021-22 school year (e.g., 23%).
RegularlyAttend	Percentage of students who regularly attended during the 2021-22 school year (e.g., 89%).
EnrollmentSize	Number of students enrolled during the 2021-22 school year (e.g., 1,432).

Appendix C: Timeline & Responsibilities

RESPONSIBILITY	TASK	DUE DATE	DELIVERABLE
EVAL TEAM	EVAL TEAM provides regular communication about progress and challenges.	Biweekly	EVAL TEAM meeting attendance and/or emails, as appropriate
EVAL TEAM	EVAL TEAM provides regular expenditure and progress reports.	Quarterly	EVAL TEAM quarterly expenditure & progress reports
EVAL TEAM	EVAL TEAM creates mock-up of initial report template, including descriptions of final study design, and data collection tools as well as template language for WSC description, findings, and recommendations section, and delivers to WSC Team.	7/18/22 to 8/5/22	EVAL TEAM sends copy of updated plan to WRC TEAM
WRC TEAM	WRC TEAM develops and distributes introductory communication to all WRC sites about research study, expectations, etc.	08/05/2022	WRC TEAM sends copy of materials to EVAL TEAM
EVAL TEAM	EVAL TEAM meets with WRC program staff to finalize/approve report template, which will serve as the final study design document, and confirms details, schedule, and responsibilities for all evaluation activities.	09/01/2022	EVAL TEAM sends email to WRC Team confirming details and activity schedule
WRC TEAM	WRC TEAM adjusts structure of the intervention (i.e., modifications to the Reading Power intervention and treatment approach).	09/16/2022	WRC TEAM sends copies of the content, scope, and sequence of the intervention training and of participant literacy skills assessments to EVAL TEAM
EVAL TEAM	EVAL TEAM works with WRC to finalize agreement on data sharing tasks and responsibilities, including processes to be used by WRC staff, site staff, and members to collect and enter data on participants, services provided, and literacy skills growth measures.	09/30/2022	EVAL TEAM sends copy of agreement and details to WRC TEAM
WRC TEAM	WRC TEAM provides EVAL TEAM with a list of school sites serving K-1 students in the 2022-23 school year and hosting members engaged with Reading Power students (for random assignment of study participants).	10/07/2022	WRC TEAM sends school list to EVAL TEAM
WRC TEAM	WRC TEAM develops and distributes information about intervention and assessment requirements for treatment and comparison groups to project site staff and members.	10/07/2022	WRC TEAM sends copy of materials to EVAL TEAM for review prior to distribution to sites.

RESPONSIBILITY	TASK	DUE DATE	DELIVERABLE
WRC TEAM	WRC TEAM makes sure screening administration and student selection are executed across all participating school sites (i.e., identify which students are eligible to be in the random assignment based on their baseline assessment scores or other criteria defined by school sites?).	10/14/2022 to 10/31/22	WRC TEAM sends all lists of students eligible for random assignment (with baseline assessment scores) to EVAL TEAM (use the data file formats specified in the mock-up).
WRC TEAM	WRC trains members in Reading Power Model	11/1/22 to 11/30/22	WRC Team confirms that all members have received initial Reading Power training (K-1 content)
EVAL TEAM	EVAL TEAM completes random assignment of schools, members, and students to treatment and comparison group and returns data to WSC team.	11/15/2022	EVAL TEAM sends random assignment list to WRC TEAM
WRC TEAM	WRC members conduct initial assessment for treatment and control groups and begin 12-week Reading Power intervention for treatment group	Week of 1/2/23	WRC Team sends confirmation of completion to Eval Team
WRC TEAM	WRC TEAM tracks students attendance and member services provided (to determine participants dosage received) as well as pre-intervention and final assessments administered and communicates with EVAL TEAM as necessary to discuss implications for evaluation.	1/3/23 to 3/31/23	WRC TEAM sends copy of student attendance and member service info EVAL TEAM
WRC TEAM	WRC members complete delivery of Reading Power intervention across all participating school sites and conduct final assessment for treatment and control groups	03/31/2023	WRC TEAM sends email confirming completion to EVAL TEAM
WRC TEAM	WRC TEAM delivers all data from assessments to EVAL TEAM for analysis and reporting.	04/07/2023	WRC TEAM sends complete data file EVAL TEAM
EVAL TEAM	EVAL TEAM analyzes student assessment data (pre-intervention and final scores) and conducts statistical modeling where appropriate, to test all research hypotheses and determine whether statistically-significant differences exist between the treatment and comparison (control) groups in outcomes.	4/14/23 to 5/26/23	EVAL TEAM sends email confirming completion of data analyses and preliminary results from student data to WRC TEAM
EVAL TEAM	EVAL TEAM creates and delivers an evaluation report DRAFT and a three-year evaluation plan DRAFT that describes analyses conducted, findings that resulted, and recommendations	06/15/2023	EVAL TEAM sends draft of results from student data to WRC TEAM

RESPONSIBILITY	TASK	DUE DATE	DELIVERABLE
	for next steps, as appropriate (this evaluation covers Year 3 of the three-year project).		
WRC TEAM	WRC TEAM Reviews the (DRAFT) Evaluation Report and Three Year Plan (DRAFT) for revision, correction and works with EVAL Team throughout this period for finalization.	07/15/2023	WRC Team sends edits and suggested changes for draft to EVAL TEAM
EVAL TEAM	EVAL TEAM provides a comprehensive, clearly-written, and 100% reproducible FINAL evaluation report (i.e., all data cleaning, preparation, analyses, and report writing are created through shareable code), which incorporates edits and feedback from WRC Team and describes analyses conducted, findings that resulted, and recommendations for next steps, as appropriate.	08/15/2023	EVAL TEAM sends final evaluation report to WRC TEAM
EVAL TEAM	EVAL TEAM provides a FINAL three-year evaluation plan to be included in the subsequent WRC grant application and works with WRC and ESD to align the plan to WSC long term strategic plan evaluated. This evaluation plan will inform the 2023-26 project.	09/01/2023	EVAL TEAM sends final three-year evaluation plan to WRC TEAM
EVAL TEAM	EVAL TEAM will work with WRC's grant writing consultant and ESD staff to ensure the evaluation plan fits the constraints of the Egrants portal to ensure smooth submission of the grant application.	09/01/2023	EVAL TEAM meetings, emails, and phone calls with ESD staff
EVAL TEAM	EVAL TEAM remains available to answer questions regarding the evaluation analyses, results, and recommendations.	10/01/2023	EVAL TEAM meetings, emails, and phone calls with ESD staff
EVAL TEAM	Upon project completion, the EVAL TEAM will deliver to WSC data files collected during project, including individual level and site level data (with identifiers and contact information) and will develop and delivers a codebook that clearly explains each variable and variable value and provides a full description of the formulas and procedures used to analyze data; and the weight calculations for all data sets; and the methods used to control for attrition and non-response.	11/01/2023	EVAL TEAM sends all materials to WRC TEAM

Appendix D: Reading Power Tutoring Guide

Reading Power - A Tutoring Program Supported by: