Expanding Opportunities to Successfully Support Early Readers: A Five-Year Study of Reading Partners Colorado

May 2017

Mile High United Way Social Innovation Fund Grant
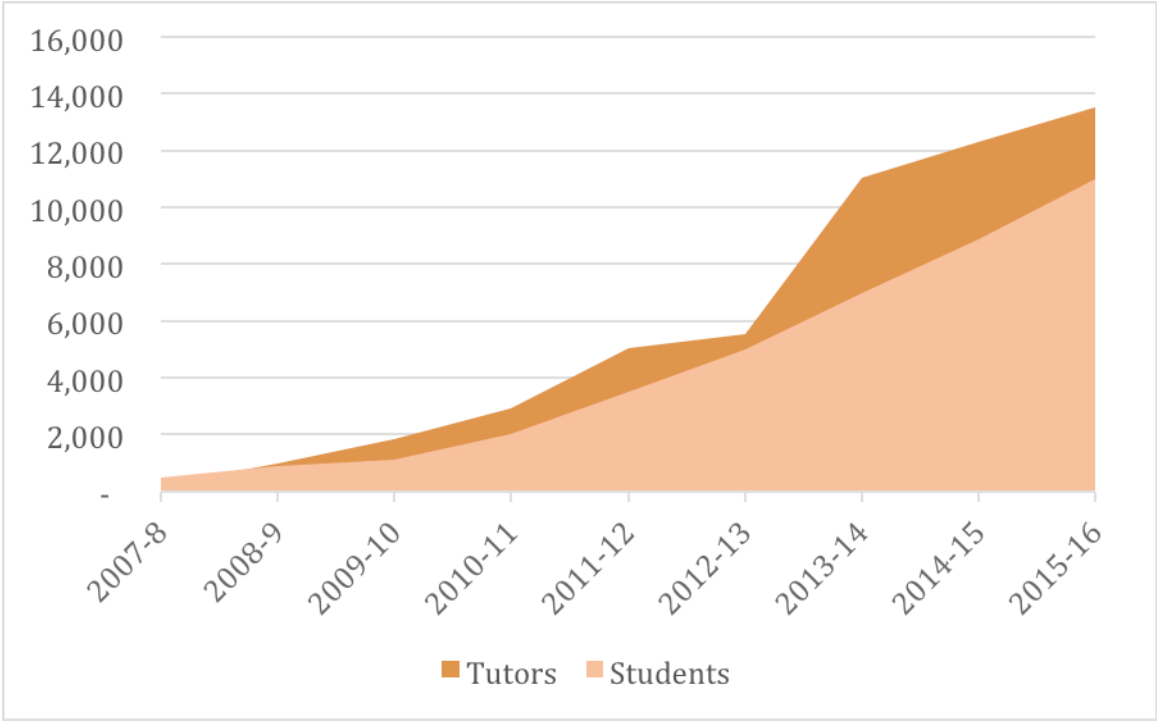Subgrantee: Reading Partners

Prepared by
Robert Reichardt, Abby McClelland, Jack Hill, and Dale DeCesare
Augenblick, Palaich and Associates,
Independent Evaluator

The Social Innovation Fund (SIF) was a program that received funding from 2010 to 2016 from the Corporation for National and Community Service, a federal agency that engages millions of Americans in service through its AmeriCorps, Senior Corps, and Volunteer Generation Fund programs, and leads the nation's volunteer and service efforts. Using public and private resources to find and grow community-based nonprofits with evidence of results, SIF intermediaries received funding to award subgrants that focus on overcoming challenges in economic opportunity, healthy futures, and youth development. Although CNCS made its last SIF intermediary awards in fiscal year 2016, SIF intermediaries will continue to administer their subgrant programs until their federal funding is exhausted.

Executive Summary

Founded in 1999, Reading Partners recruits, trains, and places community volunteers into high-need schools to provide tutoring for students who are behind grade level in reading. Reading Partners collaborates with school leaders and teachers to transform a dedicated school space into a reading center, and creates twice weekly opportunities for students to receive tailored, one-on-one instruction from a trained and supervised community volunteer. These volunteers use a structured, research-based curriculum adapted for each student. Community volunteers and students receive ongoing support from Reading Partners' Reading Center Coordinators, who provide volunteer training, observations, coaching, and progress monitoring to ensure that students are meeting their individual literacy goals. Reading Partners' vision is that one day all children will have the reading skills necessary for them to reach their full potential. Reading Partners has experienced rapid growth over the past decade, illustrated in Figure E.1, and currently serves more than 11,000 students in over 200 schools across 10 states and the District of Columbia.

Figure 1: Reading Partners Has Grown Rapidly Since its Inception

Social Innovation Fund Evaluation Overview

Reading Partners began operating in Colorado in 2012, with funding from a Social Innovation Fund (SIF) grant from the Corporation for National and Community Service (CNCS) awarded to Mile High United Way (MHUW). This five-year grant supported Reading Partners' expansion and implementation in Colorado from 2012-2017. The SIF grant required Reading Partners to engage an independent, third-party evaluator to study the program's implementation and impact. In 2012, Reading Partners hired Augenblick, Palaich and Associates (APA), a national education research and evaluation company, to serve as the external evaluator.

Over the course of its evaluation, APA produced a detailed evaluation plan outlining the study design; annual reports on project findings; and interim memoranda addressing key aspects of the study. This final summative report incorporates data from all five years of the project and has two primary components:

- An **impact evaluation,** using a quasi-experimental design with propensity score matching, designed to meet a "moderate" level of evidence under the SIF evidence framework, examining whether reading skills of students in Reading Partners improved significantly more than those of similar students not served by the program. The impact analysis draws upon Reading Partners' and school district administrative data, as well as interviews and surveys of school leaders, to examine student literacy outcomes during the 2013-14, 2014-15, and 2015-16 school years.
- An **implementation evaluation** exploring whether Reading Partners Colorado implemented the program with fidelity to Reading Partners' model. Implementation evaluation findings span the five-year period of the SIF grant and are based on data from multiple sources, including Reading Partners student folder reviews, Reading Partners administrative data, direct observations of tutoring sessions, and surveys and interviews with key stakeholders.

Impact Findings

The impact evaluation examines whether the reading skills of students served by Reading Partners improved more than those of similar students not served by the program, as measured by state-mandated, pre- and post-, school-based literacy assessments. The analysis utilizes a quasi-experimental design with a propensity score matching approach, and includes data from the 2013-14, 2014-15, and 2015-16 school years. The three-year sample includes a total of 698 Reading Partners students and 853 similar comparison students, for a total sample size of 1,551. Comparison students for the study were drawn either (1) from schools with Reading Partners sites (where comparison students included only students who were *not* served by Reading Partners) or (2) from a separate set of identified comparison schools. APA selected 30 comparison schools that were similar to Reading Partners schools in terms of geographic location, racial/ethnic makeup of students, poverty rate among students and assessment administered. Students in the final sample were matched on their assessment pre-scores and demographic characteristics. The final matched sample was very well-balanced and the two groups of students were comparable.

The impact evaluation sought to answer four research questions:

1. Does Reading Partners tutoring lead to improved near-term reading achievement for students in grades one through three when compared to similar students who do not receive tutoring?

2. Do differences in reading achievement between students who receive Reading Partners tutoring and similar students who are not in Reading Partners increase as students receive more tutoring?

3. Are there differential impacts of Reading Partners tutoring on different student groups, including English-language learners (ELL) vs. non-ELL students, boys vs girls, grade level, and different races?

4. What is the effect of participating in Reading Partners for multiple years?

Impact Finding #1
Students who participated in Reading Partners during one school year had spring reading assessment scores that are significantly higher than the scores of similar students who do not participate in the program.

Findings in response to each of these questions are summarized, in turn, below.[1]

Overall Impact: On average, students who participated in Reading Partners during one school year had spring reading assessment scores that were significantly higher than the scores of similar students who did not participate in the program, controlling for fall assessment performance. For the average Reading Partners student, this improvement was equivalent to moving from the 15th percentile to the 21st percentile. This is an effect size of approximately 0.14, which is consistent with the average effect size for one-on-one tutoring programs found in a 2009 meta-analysis (Slavin, Lake, Davis & Madden, 2009). This effect size is roughly equivalent to that found in a 2015 experimental study of the Reading Partners program conducted by MDRC, which found a significant, positive effect of the program with an effect size of about 0.10 (Jacob, Armstrong & Willard, 2015). This study differs from the MDRC study in both methodology and study population.

Program Dosage: APA used two models to investigate the influence of program dosage on student literacy outcomes. The first model included only Reading Partners students. APA did not find significant effects based on dosage using this model. However, due to its focus exclusively on Reading Partners students, this model may have lacked sufficient statistical power to detect an effect.

Impact Finding #2
For each ten additional Reading Partners tutoring sessions received, there was a statistically significant increase in student reading assessment scores.

The second model compared Reading Partners students to comparison students not served by the program. In this model, APA found that, for every ten additional Reading Partners tutoring sessions received, there was a statistically significant increase in student reading assessment scores: 0.7 point Normal Curve Equivalent increase for every ten additional sessions.

Differential Program Impact: APA did not find statistically significant differences in program impact based on grade level, gender, or race or ethnic identity. However, APA did identify a statistically significant differential effect for ELL students. Reading Partners tutoring had a significantly larger impact on ELL students than general population students in the program, illustrated in Figure E.2, below. Although ELL students who did not attend the program scored lower on spring assessments than non-ELL students who did not attend the program, that trend is reversed for students who attended Reading Partners: ELL students who attended Reading Partners scored higher on the spring assessment than their non-ELL counterparts who also attended the program, while controlling for fall assessment performance. This demonstrates that Reading Partners' program, as implemented in Colorado, was

Impact Finding #3
*As implemented in Colorado, Reading Partners was particularly effective for ELL students, helping them to outperform their non-ELL peers, both in and out of the program.*

---

[1] For each analysis model, APA accounted for both student- and school-level factors using a Hierarchical Linear Model (HLM) with student-level variables at level one and school-level indicators at level two.

particularly effective for ELL students, helping them to outperform their non-ELL peers, both in and out of the program. This is especially important as 55% of Reading Partners students included in the analyses were identified as ELL.

Figure E.2: Relative Spring Literacy Assessment Score of ELL students in and out of Reading Partners
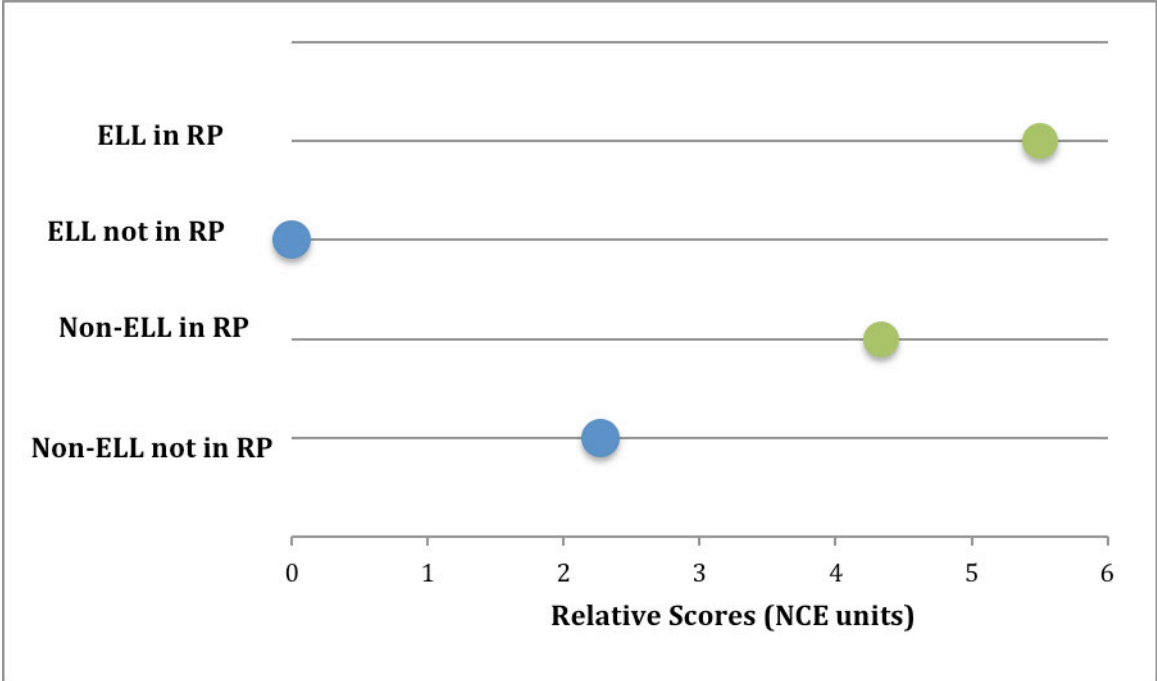
Participation Duration: The fourth and final impact research question sought to examine the effect of a student attending Reading Partners' program for more than one year, but the available sample of students who attended the program for multiple years was too small to provide sufficient statistical power to answer this question.

Implementation Findings

In brief, the implementation evaluation found that Reading Partners was able to quickly launch, sustain, and implement its program with fidelity in a new region. Reading Partners Colorado was able to secure funding, engage school partners, recruit and train volunteer tutors, identify students meeting program enrollment criteria, assess and develop reading plans for those students, and ultimately deliver literacy tutoring to those students using Reading Partners' curriculum.

Volunteers and school staff both perceived the Reading Partners curriculum as appropriate for struggling readers.

> ***Implementation Finding #1***
> *The Reading Partners program was implemented with fidelity: maintaining funding, identifying school partners, recruiting and training volunteer tutors, identifying students, developing reading plans, and delivering literacy tutoring using the Reading Partners curriculum.*

School leaders reported that Reading Partners required much less of their time and engagement to implement successfully than similar supplemental programs for students; volunteers felt supported by the program and said it was easy to use the curriculum.

Table 1 shows the school districts, number of schools and students served, number of tutors engaged, and the average number of sessions per student for each of Reading Partners' first four years in the state.

Table 1. Growth of Reading Partners Colorado

	Districts	Schools	Students Served	Tutors	Average Sessions per Student
Year 1 2012-13	**APS, DPS**	**8 (all new)**	**323**	**481**	**25**
Year 2 2013-14	APS, DPS, Sheridan[2]	11, (4 new)	558	803	32
Year 3 2014-15	**APS, DPS, Sheridan**	**13 (6 new)**	**770**	**1,332**	**31**
Year 4 2015-16	DPS	14 (9 new)	881	1,219	34

Data source: APA analysis of Reading Partners data

Reading Partners Colorado experienced constant growth in the number of students served, tutors engaged, and the average number of sessions provided to students. At the same time, Reading Partners Colorado experienced a substantial amount of transition in its school and district partnerships. Denver Public Schools (DPS) remained an active partner throughout the study period, while partnerships with Aurora Public schools (APS) and Sheridan Schools were shorter-lived. Even in DPS, Reading Partners Colorado experienced substantial turnover among school partners, with new schools added and existing schools leaving each year, in part due to changes in school needs and resources and in part due to funding issues.

> ***Implementation Finding #2*** *The majority of students participating in Reading Partners received at least 20 tutoring sessions in a school year.*

The majority of students participating in Reading Partners received at least 20 tutoring sessions per year. For example, in 2014-15: 75 percent of students received 20 or more sessions, over 50 percent received 30 or more sessions, and one-third received 40 or more sessions.

Data from reviews of Reading Partners' student folders conducted in 2015-16 also provide insight into the relationships between student enrollment, the number of tutors per student, and the rate of session delivery:

- Students who enrolled earlier in the year generally received more sessions.
- About one in ten sessions was provided by Reading Partners staff, rather than volunteers.

---

[2] Sheridan was not part of the impact or implementation studies.

- More tutoring sessions often translated to more tutors. On average, students worked with two additional tutors for every five additional sessions received.
- Working with multiple tutors did not necessarily mean that students did not have opportunities to build relationships with another caring adult. The median student received 45% of sessions from a singe tutor and over half of students had a primary tutor who provided at least 13 of their sessions.
- Reading Center Coordinators play a pivotal role in ensuring that students receive two tutoring sessions per week and establish strong relationships with their tutors. Specifically, Coordinators broker multiple schedules, including that of the tutor, student, and the student's teacher, to meet program goals for number of sessions and strong tutor-student relationships. Coordinators were effective at navigating any one scheduling challenge (e.g., a student missing a session due to illness). However, challenges with more than one schedule (e.g., volunteer cancelations and difficulty with school scheduling) often reduced the share of students receiving two sessions a week and increased the number of tutors per student.

The implementation evaluation catalogued changes, challenges, and opportunities over the first four years of implementation in Colorado. First, Reading Partners made several important programmatic changes affecting all of its regions:

- Revised curriculum for and increased emphasis on serving students in the early grades (K-3);
- Modified approach to tutor orientation and training;
- Changed the literacy assessment tools used with enrolled students; and
- Revised student enrollment criteria.

Over the same period, Reading Partners Colorado faced multiple evaluation and fundraising challenges and opportunities, including some specifically associated with being a SIF sub-grantee.

Additionally, Reading Partners experienced significant organizational change, at both the national and regional levels. APA used Stevens' *Nonprofit Lifecycle* Model as a conceptual framework to describe these changes (Stevens, 2011). The national organization evolved through the growth stage into the mature stage, a transition characterized by increasingly formalized organizational structures and policies. This included adding professional staff and growing the role, responsibilities, and supports for regional executive directors. In particular, it built professional capacity and created an internal policy system needed to support a mature nonprofit. Its relationship with regional offices evolved to include a more nuanced view of autonomy and authority. During the same period, Reading Partners Colorado navigated the growth stage, with a focus on establishing a strong presence in and relationships with the local education community, including schools and district leaders, tutors, and funders.

Conclusion

Throughout this study, Reading Partners Colorado implemented the program with fidelity. This adherence to Reading Partners' program model translated to significant positive literacy outcomes for participating students, with students who received more tutoring sessions seeing even greater literacy gains. Notably, Reading Partners Colorado was especially effective for English Language Learner (ELL) students, helping them to outperform their non-ELL peers, both in and out of the program.

Looking forward, it will be important for Reading Partners to encourage program innovation so that its staff remains engaged and the organization can capitalize on new opportunities. For example:

- Reading Partners Colorado may benefit from its new freedom to include AmeriCorps members in program delivery.[3] These new team members could be engaged as Reading Center Coordinators (potentially reducing costs) or as full-time tutors who could quickly grow their literacy expertise.
- Reading Partners could pursue new uses for its strong program curriculum, such as in a summer school setting or through licensing with other programs.
- Reading Partners may find opportunities to introduce new programs and foster future organizational growth by leveraging its expertise in volunteer engagement, school partnerships, and program expansion and replication.

Pursuing opportunities such as these will help ensure that Reading Partners can continue to evolve and grow as it seeks to reach greater numbers of students across multiple states and communities.

---

[3] With the end of the five-year SIF grant, Reading Partners Colorado will no longer face restrictions on use of AmeriCorps members as Reading Center Coordinators.

Contents

Introduction

In 2012, Reading Partners began operating in Colorado with funding from the Social Innovation Fund (SIF) grant from the Corporation for National and Community Service (CNCS) awarded to Mile High United Way (MHUW).[4] The SIF grant has supported Reading Partners' expansion and growth in Colorado for the past five years. In 2012, Reading Partners Colorado hired Augenblick, Palaich and Associates (APA) to serve as external evaluator for an impact evaluation and an implementation evaluation, both in accordance with SIF grant requirements. Over the course of this project, APA produced a highly detailed Sub-Grantee Evaluation Plan, outlining the study design; annual reports on project findings; and interim memoranda addressing key aspects of the study. This final summative report incorporates data from all five years of the project and has two primary components:

- an **impact evaluation** examining whether reading skills of students in Reading Partners improved more than those of similar students not served by the program. The impact analysis looks at student results over the 2013-14, 2014-15, and 2015-16 school years.
- an **implementation evaluation** of Reading Partners programming in Colorado over the five-year period of the SIF grant. Implementation evaluation findings are based on data from multiple sources, including Reading Partners student folder reviews, Reading Partners administrative data, and interview and survey data.

The study described in this report was designed and executed to achieve a **moderate** level of evidence within the SIF evidence framework. The report body is organized into four basic sections:

1. A brief overview of the Reading Partners program in Colorado;
2. A discussion of impact evaluation questions, approach, and findings, including information about the literacy context in comparison schools, to inform interpretation of impact findings;
3. A discussion of implementation evaluation questions, approach, and findings; and
4. Overall conclusions and next steps.

The impact and implementation components of the study are presented separately because they involve different research questions, school samples, and time periods. While the impact analysis reviews student data from the 2013-14, 2014-15, and 2015-16 school years, the implementation evaluation describes activities spanning the entire five-year funding period.

Reading Partners Overview

Founded in 1999, Reading Partners recruits, trains, and places community volunteers into high-need schools to provide tutoring for students who are behind in reading. Reading Partners collaborates with school leaders and teachers to transform a dedicated school space into a reading center, and creates twice weekly opportunities for students to receive tailored, one-on-one

---

[4] This report incorporates feedback from CNCS on earlier reports. Wherever possible, this report addresses specific recommendations from CNCS or from its evaluation consultant, JBS International.

instruction from a trained and supervised community volunteer. These volunteers use a structured, research-based curriculum adapted for each student. Community volunteers and students receive ongoing support from Reading Partners' Reading Center Coordinators, who provide volunteer training, observations, coaching, and progress monitoring to ensure that students are meeting their individual literacy goals. Reading Partners' vision is that one day all children will have the reading skills necessary for them to reach their full potential. Reading Partners has experienced rapid growth over the past decade, illustrated in Figure 1, and currently serves more than 11,000 students in over 200 schools across 10 states and the District of Columbia.

Figure 1: Reading Partners Has Grown Rapidly Since its Inception
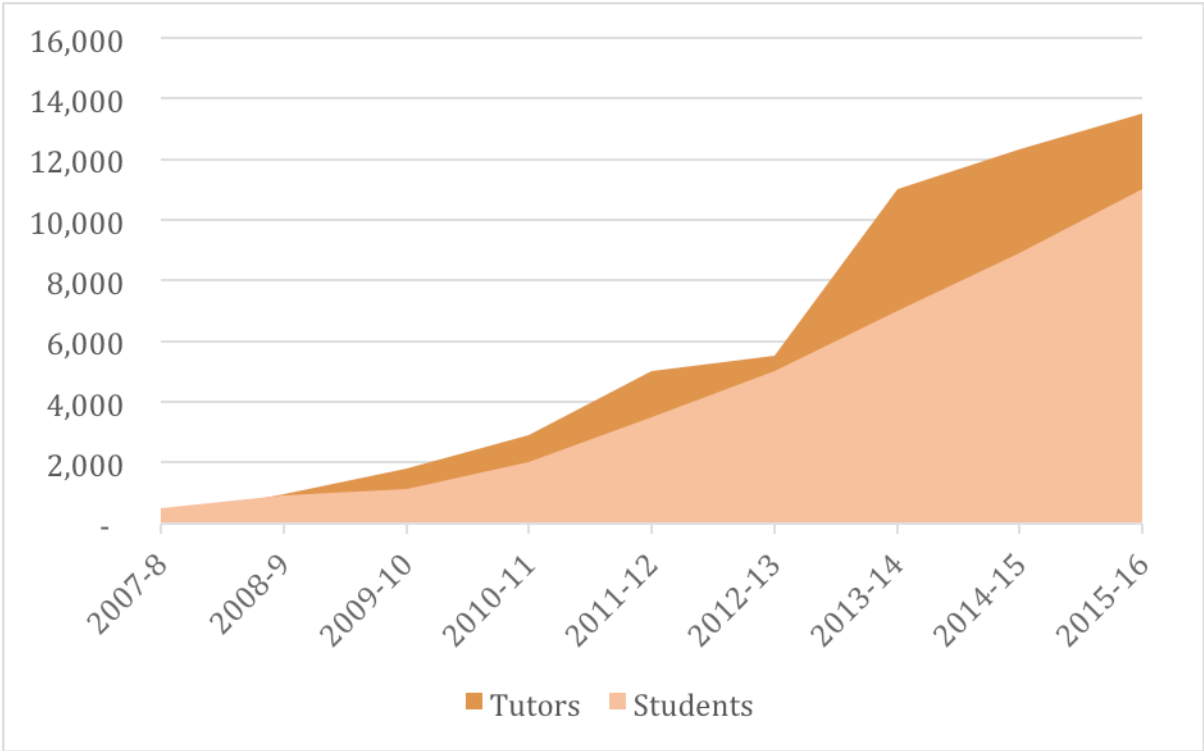
Reading Partners' Organizational Structure

Figure 2, below, provides an overview of the national and regional structure in the organization. The national staff includes Reading Partners Chief Executive Officer and Executive Team, as well as departments focused on program quality; AmeriCorps; community engagement; research and evaluation; marketing and communication; information technology; finance; and development.

Figure 2. Reading Partners Organization as of the 2016-17 School Year, by Establishment Year
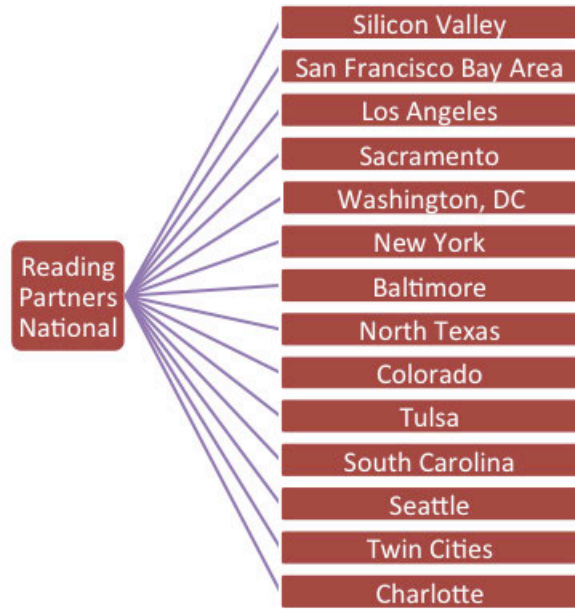
Each Reading Partners region is led by a regional Executive Director who oversees program delivery, community engagement, school partnerships, and development activities. While the specific structure varies somewhat by region, regional teams often include a Program Director, Program Managers, Program Associates, a Development Manager, and a Community Engagement Manager. Program Managers and Program Associates supervise and support Reading Partners school sites. Reading Partners Colorado's organizational structure has varied over the course of the project, but is illustrated in Figure 3.

Figure 3. Reading Partners Colorado Organization Chart (Illustrative)

The Reading Partners Program: Theory and Structure

Reading Partners' highly structured, closely supervised, volunteer-delivered program is designed to produce a set of meaningful, measurable benefits for students who participate. The Logic Model presented below (Figure 4) illustrates the needs and challenges confronting the students and communities Reading Partners serves; the key elements of Reading Partners' response to those needs and challenges; the intended benefits for students who participate in the program; and the desired impacts on the larger community.

Reading Partners Logic Model

Current Situation and Diagnosis of Need	Program Elements	Program Participation	RP Student Outcomes	Impacts on the Larger Community

Children in low-income communities read below proficiency

Schools lack sufficient resources to provide individualized support to students

Parents of these students do not have resources to help their children improve academically

Schools do not have the infrastructure to adequately build community support

---

Recruit and train caring adult volunteers to work with students needing reading support

Provide structured, curriculum-based reading support, with an individualized reading plan

Align RP curriculum with local curriculum and learning standards

Administer assessments and analyze state-sanctioned standardized tests to benchmark and track student progress

Work closely with classroom teachers to develop and modify each student's reading plan

Provide ongoing training and support to RP staff and volunteers

---

Recruit students in need of intensive intervention in reading

Conduct twice weekly tutoring sessions to provide opportunities for literacy instruction and development of reading skills

Use a one-on-one tutoring model to help students build close relationships with caring adults

---

Students gain literacy skills and improve their reading abilities

Students are better prepared to learn and demonstrate better academic behavior in the classroom.

Students gain self-confidence and develop a positive outlook on school.

Students make academic progress in other subject areas (e.g., math, science)

Students demonstrate better long-term outcomes than their non-RP peers on both academic and social indicators (e.g., lower rates of school dropout, teen pregnancy, and juvenile delinquency; higher rates of on-time grade promotion, high school graduation)

Schools experience broad improvement in academic achievement because every student is receiving an appropriate level of support

Schools recognize the value of providing rigorous individualized support to students

Parents and other community members develop greater awareness of literacy challenges and solutions

Volunteers promote greater community engagement in public education

The process model for Reading Partners in Figure 5 describes the key components of the Reading Partners program. Tutoring is the core program activity and it is supported by four components: a curriculum, trained volunteers who conduct the tutoring, in a school that provides space and access to students, and funding is necessary to support these activities. Through these four supports, the Reading Partners team works to meet the mission of helping children become lifelong readers by empowering communities to provide individualized instruction with measurable results. Each of these four supports are necessary for the implementation of the Reading Partners tutoring program.

Figure 5: Reading Partners Process Model
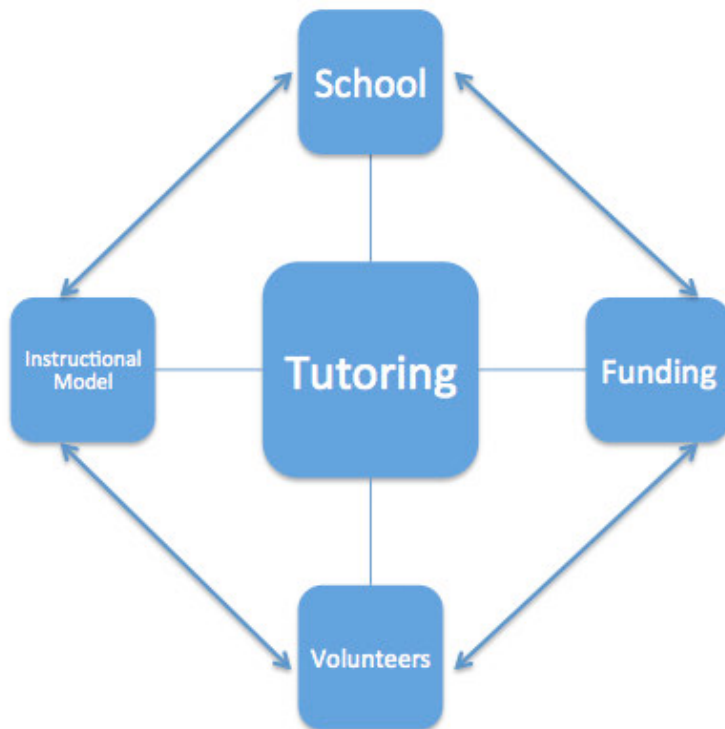
Tutoring

Reading Partners is a reading program that recruits, trains, and pairs volunteer tutors with struggling readers in elementary school reading centers. Each individual tutoring session is usually a 45-minute one-on-one experience between a volunteer tutor and a student. Reading Partners' implementation goal is that students receive two 45-minute tutoring sessions each week for a minimum of 20 sessions over the course of the school year. To the greatest extent possible, students are seen by the same tutor for each of their tutoring sessions to support development of strong relationships between tutors and students.

Schools

Schools play a crucial role in enabling Reading Partners' tutoring. Schools provide space where tutoring occurs, called a Reading Center. Schools also work with the Reading Center Coordinators to identify struggling readers for tutoring. Reading Center Coordinators work with school staff and teachers to identify times when students can be pulled out of their

regular classroom for tutoring. Schools or districts also provide funding to partially offset program costs.

Students

Students in need of intensive literacy support are identified by teachers and principals and recruited to the Reading Partners program. Reading Partners is targeted towards students who are one month to two and a half years behind grade level in reading skills, can speak conversational English, and who are typically not on an academic Individualized Education Plan (IEP).[5] Each student referred to Reading Partners is assessed by Reading Partners two or three times per year, depending on the student's enrollment date. Reading Partners personnel administer these assessments to students, who may also participate in school- or district-administered assessments. Reading Center Coordinators then use a student's initial, or "baseline," assessment results to develop an Instructional Reading Plan (IRP) for that student and to identify the appropriate curriculum unit for tutoring that student. The IRP is used throughout the year and outlines key skills and strategies for the volunteer tutor to work on with the student. The IRP is included in the student's folder with a running log of tutoring sessions. At the end of the fall semester, Reading Center Coordinators assess all students enrolled in tutoring and update student IRPs as needed.

Volunteers

The Reading Partners regional offices recruit community volunteers from many sectors, including high school or college students, retirees, and working adults. Students may volunteer individually or participate in Reading Partners as part of work-study or to fulfill community service expectations by classes, schools or colleges. Working adults may volunteer individually or as part of organized efforts by employers or volunteer organizations. Some tutoring is provided by Reading Partners staff (usually Coordinators).

Curriculum

Reading Partners uses a research-based curriculum. The program offers more than 120 different lesson plans across three curricula: Emerging Readers, Beginning Readers, and Comprehension Readers. Emerging Readers is for students just developing foundational reading skills and generally younger students. Beginning Readers is for students who are still perfecting their phonics and phonological awareness skills to read with accuracy and fluency. And, Comprehension Readers is for students who are working on applying many skills to facilitate comprehension of what they are reading.

Lessons are comprised of a combination of common activities that occur regardless of curriculum type, such as the tutor read aloud, and activities that are tailored for each curriculum level. Each curriculum consists of multiple, sequenced lessons. Students progress through each curriculum and can move to the next curriculum as they master new skills. Students can skip or repeat lessons based on their rate of progress and their lesson placement is re-evaluated by Reading Partners staff after a mid-year assessment.

---

[5] Reading Partners does serve some students with IEPs on a case-by-case basis, depending on conversations with teachers and other school leaders about whether Reading partners is an appropriate intervention for the student.

Unique Attributes of Program Implementation in Colorado

Program implementation in Colorado differs from other Reading Partners' regions in two notable ways. First, all Reading Partners Reading Centers are managed by Reading Center Coordinators. In all Reading Partners regions other than Colorado, the Reading Center Coordinator positions are held by AmeriCorps State or National members. However, the CNCS SIF grant to Reading Partners Colorado did not allow the region to engage AmeriCorps members in direct service roles, so Colorado's Reading Center Coordinators are all paid Reading Partners employees.[6]

Colorado was able to fill these positions with experienced educators with between one and twelve years of experience working in roles such as classroom teachers, pre-school teachers, and educators teaching abroad in higher education. Coordinator positions were staffed by experienced educators throughout the four years of APA's evaluation.

Second, while Reading Partners generally serves students in kindergarten through fifth grade, Colorado's implementation placed greater emphasis on serving students in the younger grades, and the impact study in Colorado focuses exclusively students in grades one through three, in response to the focus of SIF funding in Colorado on reading proficiency by grade three. Kindergarteners were excluded because literacy assessment data were either unavailable or not reliable for purposes of the impact analysis.

Overview of Prior Research

Reading Partners' model of volunteer-driven, pull-out, one-on-one literacy tutoring draws upon a research base indicating this type of program model has an impact on improving literacy skills. Reading Partners has also conducted an impact study that indicates positive impacts of the program.

The research base has several strong studies indicating the effectiveness of models similar to Reading Partners. For example, a 2000 meta-analysis of 29 studies published in the *Journal of Educational Psychology* found an average effect size of +.41 (Elbaum, Vaughn, Huges & Moody, 2000). The authors concluded that two sessions of one-on-one tutoring per week by a trained, supported, and supervised community volunteer can be an effective and affordable alternative to provide significant help to students at risk for reading failure. A later meta-analysis of seven studies of one-on-one tutoring, including four that were randomized, showed an average effect size of +.14 (Slavin, Lake, Davis & Madden, 2009). If studies of programs with fewer tutoring sessions or minimal structure are dropped from the analysis, the effect size increases to +.51. The authors conclude that:

> One-to-one tutoring is clearly very effective, and when resources are limited, well-structured programs making use of paraprofessionals and volunteers may reach more struggling readers for the same cost as serving many fewer children with certified teachers (Slavin, Lake, Davis & Madden, 2009, at p. 56).

---

[6] Reading Partners Colorado did engage AmeriCorps VISTA members, who focus specifically on organizational capacity building and not direct service, during the SIF grant. The CNCS website (www.natioanlservice.gov) provides additional information about the different types of AmeriCorps members serving in communities across the country.

These meta-analyses indicate that the basic model behind Reading Partners has evidence of effectiveness with a conservative expected effect size of +.15 to +.20. Because of this, APA will create our sample in this evaluation to allow a minimum detectable effect size (MDES) of .20.

A research team from Stanford University, led by Chiatovich, completed an impact evaluation in January 2012 that provides preliminary evidence of Reading Partners' effectiveness. The study used matched pairs that were randomly assigned to Reading Partners or a control group (total N=486) in grades one through five (Chiatovich, 2012). The study did not include questions of implementation process or quality. This study found a Reading Partners effect size of +.51 on the Rigby PM Ultra Benchmark assessment administered by Reading Partners, but no effect when assessing student learning with a California state assessment that combined reading and writing scores. Given these mixed results, the program believes this study provides a preliminary level of program effectiveness.

Other findings from the Chiatovich study, based on the analysis of the Rigby assessment, include that the impact of Reading Partners did not vary significantly according to students' gender, grade, year of study participation, race or ethnicity, English language learner level, or whether children were above the target range of beginning reading ability. In other words, the impact of Reading Partners was consistent across grade level and student status. Nor did the effect vary by school. The Stanford team suggested this lack of variation by school indicates consistent implementation; however their study did not examine program implementation directly.

Finally, and most recently, MDRC conducted a rigorous implementation and impact evaluation during the 2012-13 school year (Jacob, Armstrong & Willard, 2015). That evaluation examined program impact, cost and implementation in 19 Reading Centers, in 12 school districts, and 6 Reading Partners regions (New York City; Washington, DC; and, in California, Sacramento, San Francisco, Silicon Valley, and Los Angeles). That evaluation found Reading Partners was implemented with fidelity. In March 2015, MDRC reported results from the impact evaluation, which analyzed changes in reading scores for second through fifth graders, using a randomized control design. That study, which met requirements for the "strong" evidence level within the SIF evidence framework, found that participation in Reading Partners had a positive impact on student reading scores, with effect size impacts between 0.09 and 0.11, depending on the reading skill being measured. These results compare favorably with other volunteer tutoring programs. The MDRC report also found that session scheduling challenges made it difficult for Reading Partners to meet its goal of providing students with two tutoring sessions with a single tutor each week. Nonetheless, MDRC found that most students received at least three tutoring sessions over a two-week period, though often with multiple tutors. In addition, the cost analysis in that evaluation Reading Partners to be substantially less costly than other supplemental reading services typically offered to struggling readers.

Impact Analysis: Student Outcomes in 2013-14, 2014-15 and 2015-16

Introduction

The impact evaluation examines whether the reading skills of students in Reading Partners improved more than those of similar students not served by the program, as measured by state-mandated, school-based literacy assessments. This impact evaluation sought to answer <u>four research questions</u>:

1. Does Reading Partners tutoring lead to improved near-term reading achievement for students in grades one through three when compared to similar students who do not receive tutoring?

2. Do differences in reading achievement between students who receive Reading Partners tutoring and similar students who are not in Reading Partners increase as students receive more tutoring?

3. Are there differential impacts of Reading Partners tutoring on different student groups, including English-language learners (ELL) vs. non-ELL students, boys vs girls, grade level, and different races?

4. What is the effect of participating in Reading Partners for multiple years?

This final year of the impact analysis incorporates three years of data and leverages processes and procedures established by the APA research team during the first two years of the impact analysis, focused on establishing a moderate level of evidence.[7] In particular, APA successfully completed the following activities:

- Exchanging and managing sensitive data files with participating districts;
- Matching students to identify a balanced treatment and control sample;
- Normalizing student assessment data; and
- Applying statistical models to analyze the data.

The final analysis presented in this report combines student data from the 2013-14 and 2014-15 school years with and student data from the 2015-16 school year, resulting in a pooled data sample with additional statistical power.[8] The following report sections discuss the process that APA followed in preparing the data for analysis, including: a description of the data sources used, particularly the literacy assessment data that serve as the study outcome measure; the steps taken to normalize outcome data in preparation for analysis; the procedures used to construct the student comparison group; and the results of the Propensity Score Matching process that yielded the final dataset for analysis.[9]

---

[7] See Appendix A for more information about the targeted level of evidence.

[8] Data from 2013-14 and 2014-15 were previously analyzed and reported in the Year 3 and Year 4 reports.

[9] Reading Partners operated in both Aurora Public Schools (APS) and Denver Public Schools (DPS) in 2013-14 and 2014-15. In 2015-16, Reading Partners operated only in DPS. However, APA has been unable to obtain complete student data from the 2014-15 school year from APS. APA has a data agreement in place with APS. APA did receive a partial dataset from APS in April 2016, but it did not

Student Impact Sample:
The complete pooled sample contained 698 Reading Partners students and 853 similar comparison students. The 2015-16 student sample had 399 Reading Partners students and 399 comparison students.

The total pooled sample of 2013-14, 2014-15, and 2015-16 students contained 698 Reading Partners students and 853 similar comparison students. The pooled sample is smaller than the sum of the samples in each year because some Reading Partner students were served in two or three of the years but were only included in the pooled sample once. This pooled sample includes the balanced sample of 2015-16 students, with 399 Reading Partners students and 399 similar comparison students. These 2015-16 students were then pooled with the 2013-14 and 2014-15 students, as described in Table 1, below. As described further below, students who participated in Reading Partners in both years were included in the analytic model for the primary research question on program effect only during their initial year of program participation, because that question focused only on the impact of participating in Reading Partners during a single year.

Table 1: Sample sizes for the impact analysis

	2013-14	2014-15	2015-16	Total[10]
Reading Partners students	233	291	399	923
Comparison students	233	291	399	923
Total	466	582	798	1,846

In brief, APA found statistically significant positive results in three analysis models using the three-year pooled sample. First, APA found that, on average, students who participate in Reading Partners during one year have spring assessment scores that are significantly higher than the scores of similar comparison students who do not participate in Reading Partners. Second, when comparing Reading Partners students to all comparison students, there is a significant positive effect of each additional tutoring session. Finally, Reading Partners students designated

Overall Finding: On average, students who participate in Reading Partners for one year have spring assessment scores that are significantly higher than the scores of similar comparison students who do not participate in Reading Partners.

---

include fall assessment scores for students, which meant APA did not have pre-test data for APS students so could not perform propensity score matching with the APS students. This meant that no APS data could be included in the analysis for 2014-15. Despite continued contacts with APS since then, APA still has been unable to obtain a complete dataset. Therefore, the data for APS students in 2014-15 has not been included in this impact analysis. The pooled data includes APS students from 2013-14 and DPS students from all three years. While APA would prefer to have APS data from 2014-15, the 2013-14 analysis indicated that district was not a significant predictor of student performance, indicating that there are no significant differences between APS and DPS students.

[10] These numbers differ from the sample sizes reported in the report text because students who attended the program for multiple years are included in each yearly count. The numbers in the report text count the number of total unique participants over the study period.

as English language learners (ELL) grew significantly more than non-ELL Reading Partners students. These findings are discussed in detail below.

Data Sources and Methods

This final impact evaluation describes the impact for the pooled group of students who participated in Reading Partners during the 2013-14, 2014-15, and 2015-16 school years. As described above, while Reading Partners generally serves students in kindergarten through fifth grade, this study only includes students in grades one through three because the SIF funding in Colorado focuses on reading proficiency by grade three and literacy assessment data for kindergarteners were either unavailable or not reliable for purposes of the impact analysis.

Data Sources

The impact evaluation data for this analysis came from two sources:[11]

- First, for both treatment and comparison students, participating school districts provided information on student literacy assessment scores, demographics, participation in select school programs, and grade levels. Demographic data described student gender, race, and ethnicity.[12] The school program participation information included whether a child had an IEP, was an English Language Learner (ELL), and/or was eligible for the federal free or reduced-price lunch program (FRL).

- For treatment students only, Reading Partners provided information on program participation, including the number of Reading Partners tutoring sessions (dosage) received in the year(s) during which the student participated in the program.

> **Data Sources:** APA gathered demographic and assessment score information on both Reading Partners students and comparison group students from participating school districts. Reading Partners provided information on Reading Partners students' participation in the program. For this final impact analysis, APA pooled data on students from the 2013-14, 2014-15, and 2015-16 school years.

Schools administer literacy assessments to elementary grade students multiple times each year, including benchmark administrations in the fall and spring, as part of their compliance with Colorado's Reading to Ensure Academic Development Act (Colorado READ Act, or CO HB 12-1238). These assessments act as tools to help measure student reading progress across school districts and serve as the primary outcome measure for impact analysis. However, the READ Act allows schools and districts to select among a menu of seven approved literacy assessments. As a result, the specific assessments administered varied across schools, districts, and school years:

---

[11] Please see Appendix B for more information about the administrative data process.
[12] The race and ethnicity data fields included white, black, Latino, Asian, and multiple race or ethnicities.

- In 2013-14, DPS and APS both assessed literacy using the Diagnostic Reading Assessment, Second Edition (DRA2).
- In 2014-15, DPS allowed district elementary schools to choose literacy assessments from a list of seven approved possibilities. While some non-Reading Partners schools opted to discontinue use of the DRA2, the Reading Partners schools and comparison schools continued to use only the DRA2 assessment.
- In 2015-16, DPS again allowed district elementary schools to choose literacy assessments from a list of approved possibilities. In 2015-16, students in Reading Partners schools took three different assessments: The DRA2, the iStation, and the DIBELS Next assessments.[13] As in previous years, schools administered these assessments multiple times a year as part of their responsibilities under the Colorado READ Act, including administrations in the fall and spring.

The final impact analysis includes two literacy outcomes measures due to changes in assessment practices within partnering school districts:[14]

- First is the DRA2, published by Pearson. Schools currently administer the DRA2 in the fall and spring in compliance with the Colorado READ Act. The DRA2 is designed so teachers can use the assessment data to "make sound decisions about student reading levels and development, and [to] inform subsequent instruction" (Pearson, 2009). The DRA2 is a criterion-referenced test that has multiple scales and forms, with testing procedures that vary depending on student age and development range (McCarthy & Christ, 2010). Although the DRA2 has a number of subcomponents, the DRA2 scores provided by DPS and used for this analysis were the composite scores (Independent Reading Level). These scores were on an ordinal scale.
- Second is the iStation Early Reading Assessment. This assessment is designed to provide information on the literacy skills for students in pre-K through 3rd grade. The iStation assessment is a computer adaptive assessment that is administered on a computer. After each question, the assessment selects a following question for the student based on their responses to all the questions so far, tailoring the specific questions to the student's performance. The iStation assessment produces a scale score for each child, which is on an ordinal scale.

Both assessments have been found to possess high internal consistency and test-retest reliability (DRA2 Technical Manual, 2009; Istation's Indicators of Progress Early Reading Technical Report, Version 4).

---

[13] As discussed later in the report, so few Reading Partners students took the DIBELS Next assessment that they were excluded from the analysis rather than introducing another outcome assessment into the analysis.

[14] As discussed below, two Reading Partners students took the DIBELS Next assessment. Those two students were excluded from this analysis because of the difficulty of selecting appropriate comparison students for such a small sample and to avoid introducing another outcome measure for only two students.

Methods

Assessment Data Normalization

Combining data across school years led to a larger overall sample size but also expanded the number of different student assessments used as outcome measures in the impact analysis. Both because districts used multiple different assessments, as described above, and because the assessment score results are on a scale that is more ordinal than interval, the APA research team transformed all student assessment scores to a normal curve equivalent (NCE), which describes student performance on a 1-100 interval scale, similar to a percentile scale. APA began this process of standardization and transformation in Year 3 with the 2013-14 data and continued this process with the 2014-15 data, so also standardized and transformed the 2015-16 assessment scores before pooling them with data from the previous two years. Throughout this report, significant results have been translated back into percentile units for ease of interpretation.

> **Measuring Impact:** The impact evaluation used student scores on the DRA2 and iStation assessments administered by school personnel. To use these scores in analytic models, APA standardized the scores by school district and by grade level by transforming raw scores into percentiles and Normal Curve Equivalents (NCEs).

For assessment data from the 2013-14 and 2014-15 school years, APA normalized the DRA2 assessment scores by grade level and district. This means that the scores are normalized within the local sample, rather than representing a national normalization. Because the initial scores were not normally distributed, a z-score normalization was inappropriate. Instead, APA performed a percentile transformation on the scores, calculating the percentile rank of each student's score within their grade level. [15] The percentile scores were then transformed into NCE scores for inclusion in the analytic regression models.

When adding the 2015-16 data to the pooled 2013-14 and 2014-15 data, APA again normalized the raw assessment scores into percentiles and then transformed the percentile scores into NCEs. APA normalized the DRA2 and iStation scores by grade level. Because all of the students in the 2015-16 data were from DPS, APA did not normalize by school district. Normalization of assessment scores facilitated the inclusion of multiple assessments in the analytic model, allowing comparisons across the multiple years of data, regardless of the specific assessment administered in a school.

> **Normalizing Literacy Assessment Data:** The research team needed to normalize student assessment data to allow for inclusion of multiple different assessment measures. Normalization of assessment scores yields data that can be compared across grade-levels and across multiple different assessments. These data are normalized within the local sample, not nationally.

---

[15] The analysis of the 2013-14 data normalized student assessment scores by grade level and by school district, so percentile ranks were assigned for the group of all APS first graders, then the group of all APS second graders, and so on. In year 4, because all students in the analysis were from DPS, scores were simply normalized by grade level.

Again, the distributions of the DRA2 and iStation test scores in individual grades before normalization were not normally distributed, making a z-score normalization inappropriate. Instead, APA performed a percentile transformation on the DRA2 and iStation scores, calculating the percentile rank of each student's score within their grade level.

Although percentile transformations were the appropriate normalization technique given that the raw score data were ordinal, percentiles themselves are ordinal and therefore not appropriate for inclusion in the analytic models. Because percentiles are ordinal, the distance between percentile scores is not constant across the range of percentiles, making it is easier for a student to make a test score gain from the 50th to the 51st percentile than to move from the 90th to the 91st percentile. This property makes percentile data inappropriate for inclusion in the impact analysis models.

Because the linear models APA employed for analysis required input data that possessed the interval property, APA performed an additional transformation from percentiles to NCE scores. The NCE scores are similar to the percentiles but are a transformation specifically designed to impose the interval property on ordinal data. An NCE score is identical to a percentile at the first, 50th, and 99th percentiles.

Selecting Comparison Students

The impact analysis compares Reading Partners students to similar students who received "business as usual" reading supports. The comparison group includes students who either attended schools served by Reading Partners but did not participate in the program or attended similar schools not served by Reading Partners. To construct the final student sample for this study, APA took three key steps: (1) identified schools for inclusion in the sample; (2) applied the PSM process on student data from the 2015-16 school year to identify comparison students from those schools; and then (3) combined the PSM matched sample for 2015-16 with the pooled PSM matched students from the 2013-14 and 2014-15 school years. This section of the report describes selection of the comparison schools; a detailed description of the PSM process appears in the Student Sample Selection section below.

All treatment students were served by Reading Partners. As noted in the text box above, comparison students for the sample were drawn either from the treatment schools, including students at these schools that were not served by Reading Partners, or from a separate set of identified comparison schools.

Comparison Students: Comparison students for the study were drawn either (1) from schools with Reading Partners sites (where comparison students included only students who were *not* served by Reading Partners) or (2) from a separate set of identified comparison schools. APA selected 30 comparison schools that were similar to Reading Partners schools in terms of geographic location, racial/ethnic makeup of students, poverty rate among students and assessment administered.

Using the same procedure followed to identify comparison schools in the 2013-14 and 2014-15 school years, APA identified comparison schools for 2015-16 that were sufficiently similar to the schools served by Reading Partners on the following characteristics: school

district, geographic proximity to one or more treatment schools, use of traditional academic programs (Montessori or international academies were not included), school enrollment, and similar school-level student demographics, where demographic comparisons focused on measures of FRL eligibility and percentages of Latino and black students.[16] APA's reasoning was that schools with similar student composition are more likely to have similar reading programs. For the 2014-15 school year, APA collected data from a sample of comparison schools that suggested the reading programs were generally similar in these schools, but may have had somewhat longer literacy blocks. The APA research team identified multiple comparison schools per treatment school to provide a large pool of potential students, thereby increasing quality of the match and ensuring that matched students were as similar as possible to Reading Partners students. For the 2015-16 school year, APA also identified comparison schools based on the assessment they administered, to ensure there were sufficient students from comparison schools who had taken either the iStation or the DRA2 assessment for an adequate match.

Table 2, below, lists the 15 treatment and 30 comparison schools included in the sample for 2015-16.[17]

Table 2. Sample of Treatment and Comparison Schools in the Impact Analysis, 2015-16

School	Group	School Enrollment	FRL	Black	Latino	White
Amesse Elementary	Comparison	488	95.9%	18.8%	71.5%	4.1%
Archuleta Elementary	Comparison	564	87.9%	13.3%	75.4%	4.6%
Barnum Elementary	Comparison	457	98.2%	2.5%	90.8%	6.3%
Columbian Elementary	Comparison	240	91.7%	0.7%	87.3%	8.9%
Castro Elementary	Comparison	528	97.5%	4.4%	88.8%	1.7%
Doull Elementary	Comparison	445	89.9%	0.8%	86.9%	7.2%
Ellis Elementary	Comparison	429	86.5%	8.8%	55.6%	28.2%
Fariview Elementary	Comparison	203	99.5%	39.2%	41.0%	7.0%
Godsman Elementary	Comparison	517	93.2%	0.2%	91.1%	5.6%
Goldrick Elementary	Comparison	530	97.4%	3.8%	88.0%	2.7%
Green Valley Elementary	Comparison	657	68.6%	26.9%	53.6%	9.8%

---

[16] Schools were considered to be sufficiently similar to treatment schools if they were within 10 percentage points of the treatment school on all the matching variables (FRL eligibility and percentages Black and Latino students). Please see Appendix C for a complete list of comparison schools over the three year period.

[17] Comparison schools for each year of the impact evaluation are listed in Appendix C. Detailed information about the characteristics of comparison schools used in the 2013-14 school year and the 2014-15 school year can be found in the Year 3 and Year 4 Reading Partners Reports, respectively.

School	Group	School Enrollment	FRL	Black	Latino	White
Greenlee Elementary	Comparison	297	93.3%	26.0%	59.8%	8.1%
Greenwood Elementary	Comparison	636	93.2%	9.7%	85.2%	1.3%
Harrington Elementary	Comparison	324	97.2%	19.1%	74.9%	4.1%
Holm Elementary	Comparison	484	82.0%	19.1%	53.1%	16.6%
Johnson Elementary	Comparison	387	94.8%	1.7%	85.2%	6.0%
Marrama Elementary	Comparison	535	82.6%	27.6%	53.0%	8.1%
Maxwell Elementary	Comparison	486	92.4%	16.3%	77.1%	3.1%
McGlone Elementary	Comparison	666	94.6%	14.8%	78.6%	2.7%
McMeen Elementary	Comparison	613	79.9%	26.6%	37.5%	24.4%
Montclair Elementary	Comparison	429	73.0%	36.1%	32.3%	26.0%
Munroe Elementary	Comparison	542	97.6%	2.3%	92.0%	1.7%
Newlon Elementary	Comparison	524	92.2%	1.1%	93.0%	3.6%
Oakland Elementary	Comparison	402	91.5%	25.3%	66.7%	4.0%
Palmer Elementary	Comparison	236	48.7%	27.1%	23.1%	40.3%
Schmitt Elementary	Comparison	360	95.8%	9.2%	76.4%	4.7%
Stedman Elementary	Comparison	239	89.1%	56.9%	30.2%	7.5%
Swansea Elementary	Comparison	472	96.2%	3.6%	93.2%	2.3%
University Park Elementary	Comparison	423	25.1%	3.3%	13.5%	70.9%
University Prep Elementary	Comparison	363	81.8%	38.0%	47.9%	8.3%
Wyatt Elementary	Comparison	513	94.9%	18.5%	77.0%	1.6%
Ashley Elementary	Treatment	358	89.7%	25.9%	59.4%	9.4%
Cheltenham Elementary	Treatment	381	99.0%	12.1%	78.4%	4.1%
Cole Elementary	Treatment	468	93.2%	19.3%	71.4%	5.4%
Colfax Elementary	Treatment	289	94.8%	11.0%	76.5%	7.6%
College View Elementary	Treatment	396	96.5%	5.3%	78.7%	5.1%
Columbine Elementary	Treatment	198	89.9%	45.8%	41.3%	8.3%
Dora Moore Elementary	Treatment	360	75.8%	21.9%	38.9%	27.2%

School	Group	School Enrollment	FRL	Black	Latino	White
Force Elementary	Treatment	529	94.9%	1.4%	89.9%	5.1%
Garden Place Elementary	Treatment	336	94.0%	3.1%	91.3%	3.8%
Kaiser Elementary	Treatment	315	75.6%	5.0%	59.1%	29.7%
Samuels Elementary	Treatment	454	78.2%	20.4%	47.9%	19.8%
Smith Elementary	Treatment	335	94.9%	29.8%	63.3%	2.8%
TreVista Elementary	Treatment	320	97.2%	15.3%	77.8%	4.7%
Valdez Elementary	Treatment	369	55.8%	0.8%	62.3%	34.7%

Data source: Colorado Department of Education, Reading Partners Colorado

In 2013-14, APA gathered information on "business as usual" literacy practices by interviewing literacy leaders in the schools. In 2014-15 and 2015-16, APA surveyed literacy leaders at Reading Partners schools and comparison schools to determine the "business as usual" supports that students received during the school year.[18] The literacy leaders consistently described using the response to intervention (RtI) framework for their students. The RtI framework has three tiers that support different school supports to meet student needs. Tier 1 is the universal level that describes the instruction and supports that the large majority of students (i.e. general population students) receive. Tier 2 is the intervention level for students who are behind or not making expected progress. In many treatment schools, the Reading Partners program was described as the main Tier 2 intervention. Tier 3 is the intensive intervention level, generally focused on students with IEPs.

School staff at Reading Partners schools described Tier 2 students, or "cusp" students, as prime candidates for enrollment in Reading Partners. Reading Partners provided the primary support for Tier 2 students in the treatment schools, and based on interviews with literacy leaders in Reading Partners schools, APA believes that few Reading Partners students accessed additional Tier 2 supports because literacy leaders in schools served by Reading Partners reported that the school used Reading Partners as their Tier 2 support for eligible students.

The literacy block period at Reading Partners schools ranged from 90 to 150 minutes. Over the three years of data collection, in the majority of schools working with Reading Partners, students consistently received tutoring services outside of the literacy block period. Only two Reading Partners schools had students receive Reading Partners tutoring during the literacy block period. There was more variation in whether Reading Partners schools also

---

[18] Because comparison schools were not participating in the study, APA often had difficulty obtaining information from literacy leaders at those schools. APA was able to interview literacy leaders at 13 of the 21 comparison schools in 2014-15 (62%). In 2015-16, APA was able to interview literacy leaders at 9 of the 30 comparison schools (30%). In both years, APA was able to interview literacy leaders at all schools participating in Reading Partners.

worked with other outside reading programs, such as Reading Corps. Over multiple years, about half of the schools also worked with another outside reading program.

There was similar variation in the "business as usual" operations at comparison schools over the three-year period. These schools incorporated a literacy block between 90 and 180 minutes in duration. In some cases, students who were behind (Tier 2) received additional reading supports beyond the reading block. In most cases, these reading supports occurred in small groups or within the classroom. Comparison schools not served by Reading Partners generally did not offer a supplemental reading program commensurate to Reading Partners, with individualized 1 on 1 tutoring instruction.

Within APA's sample, the students in both treatment and comparison groups generally received similar supports. The largest difference was the disparity in the length of the literacy blocks. Literacy blocks in comparison schools were roughly 150 minutes, while the average in treatment schools was 120 minutes. Furthermore, two treatment schools have uniform literacy models, whereas no comparison schools have a uniform literacy model. Besides these minor differences, supports were very similar across treatment and comparison schools.

Student Sample Construction

As in prior years, APA used one-to-one Mahalanobis Propensity Score Matching (PSM), without replacement to construct the sample for analysis.[19] This matching technique takes individual Reading Partners students and matches them with a comparison student who is similar in terms of demographics and baseline assessment score. Each comparison student can only be matched with one treatment student, so this matching technique generates treatment and control groups of equal size. Using this technique, the APA research team created a balanced sample, as shown in Tables 6 and 7 and discussed in greater detail below.

> **Student Sample:** To construct the student sample for the impact analysis, APA matched Reading Partners students and comparison students who had both fall and spring DRA2 or iStation scores. Students were matched on their assessment pre-scores and demographic characteristics. The final matched sample was very well-balanced and the two groups of students were comparable.

PSM is a statistical technique for identifying comparison students with similar characteristics as students who are in the treatment group. Propensity scores are estimated using logistic regression in which the treatment assignment is the outcome variable and the covariates related to reading scores are used as predictor variables. To perform the PSM, APA gathered student data from partner school districts on variables believed to predict participation in Reading Partners. Specifically, DPS provided a set of student-level variables describing the demographic characteristics of students, their school location, and their academic performance at the beginning of the school year. APA used all of the available variables for the PSM.

A logistic regression predicting Reading Partners participation describes the relationships between each student-level variable and the probability of participating in Reading

---

[19] See Appendix D for technical details of the propensity score match, including the distribution of propensity scores across comparison and treatment groups.

Partners, controlling for the other student-level variables. The model for this prediction is described below:

$$Y_i = \beta_0 + \beta_{fall\ DRA(i)} + \beta_{FRL(i)} + \beta_{ELL(i)} + \beta_{SpEd(i)} - \beta_{male(i)} + \beta_{Latino(i)} + \beta_{Black(i)} - \beta_{Asian(i)} + \beta_{other(i)} + e_i$$

The propensity model predicts whether a student (i) is in the treatment group (Y_i) as predicted by their Fall assessment test score prior to treatment (normalized to NCE scores), and a set of indicator variables describing whether the student qualified for FRL, qualified for special education (SpEd), or was identified by the district as an ELL, male, Latino, black, Asian, or from another race (with white being the reference variable). Some of these variables were selected for inclusion in the propensity score match because Reading Partners policies mean they are associated with treatment group membership, such as the requirements that participating students be fluent in conversational English. Other variables were selected because literature indicates they are predictive of academic achievement and need for supplemental literacy services.

When preparing the 2015-16 data, APA ran the PSM model separately for each grade to reduce bias due to grade-specific assessment scales. APA also ran the PSM model separately for each assessment – so treatment students who took the iStation could only be matched with comparison students who also took the iStation, and treatment students who took the DRA2 could only be matched with comparison students who took the DRA2. This within-assessment matching controlled for any variations or bias due to assessment differences. Additionally, potential match students were drawn only from schools implementing Reading Partners or comparison schools, as described above.

In providing assessment scores, DPS provided scores for the single test administration of the DRA2 and the iStation that occurred in September. DPS also provided spring test scores for administration of the iStation in April and May, and scores from the May DRA administration.

Of the 498 Reading Partners students who had at least one DRA2 score in 2014-15, 333 were used in the analysis. The 165 excluded students were missing either fall or spring assessment scores, which made it impossible to perform a PSM for them or to include them in the analysis model. Because the Spanish version of the DRA2, the EDL, is very similar to the DRA2 and measures the same underlying constructs, APA used EDL scores for students who had no DRA2 scores available. Fortunately, the group of Reading Partners students with both DRA2 assessment scores was demographically very similar to the group of Reading Partners students who took the DRA2 assessment as a whole, as shown in Table 2.[20]

Table 2 provides the means and standard deviations for the group of Reading Partners first through third grade students who took the DRA2 assessment and could be included in the sample used in the PSM because they had data for both test scores and all demographic variables, compared with all first through third grade students served by Reading Partners who took the DRA2 in 2015-16. The table also includes the standardized mean difference, or

---

[20] Please see the Year 3 and Year 4 reports for an analysis of the similarity of treatment and comparison students for the 2013-14 and 2014-15 data.

the difference in the demographics in the two groups divided by the standard deviation for the whole DRA2-taking Reading Partners group. A standardized mean difference of 0.05 indicates that the mean for Reading Partners students with both DRA2 test scores was 0.05 standard deviations larger than the mean of all Reading Partners DRA2 students. When using the standardized mean difference to identify similar groups, the absolute magnitude should be no larger than 0.25 and preferably less than 0.1. All of the standardized mean differences met the .1 threshold, indicating that the sample with Reading Partners students with both DRA2 test scores was demographically comparable to the group of all DRA2 Reading Partners students.

Table 3. Description of the 2015-16 Reading Partners Analysis Sample who Took the DRA2 Assessment

Demographic Variable	All Reading Partners Students in Grades 1-3		Reading Partners Students in Grades 1-3 with Only One Test Score		Absolute Standardized Mean Difference
	Average	Standard deviation	Average	Standard deviation	
FRL	90.2%	29.8%	92.1%	27.0%	0.066
ELL	40.2%	49.1%	37.6%	48.6%	0.053
Special Education	8.4%	27.8%	8.5%	28.0%	0.002
Male	51.4%	50.0%	49.7%	50.2%	0.034
Asian	2.8%	16.5%	1.8%	13.4%	0.060
Other Race	6.2%	24.2%	7.3%	26.0%	0.043
Latino	64.1%	48.0%	63.0%	48.4%	0.021
Black	16.7%	37.3%	19.4%	39.7%	0.073
Count (n)	498		165		

Data source: APA analysis of DPS and Reading Partners data

APA was able to create complete iStation records with a September and May administration for 91 of the 204 Reading Partners students who took the iStation. Thirteen Reading Partners students took the iStation Español, a Spanish-language assessment by the creators of iStation. However, the iStation Español measures different underlying constructs than the iStation literacy assessment. DPS does not consider the iStation Español to be an equivalent assessment to the English-language iStation assessment included here. Thus, these thirteen Reading Partners students who took only the iStation Español were excluded from the analysis. The 113 Reading Partners students who took the iStation either in fall or spring but who did not have a complete test record were also excluded from the analysis. As with students excluded from the DRA2 analysis, the Reading Partners students who took the iStation but were excluded due to not having a complete test record were very similar to the complete pool of Reading Partners students who took the iStation.

Table 4, below, shows the demographic characteristics of the Reading Partners students who took the iStation as a whole, compared to those who were excluded from the analysis for having only one iStation assessment. Again, all the standardized mean differences are below the 0.1 threshold, indicating that the two groups are very similar.

Table 4. Description of the 2015-16 Reading Partners Analysis Sample who Took the iStation Assessment

Demographic Variable	All Reading Partners Students in Grades 1-3		Reading Partners Students in Grades 1-3 with Only One Test Score[21]		Absolute Standardized Mean Difference
	Average	Standard deviation	Average	Standard deviation	
FRL	91.2%	28.4%	92.0%	27.2%	0.030
ELL	33.8%	47.4%	37.2%	48.5%	0.071
Special Education	11.8%	32.3%	9.7%	29.8%	0.063
Male	56.4%	49.7%	58.4%	49.5%	0.041
Asian	2.0%	13.9%	2.7%	16.1%	0.050
Other Race	3.4%	18.2%	3.5%	18.6%	0.006
Latino	66.2%	47.4%	64.6%	48.0%	0.033
Black	14.2%	35.0%	15.9%	36.8%	0.049
Count (n)	204		113		

Data source: APA analysis of DPS and Reading Partners data

APA performed the PSM separately for each grade within each district. An appropriate match was found for each treatment student, meaning no cases needed to be excluded for failure to find a match. After performing the PSM, APA pooled the data from 2015-16 students for all grade levels and both districts. Table 5 describes the size of the 2015-16 samples, by grade and overall.

---

[21] The thirteen Reading Partners students who took the iStation Spanish are not included in this table, only Reading Partners students who were excluded because they did not have a complete test record.

Table 5. Description of the 2015-16 Treatment and Comparison Group Samples by Grade and Assessment

	Reading Partners iStation	Comparison iStation	Reading Partners DRA2	Comparison DRA2	Total Reading Partners	Total Comparison
Grade 1	27	27	103	103	130	130
Grade 2	20	20	133	133	153	153
Grade 3	34	34	97	97	131	131
Total	91	91	333	333	414	414

Data source: APA analysis of DPS and Reading Partners data

As noted previously, only two Reading Partners students in the grade range examined by this study, first through third grade, had complete test records for the DIBELS Next assessment, another assessment that DPS elementary schools could use to meet their READ Act requirements. Because of the difficulties of making and evaluating a propensity score match with a sample that small, and the additional complexity from adding another assessment to the outcome literacy measure, APA excluded these two Reading Partners students from the analysis. Three students had complete test record for IDEL assessments, the Spanish version of the DIBELS Next assessment. However, the IDEL reports only proficiency bands and does not generate a scale score, so those assessment outcomes could not be included in the analysis.

To test the quality of the match between treatment and comparison students, APA calculated the standardized mean difference of each demographic variable for the two groups. As discussed above, when reviewing the quality of a match, the absolute magnitude of the standardized mean difference should be no larger than 0.25 and preferably should be less than 0.1.

APA then pooled the matched data from 2015-16 with the already pooled data from the 2013-14 and 2014-15 years. The following two tables describe the overall pooled samples including data for all three years. Tables 6 and 7 detail the characteristics of the pooled treatment and pooled comparison groups generated by the PSM and includes all variables used in the match; Table 6 presents data for students assessed with iStation and Table 7 presents data for students assessed with DRA2.

Table 6. Characteristics of the Pooled (3-year) Treatment and Comparison Samples for iStation Students

Demographic Variable	Comparison Students		Reading Partners Students		Absolute Standardized Mean Difference
	Average	Standard Deviation	Average	Standard Deviation	
FRL	85.7%	35.2%	90.1%	30.0%	0.13
ELL	30.8%	46.4%	29.7%	45.9%	0.02
Special Education	12.1%	32.8%	14.3%	35.2%	0.06
Male	52.7%	50.2%	53.8%	50.1%	0.02
Asian	0.0%	0.0%	1.1%	10.5%	0.15
Other Race	1.1%	10.5%	3.3%	18.0%	0.15
Latino	69.2%	46.4%	68.1%	46.9%	0.02
Black	15.4%	36.3%	12.1%	32.8%	0.10
Fall iStation	202.5	18.2	201.6	17.8	0.05
Count (n)	91		91		

Data source: APA analysis of DPS and Reading Partners data

While the comparison and Reading Partners groups have absolute standardized mean differences greater than 0.1, such as difference in FRL receipt, Asian race status, and other race status, those absolute standardized mean differences still fall below 0.25, so indicate an acceptable match.

Table 7. Characteristics of the Pooled (3-year) Treatment and Comparison Samples for DRA2 Students

Demographic Variable	Comparison Students		Reading Partners Students		Absolute Standardized Mean Difference
	Average	Standard Deviation	Average	Standard Deviation	
FRL	88.7%	37.1%	90.4%	29.5%	0.06
ELL	45.0%	49.8%	46.6%	49.9%	0.03
Special Education	16.6%	37.3%	11.9%	32.4%	0.13
Male	51.0%	50.0%	53.1%	49.9%	0.04
Asian	5.7%	23.3%	3.4%	18.0%	0.11
Other Race	4.3%	20.2%	4.3%	20.3%	0.00
Latino	65.4%	47.6%	70.6%	45.6%	0.11
Black	14.7%	35.4%	13.5%	34.2%	0.04
Fall DRA2	58.0	18.6	60.8	19.3	0.15
Count (n)	680		654		

Data source: APA analysis of DPS and Reading Partners data

Again, there are some absolute standardized mean differences that exceed 0.10, but all still fall below the absolute threshold of 0.25. For students who took the DRA2, the Reading Partners students were slightly less likely to be Asian and slightly more likely to be Latino than comparison students who took the DRA2. The fall DRA2 scores of Reading Partners students were slightly higher than those of comparison students, but not different enough to worry about the introduction of bias.

As noted above, some of the students in the comparison group attended treatment schools. This factor is discussed in more detail in the Year 3 report, including information about guidance received from SIF's technical reviewers encouraging APA to include these students. Table 8 reviews the number of matched comparison group students drawn from each treatment school. Overall, in 2015-16, a small proportion of comparison students were drawn from the treatment schools: 51 students, or 12.8 percent of the comparison sample were drawn from treatment schools.

Table 8: Comparison Students in Treatment Schools

Treatment Schools	Comparison Students	Proportion of Comparison Students	Assessment Taken
Cheltenham Elementary	4	1.3%	DRA2
College View Elementary	1	0.3%	DRA2
Dora Moore Elementary	1	0.3%	DRA2
TreVista Elementary	18	5.9%	DRA2
Valdez Elementary	8	2.6%	5 iStation, 3 DRA2
Total	32	10.4%	

Data source: APA analysis of DPS and Reading Partners data

With a few notable exceptions, the comparison students drawn from the treatment schools were largely similar to the comparison students drawn from non-treatment schools. Tables 9 and 10 review the group means for the demographic variables used in the match and the pre-test score for students in the overall comparison group, those from treatment schools, those from non-treatment schools, and treatment students, by assessment taken. The tables also present the standardized mean difference, discussed above, between comparison students from treatment and non-treatment schools. As discussed above, when reviewing the quality of a match, the absolute magnitude of the standardized mean difference should be no larger than 0.25, and preferably less than 0.1.

While the groups are similar on many demographic variables, comparison students from treatment schools who took the iStation assessment were much more likely to be female and eligible for FRL than comparison students not from treatment schools. Comparison students from treatment schools who took the DRA2 were more likely to be in Special Education. For both assessments, comparison students from treatment schools were more likely to be Latino and had slightly higher fall pre-test scores than comparison students from non-treatment schools.

Table 9. Comparison Students in Treatment Schools and Non-Treatment Schools, iStation Assessment

Variable	Comparison Students				Treatment students mean
	All students mean	From treatment schools mean	Not from treatment schools mean	Absolute Standardized Mean Difference*	
FRL	85.7%	94.4%	83.6%	0.31	90.1%
ELL	30.8%	22.2%	32.9%	0.23	29.7%
Special Education	12.1%	16.7%	11.0%	0.17	14.3%
Male	52.7%	22.2%	60.3%	0.76	53.8%
Asian	0.0%	0.0%	0.0%	0.00	1.1%
Other Race	1.1%	0.0%	1.4%	0.13	3.3%
Latino	69.2%	77.8%	67.1%	0.23	68.1%
Black	15.4%	11.1%	16.4%	0.15	12.1%
Fall iStation	202.5	207.6	201.2	0.35	201.6
Count (n)	91	5	86		91

* Standardized mean difference between comparison students from treatment and non-treatment schools.
Data source: APA analysis of DPS and Reading Partners data.

Table 10. Comparison Students in Treatment Schools and Non-Treatment Schools, DRA2 Assessment

Variable	Comparison Students				Treatment students mean
	All students mean	From treatment schools mean	Not from treatment schools mean	Absolute Standardized Mean Difference*	
FRL	90.6%	93.9%	90.2%	0.13	89.3%
ELL	37.7%	39.4%	37.5%	0.04	39.9%
Special Education	9.4%	18.2%	8.4%	0.34	8.8%
Male	52.6%	48.5%	53.1%	0.09	51.9%
Asian	2.6%	3.0%	2.6%	0.03	3.6%
Other Race	5.5%	3.0%	5.8%	0.12	5.5%
Latino	64.9%	81.8%	62.9%	0.40	63.3%
Black	15.9%	9.1%	16.7%	0.21	15.9%
Fall DRA2	11.3	13.9	10.6	0.32	11.0
Count (n)	308	27	281		308

* Standardized mean difference between comparison students from treatment and non-treatment schools.
Data source: APA analysis of DPS and Reading Partners data.

Analytic Approach to the Impact Evaluation

For each analysis model, APA used a Hierarchical Linear Model (HLM) with student-level variables at level one and school-level indicators at level two. The HLM technique is intended to ensure that the findings focus only on the effect of the treatment variable, rather than effects of the different schools attended by Reading Partners and comparison group students. The level one model followed the same basic model as the PSM.

Analytic Approach: APA compared the spring DRA2 or iStation scores of the treatment and comparison students, controlling for student-level demographic variables and fall academic performance and using a hierarchical linear model (HLM) to control for school-level variables. Using this basic model, APA examined the overall effect of Reading Partners participation on spring assessment scores, as well as the effect of additional treatment sessions, the differential effect of the program on students with different demographic characteristics, and students with multiple years of Reading Partners participation.

$$Y_i = \beta_0 + \beta_{RP(i)} + \beta_{fall\ DRA(i)} + \beta_{FRL(i)} + \beta_{ELL(i)} - \beta_{sped(i)} + \beta_{male(i)} + \beta_{Latino(i)} - \beta_{black(i)} + \beta_{Asain(i)} + \beta_{other(i)} - \beta_{gr2(i)} - \beta_{gr3(i)} - \varepsilon_{(i)}$$

In this model, the outcome variable Y_i is the spring DRA2 or iStation score (normalized to NCE) for student *i* as predicted by the treatment variable, fall assessment score prior to participation in Reading Partners (normalized to NCE scores), and a set of indicator variables describing whether the student qualified for FRL, qualified for Special Education, identified by the district as an ELL, male, Latino, black, Asian, from another race (with white as the reference variable), in second grade (gr2) or third grade (gr3) (with first grade as the reference variable).

The treatment variable varies by analysis model and evaluation question. Treatment variables included:

- Impact question 1: an indicator of participation in Reading Partners;
- Impact question 2: number of tutoring sessions received;
- Impact question 3: an interaction variable of Reading Partners participation and other student characteristics; and
- Impact question 4: an indicator for number of years of Reading Partners participation.

The second level of the HLM controlled for school-level effects, including the school's district. Because the schools that comparison students came from had been specifically selected to be similar to treatment schools in terms of geographic proximity, district,

> **Summary of findings from HLM models:** The majority of variance in student assessment outcomes scores comes from differences in individual students, rather than the school they attended.

traditional curriculum model, and proportion of students eligible for free and reduced lunch and who were black and Latino, those school-level covariates were not included in the model. School-level variance was relatively small in all of the analytic models. In none of the models did school-level variance account for more than eight percent of total variance. This means that the overwhelming majority of variance in outcome scores is determined by the individual student, rather than the school they attend.

Appendix B includes full results of each model. Overall findings of each model are discussed below.

Impact Question 1 (Confirmatory): Impact of Reading Partners on Student Reading Scores

Does Reading Partners' tutoring lead to improved near-term reading achievement for students in grades one through three when compared to similar students who do not receive tutoring?

To evaluate this impact question, APA used an HLM model that compared the NCE-adjusted spring DRA2 or iStation scores for Reading Partners students to those for comparison group students, controlling for both the demographic characteristics of the students and the school-level effects. In this model, the coefficient for Reading Partners participation was

statistically significant, indicating that APA found a meaningful difference between test score outcomes for Reading Partners students and comparison group students.

Reading Partners students scored, on average, 4.42 NCE points higher than similar comparison students (Table 11).[22] This means students who participated in Reading Partners were more likely to display improved reading achievement compared to students who did not participate. Because of the difference between NCE and percentile scores, the percentile points gained depends on where the student started. See Table 12 for some illustrative examples. The median Reading Partners student started the school year in the 15th percentile in reading scores and moved up to the 21st percentile after participating in Reading Partners, while non-participating comparison students stayed at the 15th percentile. These percentiles indicate a student's relative achievement within the districts where Reading Partners was operating, not a national sample. This is equivalent to an effect size of 0.14.

Summary of Findings for Question 1: Students who participated in Reading Partners showed significantly greater improvement on the DRA2 or iStation assessments than comparable students who did not participate in Reading Partners. For the average Reading Partners student, this improvement is the equivalent of moving from the 15th percentile to the 21st percentile. This is an effect size of approximately 0.14.

In comparison, the MDRC study found an effect size of 0.10 for reading comprehension, 0.09 for reading fluency, and 0.11 for sight word efficiency (Jacob, Armstrong & Willard, 2015). The effect size found in this analysis is roughly equivalent to that found in the MDRC study, though the two studies differ in both their methodology and study population. The effect size of the Colorado program is in line with the average effect size of 0.14 found by Slavin *et al* in their meta analysis of one-on-one tutoring programs (Slavin, Lake, Davis & Madden, 2009).

Table 11. Participation in Reading Partners and DRA2 or iStation Scores

	Pooled Model	
	Coefficient	P-value
RP Participation	4.42	0.002

Source: APA analysis. N = 1,178

---

[22] Please see Appendix E for the complete coefficients from this model.

Table 12. Equivalent Percentile Increases of 4.42 NCEs

Starting Percentile	Ending Percentile	Percentile Difference
10th	15th	5
15th *(median starting percentile of RP students)*	21st	6
25th	32nd	7
50th	58th	8

Source: APA analysis

Impact Question 2 (Exploratory): Impact of Reading Partners on Student Reading Scores as Dosage Increases

Do differences in reading achievement between students who receive Reading Partners tutoring and similar students who are not in Reading Partners increase as students receive more tutoring?

To evaluate this impact question, APA used a model similar to that for impact question 1, above. Instead of a categorical variable indicating whether the student participated in Reading Partners or not, the variable of interest for this model was a continuous variable measuring the number of Reading Partners tutoring sessions a student received. The number of sessions received by Reading Partners students in the analysis ranged from one to 62 sessions, with an average of 37.5 sessions. As seen in Figure 6, the distribution was skewed, with a large group of students receiving more than 40 sessions.

Summary of Findings for Question 2:
APA did not find any significant effects for increased hours of tutoring in the model using only Reading Partners students, which likely lacks statistical power to detect an effect. However, APA found a significant effect of 0.7 NCE point increases for each 10 additional tutoring sessions, when examining outcomes for both Reading Partners and comparison students.

Figure 6. Number of Tutoring Sessions Per Student in the Impact Sample
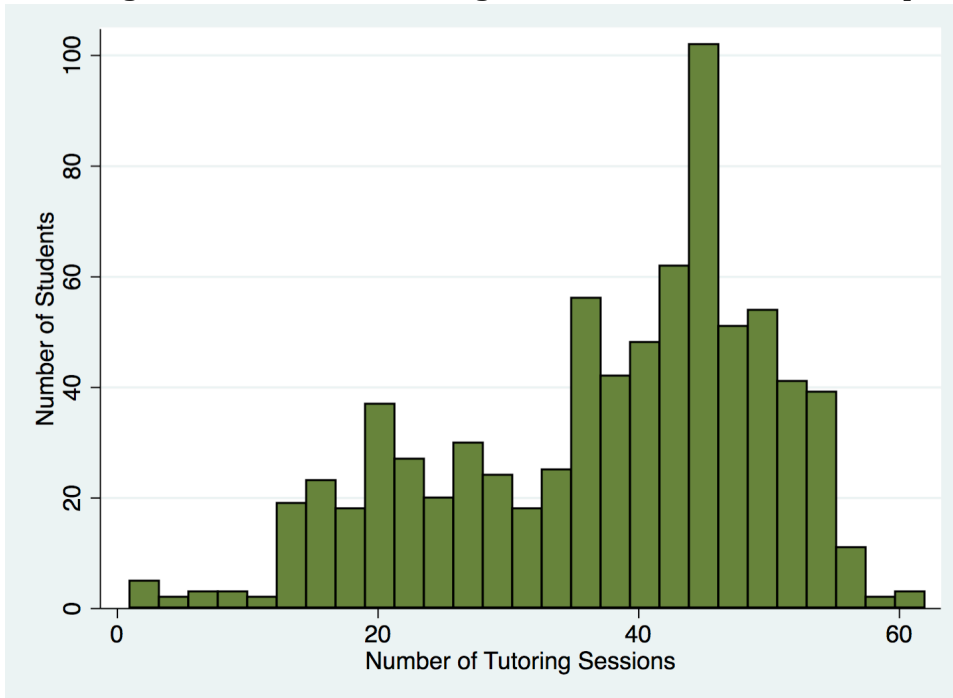
N = 399

APA ran this analysis model two separate ways: First, APA compared Reading Partners students against themselves, excluding comparison students. In this model, the effect of an additional tutoring session was not statistically significant, controlling for student demographic characteristics and school-level effects.[23]

There are multiple reasons why this analysis could have failed to detect a significant marginal effect of an additional tutoring session, even if such an effect exists. First, as demonstrated in Figure 1, the majority of students received above 40 tutoring sessions. Only 10 percent of students received fewer than 20 tutoring sessions. This relative lack of variance in the number of tutoring sessions received makes it more difficult for the analysis to detect the marginal effect of an additional session.

Second, and most important, it is likely that even with three years of pooled data, this analysis is underpowered, meaning that it lacks the statistical power necessary to detect an effect, even if one exists. This is partly because this analysis looks only within Reading Partners students and does not include comparison students, meaning that the sample size is less than half that of the main analysis in impact question 1.[24] A power analysis assuming an alpha level of 0.05 and a power of 0.8 indicated that an analysis with this sample size has a minimum detectible effect size of about 0.28. In comparison, the effect size for the primary research question is 0.14, which is the combined effect of an average of 37.5 tutoring sessions. This means it is likely that the effect of a single tutoring session is too small to

---

[23] Please see Appendix F for output for all remaining models.

[24] The total number of students included in the initial impact model is 1,178, while the number of Reading Partners students who are included in the model examining the effect of additional sessions is 399.

detect with an analysis with 399 students, even if it is likely that a positive effect exists for additional tutoring sessions.

The second model APA used to investigate this research question compared Reading Partners students against themselves and comparison students, who had a value of zero for the dosage variable. This model found a positive and statistically significant effect of an additional tutoring session of 0.07 NCE points, controlling for student demographic characteristics and school-level effects (Table 13). In other words, ten additional tutoring sessions would increase a student's spring score by 0.7 NCE points, again controlling for student demographic characteristics and school-level effects.[25]

Table 13. Additional Tutoring Sessions and DRA2 or iStation Scores, Second Model

	Pooled Model	
	Coefficient	P-value
Additional RP Tutoring Session	0.074	0.003

N = 1,178

It is not surprising that these two dosage models produce different results. The first model, which compares within Reading Partners students, is trying to find the marginal effect of receiving another Reading Partners tutoring session among students who generally received at least 20 tutoring sessions. The lack of a significant result from that model tells us that as long as a student receives the average of about 40 sessions, adding an additional tutoring session does not have a detectible effect given the existing statistical power. In contrast, the model comparing Reading Partners students to comparison students is likely picking up the overall benefit of participating in the Reading Partners program, as was demonstrated in the analysis model addressing research question 1.

Impact Question 3 (Exploratory): Differences in Reading Partners Effects on Different Student Groups

Are there differential impacts of Reading Partners tutoring on different student groups including English-language learners (ELL) vs. non-ELL students, boys vs. girls, grade level, and different races?

---

[25] Readers may notice that multiplying this coefficient by the average number of tutoring sessions does not produce the same overall program effect as the coefficient from the model in research question 1. This is likely because the relationship between number of sessions and spring score increases is not linear, so this linear regression is underestimating the effect of additional tutoring sessions.

To evaluate this impact question, APA built upon the model used for Question 1 to run separate models for each demographic characteristic, adding an interaction term to measure the interactive effect of participating in Reading Partners and being a member of each demographic group. APA also ran separate models to estimate the interaction between Reading Partners participation and each grade level. Each interaction term was the product of the student's value for the indicator variable for Reading Partners participation and the dummy variable for their inclusion in a demographic group.

Summary of Findings for Question 3: APA was able to determine a statistically significant differential effect for ELL students. The gap between ELL students and non-ELL students was smaller for students who participated in Reading Partners than for those who did not. APA was unable to detect any effects for other demographic and

With the increased statistical power provided by the pooled sample, APA was able to detect a differential effect with regards to one demographic characteristic: students who are English language learners (ELL). There was no finding of a differential effect of Reading Partners on students by race or ethnicity, by special education status, by gender, or by grade level.

A significant number of Reading Partners students are classified by DPS as ELL, with 55% of Reading Partners students included in the analysis classified as ELL. As described above, a student must have conversational English skills in order to participate in Reading Partners. The demographic characteristics of ELL and non-ELL Reading Partners students are shown in Table 14, below. As shown, ELL students in Reading Partners are more likely to qualify for free and reduced lunch, less likely to be in Special Education, much less likely to be Black, and much more likely to be Latino than non-ELL Reading Partners students. ELL Reading Partners students also had higher fall assessment scores (in NCE units) than non-ELL students participating in Reading Partners.

Table 14. Comparing ELL and non-ELL Reading Partners Students

Demographic Variable	ELL RP Students		Non-ELL RP Students		Absolute Standardized Mean Difference
	Average	Standard Deviation	Average	Standard Deviation	
FRL	93.9%	29.3%	87.2%	33.4%	0.23
Special Education	8.9%	29.3%	14.8%	35.6%	0.18
Male	48.1%	50.0%	54.7%	49.8%	0.13
Asian	6.1%	15.9%	0.5%	0.5%	0.13
Other Race	0.0%	22.9%	7.8%	7.8%	0.32
Latino	85.7%	47.8%	56.5%	56.5%	0.38
Black	6.1%	36.6%	19.5%	19.5%	0.63
Fall Assessment (NCE)	61.8	20.1	59.3	19.4	0.39
Count (n)	314		384		

Source: School district administrative data, APA analysis

Overall, Reading Partners students showed larger test score gains than those of similar comparison students. In the model for impact question 1, ELL students did not perform significantly differently from non-ELL students. While all Reading Partners students had stronger growth than students who did not participate in the program, APA's analysis found that Reading Partners tutoring had a significantly larger impact on ELL students than general population students in the program.

This finding makes intuitive sense, as Reading Partners is, at its heart, a language-based program. It is easy to imagine that a tutoring session that helps a student learn to read also provides an hour of exposure to and participation in English language development. This is especially true because the majority of the reading instruction in Reading Partners involves oral language, with the student and tutor reading aloud and conversing with one another.

Table 15. Growth for ELL Reading Partners Students and General Population Students

	Pooled Model	
	Coefficient	P-value
Reading Partners x ELL	3.44	0.042
Reading Partners	2.07	0.231
ELL	-2.27	0.037

Source: APA analysis

Table 15 above indicates that, on average, ELL students score 2.27 NCE points lower on the spring literacy assessment than non-ELL students. However, ELL students who also attend Reading Partners score 3.44 points higher, meaning that they score 1.17 NCE points higher on the spring assessment than ELL students who did not attend Reading Partners. The significant coefficient for the interaction term indicates that this 1.17 NCE point difference between ELL students who did and did not attend Reading Partners is statistically significant. These differences are illustrated in Figure 2, below.

Figure 7: Relative Spring Literacy Assessment Score of ELL students in and out of Reading Partners
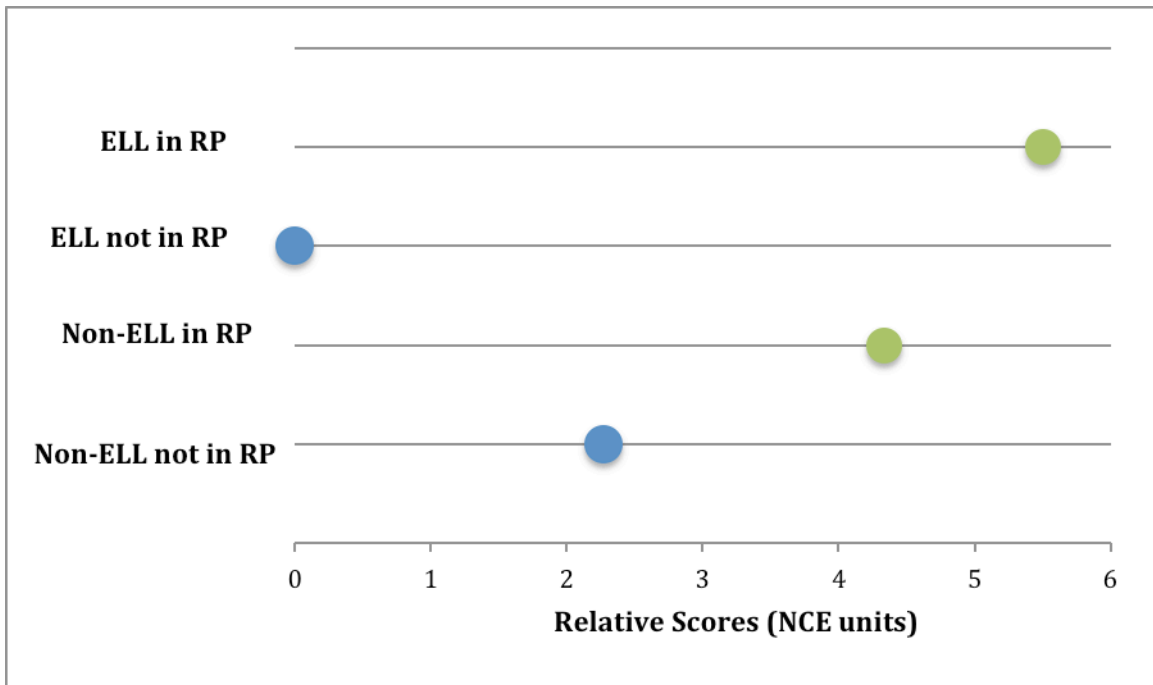

Figure 7 illustrates the relative differences between ELL students, both those who participate and who do not participate in Reading Partners, and similar non-ELL students. ELL students who participate in Reading Partners score almost 6 NCE points higher on the spring literacy assessment than ELL students who do not participate in Reading Partners. The figure also shows the overall effect of Reading Partners, as non-ELL students who participate in the program also outperform similar non-ELL students who do not receive Reading Partners tutoring. Although ELL students who do not attend the program score worse than non-ELL students who do not attend the program, that trend is reversed for students who attend Reading Partners. In other words, ELL students who attend Reading Partners score higher on the spring assessment than their non-ELL counterparts who also attend the program.

This demonstrates that the Reading Partners program, as implemented in Colorado, is particularly effective for ELL students who participate, helping them outperform their non-ELL peers, both in and out of the program.

Impact Question 4 (Exploratory): Impact of Modeling from Multiple Years of Reading Partners Treatment

How do the differences or similarities in the results using the impact and exploratory samples impact judgments about Reading Partners impact on near-term reading achievement?

This question sought to address the differential effects for students who participated in Reading Partners for multiple years. When pooling the three years of data, APA found 95 students who had participated in the program for 2 years and 6 who had participated in the program for all 3 years included in this study. While this model confirmed the positive and significant effect of participation in one year of Reading Partners, the number of students who had participated in two or three years of Reading Partners was too small to provide sufficient statistical power to determine the effect of multiple years of Reading Partners treatment.

Implementation Evaluation: Reading Partners Implementation from 2012-13 to 2015-16

This portion of the final report focuses on overarching trends and change over the first five years of Reading Partners Colorado. The analysis draws upon the four prior years of implementation study and analysis and includes information from a set of summative key stakeholder interviews conducted in fall 2016 and winter 2017, which add to the overall narrative of change and development that occurred over the course of the SIF project. It does not include analysis of data for students served by Reading Partners Colorado in 2016-17.

The implementation evaluation has four sections:

1. a summary of implementation evaluation data collection activities;
2. an overview of program implementation that draws from and synthesizes the four years of program data collected during the study;
3. a description of the organizational context and evolution in Reading Partners Colorado and National; and
4. a summary of insights derived from this study that may be relevant to other volunteer education programs operating within schools.

In brief, the implementation evaluation found that Reading Partners was able to quickly launch, sustain, and implement its program with fidelity in a new region. As is shown in the impact analysis, this implementation led to positive outcomes for the average student in the program.

> **Regional Program Launch:** Reading Partners was able to quickly launch, sustain and implement its program with fidelity in a new region.

APA utilized a life-cycle model as a structure to examine the organizational aspects of Reading Partners National and Colorado (Stevens, 2001). This model is described and discussed later in the report. A key finding from this component of the implementation study is that Reading Partners National and Reading Partners Colorado were at different stages in organizational development during the SIF project in Colorado. Specifically, Reading Partners National was in the growth stage, grappling with the challenge of developing formal organizational structures, while Reading Partners Colorado was in the start-up phase, focused on sharing organizational vision and organizational responsibility with staff and key constituencies.

This report covers all five years of the study, highlighting findings from the first four years of implementation studies and integrating them with information from the fifth year of the study. This report is a supplement to the prior reports and not all of the findings and insights from those studies are repeated here.

Implementation Evaluation Approach and Data Collection Methods

The focus of the implementation evaluation evolved over the five-year period of this study. As described in the sub-grantee evaluation plan (SEP) submitted to CNCS and MHUW, the initial implementation evaluation was focused on monitoring the continued implementation of Reading Partners' program and business-as-usual reading supports in the sample of schools in the impact evaluation. This descriptive monitoring was intended to provide

contextual information for interpreting the results from the impact study. During the period of the study, Reading Partners chose to expand the scope of the implementation evaluation to include new questions intended to inform and support program improvement and management.

This final implementation report draws upon data and information collected over the course of the five-year study period. In conducting its implementation evaluations during each year of the evaluation period, APA used a variety of data collection methods, including review of Reading Partners administrative data, interviews, surveys, and review of student folders.

Throughout the implementation study, APA reviewed Reading Partners program documentation on operations. APA also collected data from Reading Partners Colorado staff through interviews or surveys. Reading Partners National staff provided information for the initial implementation study and this final report. This final report also includes information gathered through interviews with Mile High United Way (MHUW) staff.

Reading Partners Colorado provided administrative data for the first four years of the evaluation. This information included number of tutoring sessions for students and when students enrolled. For the 2015-16 report, the administrative data also included select student demographic information. APA collected data from literacy leaders in the Reading Partners schools through interviews or surveys during the first four years of the study.

Table 16 summarizes the data sources used throughout the five-year study period.[26]

---

[26] Please see Appendix G for detail on data collection activities and responses.

Table 16. Implementation Evaluation Data Sources

Data Sources	Year 1 2012-13	Year 2 2013-14	Year 3 2014-15	Year 4 2015-16	Year 5 2016-17
Document review	X	X	X	X	X
Reading Partners' Colorado staff interviews	X	X	X		X
Survey of Reading Partners Colorado staff				X	
Reading Partners' National interviews	X				X
Interviews of MHUW staff					X
Reading Partners' administrative data	X	X	X	X	
Student folder data			X	X	
Interviews with school Literacy Leaders	X	X	X		
Surveys of school Literacy Leaders				X	
Survey of volunteer Tutors	X		X		
Session observations	X	X			
Literature review			X		
Interviews of leaders in comparison schools			X	X	

The 2013-14 implementation evaluation was limited and was intended to contextualize the impact analysis by describing any changes to the program. After the relatively minimal implementation evaluation in 2013-14, APA expanded the evaluation activities for the 2014-15 implementation evaluation to meet Reading Partner's needs. In response to lessons learned from the MDRC study, Reading Partners asked APA to collect more information from students' folders. This folder information provides information on the pace of tutoring and the number of different tutors working with students. APA reviewed 20 randomly selected student folders per site in the spring and the fall at all of the APS and DPS Reading Partners sites (a total of 452 folders were reviewed), reviewing tutoring data from the last four tutoring sessions recorded in the folder. In 2014-15, APA also interviewed Reading Partners staff both in the spring and fall and conducted a small literature review on school partnerships with outside providers.

The 2015-16 implementation evaluation data collection continued to focus on information gathered from student folder review. APA made slight changes to the methodology for review used in 2013-14. Unlike 2013-14, when APA examined folder data from the prior four tutoring sessions, in 2014-15 APA reviewed all of the sessions across the entire year for every folder reviewed. This look at tutoring notes over the course of the year allowed

APA to describe the pace and rhythm of tutoring throughout the year. However, APA only reviewed ten folders per school, which was not a large enough sample to describe school level factors and their relationship to the pace and rhythm of tutoring. APA was, however, able to link information from the folders to student demographic data collected by Reading Partners.

Implementation evaluation activities for 2016-17 focused exclusively on conducting interviews with staff from Reading Partners National and Reading Partners Colorado about implementation challenges and successes over the entirety of the grant period. APA did not conduct a folder review or other student-level implementation data collection in the 2016-17 year.

Overview of Reading Partners' Program Implementation in Colorado

Throughout all implementation evaluations during each of the years reviewed, APA found the Reading Partners program was implemented with fidelity. The team at Reading Partners Colorado was able to maintain funding to support the program, identify school partners, recruit and train volunteer tutors, identify students that fit Reading Partners' enrollment criteria, assess and develop reading plans for those students, and ultimately providing literacy tutoring to those students using the Reading Partners curriculum.

Volunteers and school staff both saw the Reading Partners curriculum as appropriate for struggling readers. Volunteers felt supported by the program. School leaders reported Reading Partners required much less of their time and engagement to successfully implement compared to similar supplemental programs for students. Schools and volunteers both reported their appreciation for the turnkey nature of the program: school leaders felt that the program required relatively little support from staff to be successful and volunteers said it was easy to participate in the program and use the curriculum.

Table 17 shows the school districts, number of schools and number of students served by Reading Partners, along with the number of tutors engaged and the average number of sessions per student, for each of Reading Partners' first four years in the state.

Table 17. Reading Partners Colorado has Grown

	Districts	Schools	Students Served	Tutors	Average Sessions per Student
Year 1 2012-13	APS, DPS	8 (all new)	323	481	25
Year 2 2013-14	APS, DPS, Sheridan[27]	11 (4 new)	558	803	32
Year 3 2014-15	APS, DPS, Sheridan	13 (6 new)	770	1,332	31
Year 4 2015-16	DPS	14 (9 new)	881	1,219	34

Data source: APA analysis of Reading Partners data

---

[27] Sheridan was not part of the impact or implementation studies.

During this period, Reading Partners had constant growth in the number of students served and number of sessions provided to students. At the same time, Reading Partners Colorado experienced a notable element of change in the schools and districts it worked with. DPS remained an active partner throughout the study period; however, partnerships with Aurora Public schools (APS) and Sheridan Schools were more short-lived. By 2015-16, Reading Partners Colorado was working exclusively in DPS. While Reading Partners did serve at some DPS schools through the entirety of the evaluation period, there was also yearly change, with new schools added and existing schools leaving Reading Partners every year after 2012-13, in part due to changes in school needs and resources and in part due to funding issues.

Consistent Program Delivery: Reading Partners served students of different races and ethnicities equally: there were no differences in the rate of tutoring or number of tutors students had by students' race or ethnicity.

In 2015-16, APA also examined relationships between program delivery and key student characteristics by connecting the student folder data with demographic data provided by Reading Partners Colorado. APA found there were no differences in the rate of tutoring or number of tutors students had by students' race or ethnicity. Further, the racial and ethnic demographics of students served by Reading Partners Colorado were similar to those of the larger student bodies at the schools served by Reading Partners.

Over the first two years of implementation, Reading Partners Colorado staff developed capacity to smoothly implement the tutoring program. By 2014-15, staff understood the program and priorities for good implementation. For example, a Program Manager remarked that space for Reading Centers was less of a problem because they were better at negotiating with schools for adequate space.

At the same time, Reading Center Coordinators described growth in their support to tutors. In 2014-15, APA collected information on how the focus of a Coordinator's work changed throughout the year. At the beginning of their year, their focus is on setting up the Reading Center and supporting students. The focus of most Reading Center Coordinators evolved throughout the year to place greater emphasis on supporting tutors, both in terms of their instruction and their relationships with students.

The next sections provide highlights from data collected during the five years of implementation analysis. The first section provides information on the dosage, pacing and number of tutors. Included in this section is a discussion about the role played by the Reading Center Coordinator in determining the student experiences in terms of dosage, pacing and number of tutors. This is followed by a discussion of programmatic changes made by Reading Partners National. The third section discusses changes faced by Reading Partners Colorado during implementation. The final section describes organizational challenges and opportunities that occurred as Reading Partners National and Colorado developed as organizations.

Tutoring: Dosage, Pacing and Number of Tutors

Administrative data provided by Reading Partners illustrates the number of sessions students received, which is a key part of program implementation. The program has a target that each student should receive at least 20 sessions over the course of the school year. Figure 8 uses Reading Partners administrative data from 2014-15 to illustrate the count of students by the number of sessions each student received, with Reading Partners' target of 20 sessions highlighted with a black bar. The majority of students received at least 20 sessions: 75 percent of students received 20 or more sessions, over 50 percent of the students received 30 or more sessions, and one-third participated in 40 or more sessions. The average 2014-15 student received 31 sessions.

> **Tutoring Dosage:** Throughout the Reading Partners Implementation, most students received more than 20 tutoring sessions during each school year.

Figure 8. Most Reading Partners Students Received 20 or More Sessions
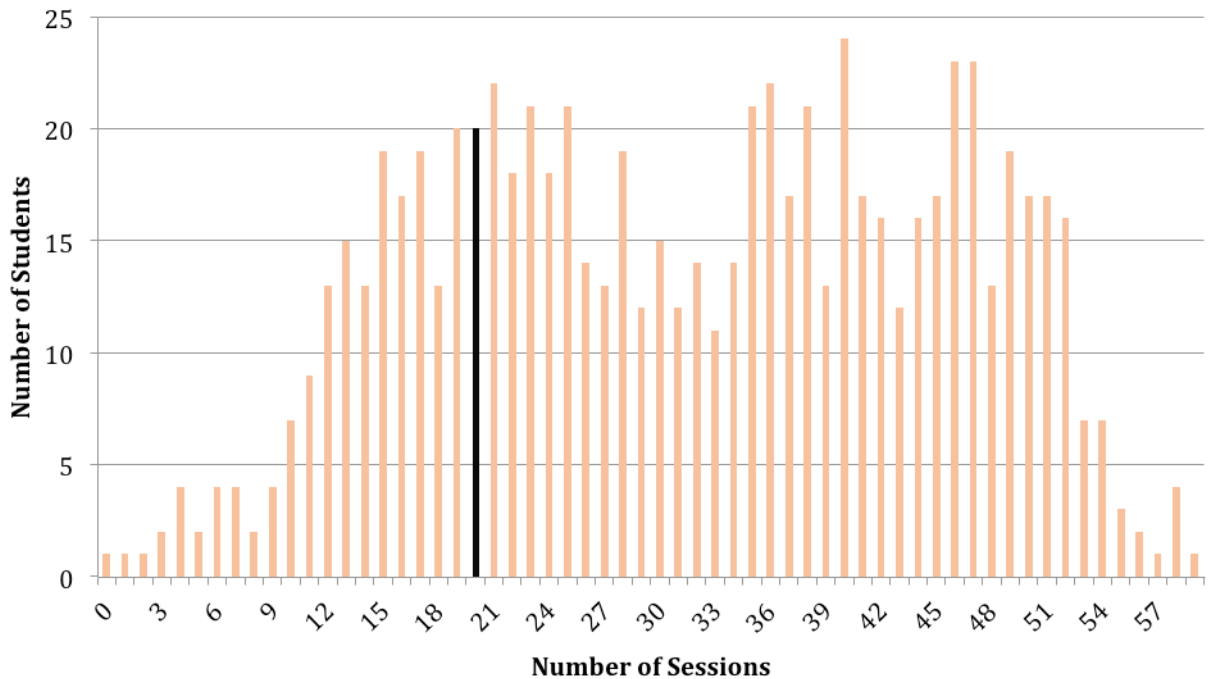
Source: Reading Partners Administrative Data, 2014 15, n=770

Pacing

To provide a more detailed picture of program delivery, particularly the pace and rhythm of tutoring sessions, APA reviewed information recorded in Reading Partners student folders during the 2014-15 and 2015-16 school years. In 2014-15, APA's folder review focused on program delivery during a two-week period in the fall and another in the spring, providing a snapshot of tutoring activities at two points in the school year.

APA found the pace of tutoring varied between schools. As Table 18 shows, during the fall, the proportion of students receiving two sessions a week ranged from 63 percent to 95 percent. Similar ranges were seen in the spring, with 60 percent to 95 percent receiving at

least two sessions per week. Over the course of the school year, on average, the proportion of students receiving two sessions a week remained steady at 75%.

Table 18: Proportion of Students Receiving Two Sessions a Week (2014-15 data)

Data Collection Period	Minimum	Maximum	All Schools	Count
Fall 2014	63%	95%	76%	211
Spring 2015	60%	95%	75%	247
Combined	67%	84%	75%	456

Source: APA extracted data from Reading Partners Colorado student folders, 2014 15

In 2015-16, APA modified its approach to student folder reviews, shifting from collection of data for fall and spring snapshots to collection of data for the entire program year. This allowed more in-depth analysis of the ebb and flow of tutoring sessions over the course of the school year. Figure 9, below, uses folder review data from the 2015-16 folder review to illustrate the rate of tutoring sessions per week over the course of the school year. As shown, Reading Partners Colorado had a quick start to tutoring. Cohorts of tutors were trained during the summer and when the school year began, Coordinators worked with principals to identify students for tutoring based on student performance in the prior year. This lead to a quick ramp-up of tutoring, which helped Reading Partners Colorado be the first region to hit enrollment goals for the year. Tutoring was maintained at a relatively high rate throughout the spring semester, peaking in March.

Figure 9: Tutoring Sessions Per Week for the 210 Students in the folder sample in 2015-16
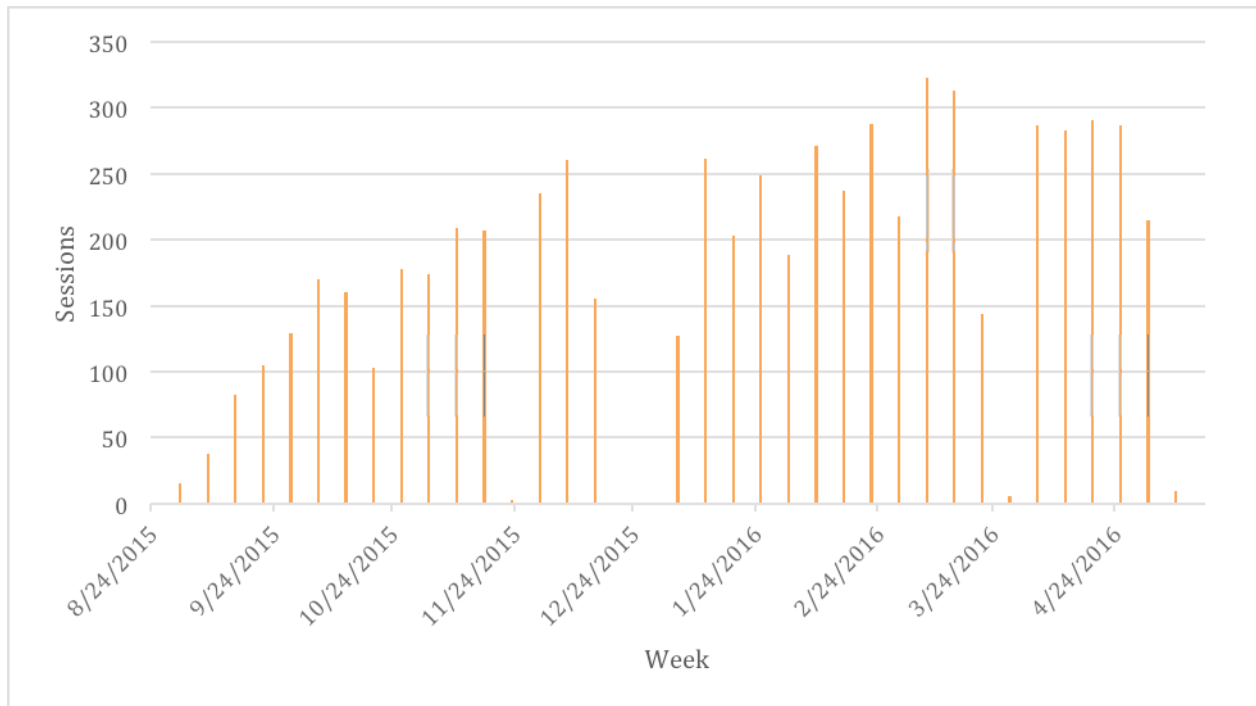
Source: APA extracted data from Reading Partners Colorado student folders, 2014-15, n=210

The 2015-16 folder review data also provided insight on general relationships between student enrollment, the number of tutors, and rate of providing sessions. In general, students who enroll in Reading Partners earlier in the year generally receive more tutoring sessions. There was also evidence of Reading Partners staff serving as tutors to ensure that students received the desired number of tutoring sessions per week and over the course of the year. Reading Partners staff, such as Reading Center Coordinators or Program Managers, provides about one in ten tutoring sessions.

Reading Partners seeks to provide students with two tutoring sessions every week. APA, like MDRC, analyzed available data to access the extent to which this goal is achieved for students. Although APA and MDRC used different metrics when analyzing the folder results, APA's findings were similar to MDRC's with respect to the pace of tutoring for students. MDRC found that students received an average of 1.5 sessions every week (Jacob, Armstrong & Willard, 2015). Similarly, APA found that 75% of students received two sessions per week. Both measures roughly equate to about 3 sessions every two weeks.

Number of Tutors

Another primary goal of the Reading Partners program is building caring relationships between tutors and students. This means that Reading Partners works to minimize the number of different tutors that work with an individual student to allow formation of stable relationships between students and tutors.

A key issue in analyzing tutor consistency is that the more sessions a student has, the more opportunities she or he has to see multiple tutors. This relationship is illustrated in Figure 10, which shows that as the number of sessions increase, the number of tutors tends to increase. This relationship is confirmed by regression analysis, which indicates that on average, each five additional sessions for a student is associated with meeting two new tutors.

Figure 10. Number of Tutors, Plotted by Number of Sessions
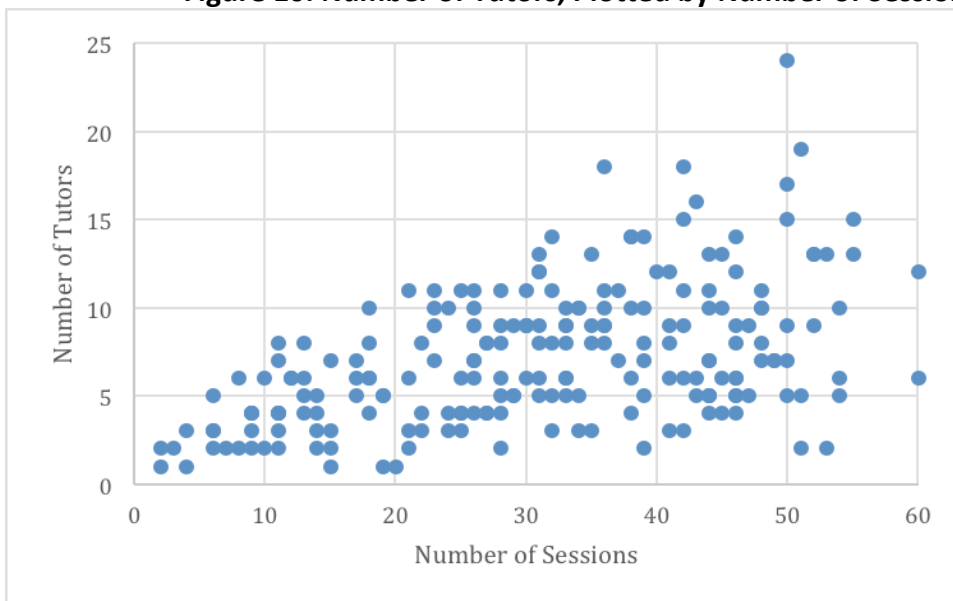
Source: APA extracted data from Reading Partners Colorado student folders, 2014-15, n=210

Seeing multiple tutors does not mean that students did not have opportunities to build relationships with another caring adult. To examine this issue, APA analyzed how many sessions students had with their primary tutor, the tutor the student worked with the most. Figure 11 shows the percentage of sessions provided by the student's primary tutor, while Figure 12 shows the number of sessions provided by that primary tutor. As shown in Figure 11, the median student received 45% of sessions from their primary tutor, marked with a black bar, while Figure 12 shows that over half of students had a primary tutor who provided at least 13 sessions.

Figure 11. Percentage of Sessions with The Primary Tutor
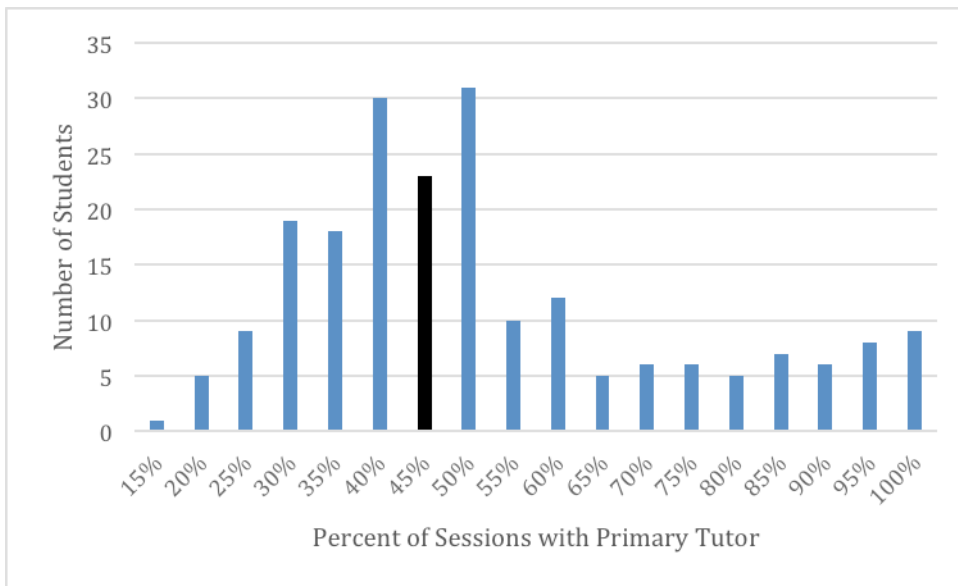
Source: APA extracted data from Reading Partners Colorado student folders, 2014 15, n=210

While the majority of students had a significant number of tutoring sessions with a single tutor, 20 students had a primary tutor who provided fewer than five sessions. However, five of those 20 students received five or fewer tutoring sessions over the course of the year.
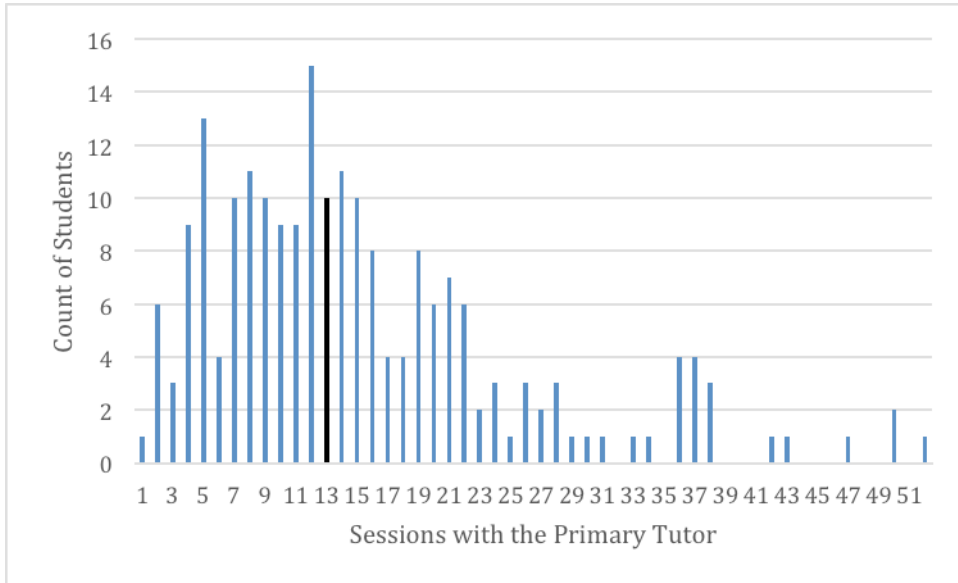
Figure 12. Number of Sessions with Students' Primary Tutor

Source: APA extracted data from Reading Partners Colorado student folders, 2014 15, n=210

It is important to note that these figures present the overall experiences of sampled students, while individual experiences can vary widely. For example, sampled data included one student who had a single tutor for over 50 sessions and a student who had 24 different tutors over the course of the year.

Again, APA and MDRC found similar results around the number of tutors students saw. MDRC assessed the goal to minimize the number of different tutors that an individual student works with by counting how many tutors students saw over four sessions. They found that 59% of students saw three or fewer tutors over those four sessions. APA used a slightly different metric, showing that about half of the students in Colorado saw one or two tutors over a four-session period.

Role of Reading Center Coordinators

MDRC identified unreliable volunteers as a primary barrier to Reading Partners' ability to ensure that students receive at least two sessions a week from a consistent adult. Although volunteers are an important component of meeting those goals, APA found the challenge was more complex and closely linked to the role of the Reading Center Coordinator. At each school site, Coordinators broker a number of schedules, including that of the tutor, student, and the student's teacher, in order to reach the program goals. Figure 13, below, depicts how these factors interact to affect the number and frequency of tutoring sessions. The importance of each of these factors means that consistent tutoring depends not only on consistent volunteer participation, but also on solid communication and clear scheduling among Reading Center Coordinators, tutors, teachers, and schools.

Figure 13: Coordination Required for Tutoring to Occur
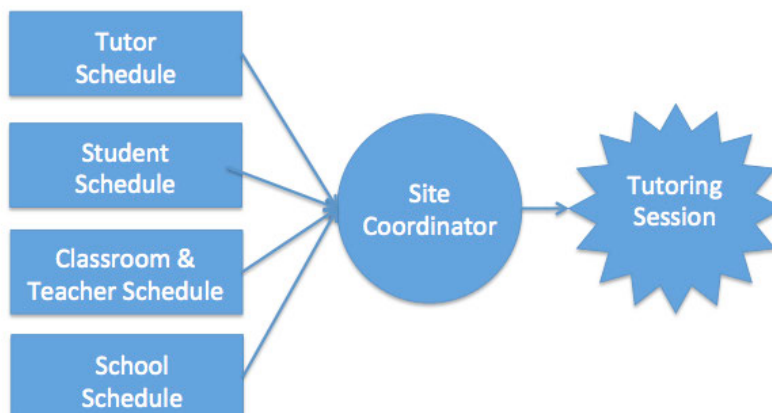
The Reading Center Coordinator serves as the broker of all this communication and scheduling, depending on information from volunteers, teachers, and students to manage and respond to changes that might interfere with the daily tutoring calendar. Good relationships between site coordinators and tutors, teachers, and school staff can facilitate many of these scheduling issues. These challenges diminished over time as Coordinators, tutors and school staff became more familiar with the program and expectations.

APA's interviews with Coordinators and other Reading Partners staff found that the Coordinators were fairly effective at navigating any one challenge, such as a rigid school schedule limiting when students can be tutored. However, depending on the school site and context, Coordinators often encountered multiple schedule complications. For example, one Coordinator worked at a school that had both many unscheduled field trips by teachers and college student volunteers who were unable to tutor during finals. Another Coordinator was at a school with frequent student absences and teachers who were inflexible about allowing students to be pulled from classrooms for make-up sessions. Challenges with more than one group of people (volunteers, students, classroom teachers, and school administrators) reduced the number of students receiving two sessions a week and increased the number of tutors seen by each student.

APA conducted a brief literature review on developing effective partnerships between outside providers and schools. Existing literature indicates that is important to have program and school staff that take on liaison-type roles, acting as "boundary spanner[s]" or bridges between programs and schools (Firestone & Fisler, 2002, 449; Goldring & Sims, 2005, 245). Securing buy-in from classroom teachers is especially important. Finally, programs working in schools should periodically evaluate their own progress, making success as visible as possible. These findings support Coordinators' focus on building relationships with staff. Their communication efforts include mid-year progress reports and Reading Center celebration events with tutors and school staff.

Yearly Programmatic Changes

The implementation evaluation catalogued changes, challenges and opportunities that occurred over the first four years of the implementation in Colorado. Reading Partners National implemented several important programmatic changes:

- Revised curriculum for and increased emphasis on serving students in the early grades (K-3);
- Modified approach to orienting and training new Reading Partners volunteer tutors;
- Change in the literacy assessment used by Reading Partners with enrolled students; and
- Revised student enrollment criteria.

At the same time, Reading Partner's Colorado also faced a set of evaluation and fund-raising challenges and opportunities including those around being a SIF sub-grantee. Additional challenges are also discussed in the organizational development context.

Reading Partners National Programmatic Changes.

Revised curriculum: Because of the SIF focus on early reading, the program in Colorado focused more heavily on serving students in the earlier grades (kindergarten through third) than did other Reading Partners regions, which hastened an organization-wide redesign of curriculum for the earliest readers during 2012-13 school year. As a result of the redesign, Reading Partners revised the existing Alphabet Reader curriculum to the new Emerging Reader curriculum. This focus on early grades in Colorado presaged a larger shift in focus by other Reading Partners regions, which in subsequent years expanded the number of early grade students served across Reading Partners' national network.

Tutor Training: An important programmatic change during the 2013-14 school was a shift of tutor training from one-on-one tutor shadowing to more group trainings, during which multiple new tutors receive training together. This development supported Reading Partners Colorado's ability to quickly meet enrollment goals in 2014-15 and 2015-16.

Literacy Assessment: During the 2015-16 school year, Reading Partners made a change to the literacy assessment used to place students in the tutoring curriculum and to internally measure progress[28]. Prior to the 2015-16 school year, Reading Partners used the Rigby Ultra PM assessment, which the Reading Center Coordinator administered to each student, one-on-one. In 2015-16, Reading Partners National changed the assessment to the STAR assessment, which is an online assessment. According to Reading Partners National staff, the assessment transition went smoothly in Colorado and the new assessment took less Coordinator time to administer.

Colorado prepared for the shift piloting use of the assessment in 2014-15 and had experience with the assessment because it was already used in some of its partner schools. However, staff of both Reading Partners National and Colorado indicated that there was some confusion early in the year about how to use assessment scores to identify the appropriate curriculum unit for a student. This confusion arose because Colorado schools

---

[28] The student assessments used for the impact analysis were administered by the school districts in compliance with state law. These changes to Reading Partners' internal assessments did not affect the impact analysis.

start earlier in the year than schools in other regions and Reading Partners National had not yet completed the crosswalk between assessment scores and curriculum units. This challenge was resolved when Reading Partners National issued a new assessment score to curriculum crosswalk.

Program Enrollment Criteria: Also during the 2015-16 school year, Reading Partners revised the criteria used to identify students who are eligible for enrollment in the program. Until 2015-16, Reading Partners enrolled students only if they were between six months and two-and-a-half years behind grade-level. However, in response to multiple factors, including the shift in literacy assessment and concern that the existing criteria excluded kindergarten and first grade students that could be effectively served by the program (because being six months behind in kindergarten or first grade represents a more significant lag than being the six months behind in third grade), Reading Partners began enrolling students who were between one month and two-and-a-half years behind.

Challenges Faced by Reading Partners Colorado

Challenges from being in the MHUW SIF Sub-grantees**:** While participation in SIF provided Reading Partners with valuable opportunities to expand its programming and continue building evidence of program impact, there were also challenges that stemmed from being in the first class of SIF sub-grantees from MHUW.

SIF's rigorous evaluation requirements proved challenging. As the SIF project in Colorado was getting underway, the federal government was in the process of increasing the requirements for rigorous evaluations of grant programs as part of a movement towards evidence based policy-making. These rigorous evaluations required technically skilled evaluators, increased data from both Reading Partners and district partners, and implementation to meet study designs. In Colorado, MHUW made the strategic decision to work with local evaluation partners instead of national evaluation partners in order to build local capacity. The SIF evaluation requirements required technical evaluation capacity and implementation of programs in a fashion that allowed for rigorous research design.

A second challenge as a SIF sub-grantee was fundraising. MHUW initially chose 11 sub-grantees to support with the grant. These sub-grantees had matching and often additional operational requirements that required them to raise funds within the Denver metropolitan area. At the same time, MHUW also had to raise funds to support both the SIF grant and their own ongoing operational costs. This sometimes resulted in competition among sub-grantees and between sub-grantees and MHUW for donor support. As the grant progressed, MHUW worked with sub-grantees to help educate the donor community about the SIF grant, which helped to relieve some of this tension.

The fundraising and evaluation challenges were insurmountable for some of the sub-grantees, which is part of the reason that only 6 out of the original 11 sub-grantees completed the five-year SIF grant period. However, while demanding, the rigorous evaluation requirements also lead to increased technical capacity at MHUW, among sub-grantees, and within the Denver evaluation community.

Reading Partners' Organizational Evolution

This section provides an overview to the organizational challenges and changes faced by Reading Partners, at both the national and regional levels. It begins with an overview of the Life-Cycle model of non-profits which provides a framework for understanding the organizational challenges and changes facing Reading Partners in Colorado and nationally.

Lifecycle model

APA has examined Reading Partners Colorado implementation during the first four years of the project (school years 2012-13 through 2015-16). A significant number of the changes facing Reading Partners, both in Colorado and nationally, reflect challenges inherent to organizational development. One tool for contextualizing those changes is the *Nonprofit Lifecycles* model (Stevens, 2001). This model identifies key challenges facing organizations as they progress through each life-cycle state from idea, start-up, growth, maturity, decline, turnaround, and if turnaround is not successful, terminal. Table 19 highlights some of the challenges associated with each stage in the model.

Table 19: Overview of the Non-Profit Lifecycles Model

Stage	Challenges
Idea	• Converting idea to action • Mobilizing the support of others
Start-up	• Sharing vision and organizational responsibility with staff, Board and constituencies • Hiring versatile staff
Growth	• Beginning to formalize organizational structure • Creating a program and strategic focus that does not trap creativity and vision
Maturity	• Keeping staff motivated around the mission • Becoming "position" rather than "person" dependent
Decline	• Reconnecting with community need
Turnaround	• Finding a turnaround champion and letting them lead
Terminal	• Accepting responsibility for the organizations renewal or termination

Source: Stevens, 2001

From the perspective of the *Nonprofit Lifecycles* model, Reading Partners Colorado and Reading Partners National were at different development stages and thus had different organizational needs and challenges over the course of the SIF project. During the evaluation period, Reading Partners National progressed through its *growth* stage and began to enter its *maturity* stage. The key challenge overcome by Reading Partners National during this growth stage was the formalization of organizational structure and policy. This formalization takes the form of both building up specialized staff for key functions and roles and development of clear operating procedures.

While Reading Partners National was in the growth stage, Reading Partners Colorado was essentially a startup organization. At the beginning of the evaluation period, Reading Partners Colorado's presence and relationships with the local education community had not been established. For Reading Partners Colorado, the key challenges were around hiring a versatile staff that could share the organizations vision and responsibilities with constituents. The goal of this shared visioning was to garner support from volunteer tutors, funders and schools.

Reading Partners National Life-Cycle Development

While Reading Partners National was expanding to three regions, it was also building specialized staff capacity. As late as 2011, Reading Partners did not have a national Human Resources Director, Information Technology Director or a standardized onboarding process for new employees. Further, the national program did not yet have an operations manual for use by new regions, instead having information about operational aspects of programs held by long-standing staff members. The main resources available to start-up regions were checklists of program and Reading Center components. In the words of one interviewee, "there was a need to put this all down on paper and capture what areas are flexible and where do you need to follow the strict program operations." According to interviews with Reading Partners National staff, documentation of start-up-processes and program fidelity began in 2013-14 (Year 2).

Evaluation data identified the challenges around developing a formal organizational structure led to a recommendation in the 2012-13 evaluation report for a purposeful examination of Reading Partners National policies, culture, and shared understanding regarding expertise within its field offices. Key questions identified by APA for Reading Partners National in the 2012-13 evaluation report were:

a. Where does expertise reside within Reading Partners? How can the organization best use that expertise?
b. What are the core concepts of the Reading Partners program that must remain consistent across sites and what can vary? If things can vary, what can that variation look like, and how can the capacity to be flexible or innovate be clearly and consistently communicated to state and national staff?
c. What are the components of the program that need to be improved and how can field offices provide input to that improvement?
d. How to best utilize expertise from their peers in other states to help the staff in start-up states without creating undue burdens?

Over time Reading Partners National refined its approach in several of these areas.

Changing Expectations for Regional Leadership

Starting in 2013-14 and through 2014-15, Reading Partners Colorado leadership also described evolving expectations of and demands upon Executive Directors as Reading Partners National reconceptualized this position as the organization developed. This evolution was described as moving from "middle manager to mini-CEO." This shift involved additional autonomy and flexibility at the local level and shifted responsibility to Executive Directors to manage tutor recruitment, fundraising and community outreach in order to support the larger program.

Allowing regional Executive Directors this autonomy and flexibility was partially driven by the acknowledgment of local differences in three of the four major program components illustrated in the process model in Figure 5:

- Tutors: effective recruiting techniques vary by region (e.g., word of mouth may work well in some regions, while digital ads work better in others);
- Fundraising: the size of donations and number of donors varies by region (e.g., some regions depend on a few large donors, while others rely on many smaller donors); and
- Schools: there is variation by district in school authority to initiate and manage relationships with service providers and variation by school in literacy program focus (e.g., some school districts manage these partnerships in a highly-centralized manner, while others allow individual principals to freely enter into agreements with service providers).

The challenge for Reading Partners National is how to provide and support regional needs.

Strategic Planning Challenges

Reading Partners Colorado's need for strategic planning support was a theme in several of the annual evaluation reports. In the 2012-13 report, a key challenge for Reading Partners Colorado's leadership was around strategic site selection, in particular selecting sites that provided enough eligible students to meet student enrollment goals while also being located near key sources of volunteers. Reading Partners Colorado had to add an additional site near the middle of the 2012-13 school year in order to address challenges with student enrollment and used paid tutors at some sites without access to an adequate pool of tutors. A formalized strategic plan or approach for site selection might have helped avoid these problems.

A similar strategic challenge was identified in 2015-16 by Reading Partners Colorado leadership, fundraising. Reading Partners Colorado lacked both a strategic plan for fundraising and the data infrastructure needed to develop and manage a diversified and stable fundraising base. The context for fundraising can vary significantly by region, which requires regional leaders to develop both a strategic approach to match their local context and have access to supporting data systems to manage relationships with the large population of possible donors. For example, Colorado has a large education related non-profit community that creates both competition for donations and a large population of professionals interested and willing to support charitable causes. This means Reading Partners Colorado needs a strategic fundraising approach in order to secure sufficient and reliable funding, as well as the data management system to manage relationships with the large population of possible supporters. The absence of strategic fundraising guidance and a strong system for managing fundraising information made it difficult for Reading Partners Colorado to establish a stable and sustainable base of funders.

Development of Program Support Tools

The introduction of new student progress monitoring tools during the course of the SIF study serves as an example of how the organization leveraged expertise held by the Colorado team to develop formal structures to ultimately meet the needs of all Reading Partners regions.

A challenge voiced in 2012-13 by both Reading Partners Colorado and National staff was around tools to support effective program implementation. In 2012-13, the primary program support tools for regions were checklists around setting up Reading Centers and ensuring that all components were in place and appropriately visible to students and tutors. While setting up the Centers correctly was an important goal to accomplish during the first year of implementation, it did not address the core activity of the organization: quality tutoring and tutor effectiveness.

The presence of experienced educators serving as Reading Center Coordinators in Colorado meant that the region was uniquely positioned to help strengthen program support tools and meet this larger organizational need. An effort that began in one Reading Partners Colorado site during the 2013-14 school year led to the development of new and enhanced program support resources that were then disseminated nationally in the 2014-15 and 2015-16 school year. At the mid-year assessment period, one Reading Partners Colorado Coordinator was surprised by the progress, and lack of progress, she was seeing in some of her students. This Coordinator, who had recently completed a master's in education degree, decided to work on systems to more closely monitor individual students so that she would not be surprised by student progress again. This Coordinator's efforts led in the following year to the development and piloting of new progress monitoring tools throughout Reading Partners Colorado; these tools were then distributed across all Reading Partners regions in following years. These tools helped Reading Center Coordinators track student progress and work with tutors to focus their work on meeting each student's needs.

The development of the progress monitoring tool is an example of effectively using local capacity to meet national needs for the development of formal structures that occurs during the growth life-cycle stage.

At the same time, the development of the progress monitoring tools also illustrates evolution of the program support focus by Reading Partners National. By 2014-15, when the progress monitoring system was being developed across Reading Partners Colorado sites, the focus of program support tools was shifting from the look of Reading Centers to strategic program components such as the growth and capacity of staff to support relationships and tutoring. The progress monitoring tool met a need created by this changing focus of quality control from checklists to supporting tutor effectiveness.

Different Reading Partners Organizations

The view of Reading Partners Colorado and National being at separate organizational development stages also highlights the two-way relationship between regional and national offices. Specifically, Reading Partners Colorado staff raised the question of whether the services provided to the regional sites were worth the fees paid to Reading Partners National out of regional fundraising support. During the four years of the evaluation, Reading Partners National grew its capacity to support regions. However, the question of value-added of national was a consistent question for Reading Partners Colorado staff.

During the evaluation period Reading Partners National was engaged in the work of a non-profit in the growth stage, in particular developing formalized structures that both defined and supported regional roles. At the same time, Reading Partners Colorado was in the start-up phase. The challenges of this phase are hiring a versatile staff and communicating the

vision and organizational responsibility with staff the board and constituencies. The key staffing challenge for reading Partners Colorado was around leadership.

Leadership Challenges

During 2013-14, Reading Partners Colorado faced the first of two leadership changes it would experience during the evaluation period. The initial Executive Director departed after the first year of operations in Colorado and an interim Executive Director was appointed in February of 2014. The interim Executive Director had worked on the National operations team for several years. The appointment was intended to be temporary, but lasted through the third year of implementation.

Reading Partners was able to successfully navigate the serious threat that these leadership changes represented to the program's ability to succeed in the Colorado region. A primary challenge to Reading Partners Colorado in the startup phase was communicating the vision of the organization to constituencies and establishing a local presence for the program, and the Executive Director is the main communicator for the regional program.

Reading Partners Colorado was able to hire a permanent Executive Director in the summer prior to 2015-16, the third leader in four years. Each Executive Director has brought different strengths and personality to the position.

A key role of non-profit leaders during the start-up phase is communicating the vision and roles of partners to secure their support of the program. Reading Partners Colorado faced both challenges and successes in securing the support of schools and funders.

Successes and Challenges in Securing School Support

Reading Partners Colorado staff describe increased focus on and sophistication of their relationships with schools. School leaders see Reading Partners as one of many vendors working in a school. For the program to have an effective relationship with schools, Reading Partners had to understand how it helped the school meet its priorities for improvement and how the program fit into their system of supports for students. The large majority of schools used the Response to Intervention (RtI) framework to structure their student supports.[29] The structure of the RtI framework allowed Reading Partners to provide essential Tier 2 supports to struggling readers, supporting the schools' existing student support structure. Reading Partners Colorado interviews indicate that building and managing relationships with schools became a primary focus during this period.

While building and managing relationships with schools was a key focus, this effort was not always successful. A surprising change at the end of the 2014-15 school year was the decision by Aurora Public Schools (APS) to end its partnership with Reading Partners Colorado. There has been no clear explanation of why the change occurred. APA interviewed school liaisons at the several APS schools where Reading Partners operated and found that school staff, including principals, were generally happy with the program. Despite this positive perception of the program at the school level, the central office decided to no longer work with Reading Partners. This sudden departure highlights the key

---

[29]This system has three levels: Tier 1, which is the instruction and supports given to the large majority (about 80%) of students.; Tier 2, which is the supports and instruction given to struggling students, usually 10-15% of students; and Tier 3 supports for students with disabilities.

challenge for regional offices in building and maintaining constituencies throughout the region.

Challenges and Successes in Securing Funding Support

Securing financial support of the program became an acute challenge in 2016-17. During planning for the 2016-17 school year, Reading Partners Colorado faced significant financial challenges, in part because of the reduction in support available as the SIF project came to an end. While SIF funding was reduced as natural evolution of the SIF process, not because of program performance, this reduction in funding increased the fundraising demands on the regional team. At the same time, DPS was facing its own budget challenges and was slow to finalize their allocation of funds to Reading Partners. These funding challenges had several impacts on the program. First, funding issues led to reductions in the number of schools the program planned on serving in 2016-17, dropping from 14 schools to 9. It also created uncertainty about whether positions would be available for all of the Program Managers who wanted to return in 2016-17.

The 2016-17 funding challenges where addressed by expanding the audience for communicating the Reading Partners vision and the role for community support. Reading Partners Colorado developed new community fundraising events and refinement of events that are targeted towards growing support from tutors and their networks.

Lessons for Other Volunteer-based Education Programs

While this implementation evaluation focused on Reading Partners' program in Colorado, findings from this study may provide important lessons relevant to other volunteer education programs.

Program Development

First, programs must emphasize development of the core content of their programs. Reading Partners National first invested in developing its core instructional program, developing a research-based program that was very easy for volunteers and recognized by teachers as appropriate instruction for struggling readers. This easy-to-use program helped develop support for the program from other core program supports: tutors and teachers. While Reading Partners National continued to refine its instructional program during the implementation period in Colorado, this core curriculum was well developed when Reading Partners began to work in Colorado.

For both tutors and schools, Reading Partners was essentially a turnkey program, requiring little startup time or investment from either volunteer tutors or participating schools. The curriculum was easy for tutors to follow and feel like they were making progress with students as they moved through the different units. For schools, Reading Partners required very little support from school staff. Reading Partners staff learned to treat schools like customers: communicating with teachers and staff about the program and its successes, and working to integrate the program into regular school operations. A core role for Coordinators was to manage and respond to the schedules of tutors, teachers and schools to make sure that tutoring occurred at times that were most convenient to these core constituents.

School Partnerships

Programs must cultivate strong working relationships with school staff and leadership. Relationships between Reading Partners and school staff are crucial to schools allowing the program to have access to students. Teachers must know that students will benefit from the tutoring more than they benefit from class time or teachers will be reluctant to allow students to attend tutoring. They need to know the curriculum is aligned with their expectations for a good reading program. Tools for establishing and building this relationship are:

- When negotiating with schools at the beginning of the relationship, it is important to be very explicit and clear about program needs for success. For Reading Partners, this includes adequate space for a Reading Center, a flexible schedule that will allow access to students at multiple points during the day, and opportunities to build relationships with teachers;
- Presentations by Reading Center Coordinators or other Reading Partners staff to school faculty before the beginning of the year;
- Visibility of Coordinators and tutoring sessions: a centrally located Reading Center can be noisy but can also allow teachers to see the tutoring and get to know the Coordinator;
- Coordinators need a regular point of contact with the school. They should meet regularly to review operational issues such as up-coming events that will require schedule modifications and student challenges or successes;
- Regular communication between Coordinators and school staff. This includes both newsletters and updates to teachers on their individual students' progress. These updates should occur at least once during the year. By discussing student's strength and challenges in reading skills and progress using a standardized assessment, the Coordinator communicates to teachers both an important knowledge of individual students, as well as program success;
- Celebration events during the year offer opportunities to tutors, Coordinators and school staff to build relations and learn about each other; and
- Participation of Coordinators in school events, as a member of the school's instructional community, require a level of buy-in by school leadership, but also further the relationship with staff.

Volunteer Engagement

Programs that depend on volunteers should be located near reliable and robust sources of volunteers: schools and colleges can provide students, established suburbs can be good sources of seniors, and downtown areas or other areas with high concentrations of office buildings can provide working volunteers. Volunteer coordinators can help build relationships with schools, colleges and large employers to tap into these sources of volunteers.

Prepare for Turnover

Programs must expect and be prepared to navigate a certain degree of turnover among school partners, school staff, program staff, students and tutors. For example, nationally, Reading Partners reports about a third of schools do not return to the program from year to year. Based on publicly available school-level data, about one in six Colorado principals left

their schools between the 2014-15 and 2015-16 school years.[30] Reading Partners Colorado's three Executive Directors over a four-year period was characterized as only somewhat exceptional by Reading Partners National staff.

[30] Colorado Department of Education Staff Statistics: Personnel Turnover Rate by District and Position Category. Available at: https://www.cde.state.co.us/cdereval/staffcurrent

Conclusions

Reading Partners has successfully navigated the start-up of its program in Colorado, establishing a strong presence within the Denver metropolitan education community. It was viewed as an easy to implement program by both tutors and school leaders. Over the first four years of the program, Reading Partners Colorado successfully grew the number of students served and the number of tutor sessions provided, while maintaining fidelity to the program model. All of this progress was attained while the program navigated a challenging fundraising context and rigorous evaluation requirements.

Over this period, Reading Partners National has evolved as an organization. APA used Stevens' *Nonprofit Lifecycle* Model as a structure to describe organizational changes to Reading Partners National and Colorado. National's movement through the growth stage into the mature stage was characterized by formalizing organizational structures and policies. This included adding professional staff and growing the role, responsibilities, and supports for regional executive directors. In particular, it has built professional capacity and created an internal policy system needed to support a mature non-profit. Its relationship with regional offices has evolved to include a more nuanced view of autonomy and authority. During the period of this study, Reading Partners Colorado was going through the growth stage with a focus on developing a presence and relationships with the local education community: building relationships with schools, tutors and funders.

Throughout this study, Reading Partners Colorado implemented the program with fidelity. This adherence to Reading Partners' program model translated to significant positive literacy outcomes for participating students, with students who received more tutoring sessions seeing even greater literacy gains. Notably, Reading Partners Colorado was especially effective for English Language Learner (ELL) students, helping them to outperform their non-ELL peers, both in and out of the program.

Looking forward, it will be important for Reading Partners to encourage program innovation so that its staff remains engaged and the organization can capitalize on new opportunities. For example:

- Reading Partners Colorado may benefit from its new freedom to include AmeriCorps members in program delivery.[31] These new team members could be engaged as Reading Center Coordinators (potentially reducing costs) or as full-time tutors who could quickly grow their literacy expertise.
- Reading Partners could pursue new uses for its strong program curriculum, such as in a summer school setting or through licensing with other programs.
- Reading Partners may find opportunities to introduce new programs and foster future organizational growth by leveraging its expertise in volunteer engagement, school partnerships, and program expansion and replication.

Pursuing opportunities such as these will help ensure that Reading Partners can continue to evolve and grow as it seeks to reach greater numbers of students across multiple states and communities.

---

[31] With the end of the five-year SIF grant, Reading Partners Colorado will no longer face restrictions on use of AmeriCorps members as Reading Center Coordinators.

References

Chiatovich, T. (January, 2012). Reading Partners Matched Pairs Evaluation Study. Stanford University, School of Education

Education, P. (2009). DRA2 K-8 Technical Manual: Developmental Reading Assessment. Upper Saddle.

Elbaum, B., Vaughn, S., Huges, M.T., & Moody, S. W. (2000). How Effective Are One-to-One Tutoring Programs in Reading for Elementary Students at Risk for Reading Failure?: A Meta-Analysis of the Intervention Research. *Journal of Educational Psychology, 92,* 4, 605-619.

Firestone, W. A., & Fisler, J. L. (2002). Politics, community, and leadership in a school-university partnership. *Educational Administration Quarterly*, *38*(4), 449-493.

Goldring, E., & Sims, P. (2005). Modeling creative and courageous school leadership through district-community-university partnerships. *Educational Policy*, *19*(1), 223-249.

Jacob, R., Armstrong, C., & Willard, J. (2015). Mobilizing Volunteer Tutors to Improve Student Literacy: Implementation, Impacts, and Costs of the Reading Partners Program.

Mathes, P., Torgesen, J., & Herron, J. (2014). Istation's Indicators of Progress (ISIP) Early Reading Technical Manual Version 4.

McCarthy, A. M., & Christ, T. J. (2010). Test Review: Beaver, JM, & Carter, MA (2006). The Developmental Reading Assessment—Second Edition (DRA2). Upper Saddle River, NJ: Pearson. *Assessment for Effective Intervention*, *35*(3), 182-185.

Rose, S. (2012). Reading/literacy. *Retrieved from Education Commission of the States: www. ecs. org*.

Slavin, R.E., Lake, C., Davis, S., & Madden, N. (2009, June) Effective programs for struggling readers: A best evidence synthesis. Baltimore, MD: Johns Hopkins University, Center for Data-Driven Reform in Education.

Slavin, R., Lake, C., Davis, S., & Madden, N. A. (2009). What works for struggling readers. *Institute for Effective Education, University of York*.

Stevens, S. K. (2001). *Nonprofit lifecycles: Stage-based wisdom for nonprofit capacity*. Stagewise Enterprises.

Appendix A: Targeted Level of Evidence

During the initial sub-grantee evaluation plan (SEP) submitted to CNCS, the study design was conceived as a randomized controlled trial. At that point, it seemed likely that Reading Partners services in schools would be significantly oversubscribed. The demand for Reading Partners services that exceeded supply would have made it possible to randomly allocate students to participate in Reading Partners, creating a randomly assigned treatment and control group. This design would have controlled for observed variables such as race and ethnicity, gender, free and reduced lunch status, and beginning literacy assessment score. It also would have controlled for unobservable characteristics such as parental support for student learning and student motivation.

However, APA and Reading Partners requested a final modification of the SEP in April 2014. At that point, it was clear that the Reading Partners program was able to accommodate all or nearly all students referred to it, meaning there was virtually no over-subscription. The low rates of over-subscription made it impossible to use random assignment ethical. It was also clear that school partners had significant concerns about denying Reading Partners services to eligible and needy students.

These practical and ethical barriers to proceeding with the planned randomized controlled trial necessitated a shift to a quasi-experimental model using propensity score matching to create a comparison group of similar students. This change was identified by multiple stakeholders as crucial to the program's continuation in Colorado. The ethical justification for random assignment is that random assignment is fair when there are more students that need services than available slots. However, the demand for services did not greatly exceed the available Reading Partner slots. Educators were opposed to allocating available services to students on a random basis. These conversations made it clear that, absent a change in research design away from an RCT, Reading Partners' relationships at several school sites was in jeopardy. Also in jeopardy was the program's ability to expand into new schools in these districts.

The design changes were made after numerous conversations between Reading Partners leadership, the Corporation for National and Community Service and its evaluation reviewers at JBS, the Butler Institute (which provides research design technical assistance to MHUW), as well as school principals, teachers, and literacy leaders from both Denver Public Schools and Aurora Public Schools. The modified quasi-experimental study design targets a moderate level of evidence within the CNCS framework.

As outlined in the body of the impact report, the study design attempts to address many threads to validity. The propensity score match includes a range of demographic variables and a fall pre-test for the spring assessment outcome variable, ensuring matched students are similar to treatment students on those observed variables. The pool of potential comparison students was also drawn from schools that are similar to treatment schools in terms of geographic location, overall academic program, proportion of Latino and Black students, and proportion of students eligible for free and reduced lunch. Drawing comparison students only from these similar schools minimizes differences attributable to school-level difference. Finally, APA's interviews with literacy leaders at both treatment and

comparison schools allowed an analysis of the similarities and differences between the baseline literacy instruction received by the two groups of students. While differences may still exist, these efforts should significantly reduce existing threats to delivery, supporting a determination of a moderate level of evidence.

Appendix B: Administrative Data Process

All administrative data used in this study was obtained from the school districts where Reading Partners was conducting operations. In 2013-14 and 2014-15, Reading partners operated in both Aurora Public Schools (APS) and Denver Public Schools (DPS). In 2015-16, Reading partners operated only in DPS. APA obtained student-level data directly from the two school districts, pursuant to a data access agreement negotiated with each individual district.

In the fall, Reading Partners provided APA with a data file on all students who had received Reading Partners services during the previous school year. This contained some program information on students, such as the date they entered and left the program and number of tutoring sessions received, as well as their district- and state-assigned student number.

APA usually obtained the data in the early winter, after the districts had time to process and analyze the assessment data from the previous school year. This meant that APA received data with the spring 2015 assessment scores in December 2015, for example. Each fall, APA would submit an official data request to the school district, including a file containing the student ID numbers and relevant program information for students who received Reading Partners tutoring during the relevant school year. Using this list of ID numbers, the district would identify the assessment and demographic information for students who received Reading Partners services. APA also requested data on all other students who attended first, second, or third grade in the district.

APA provided the district with a list of requested demographic and assessment variables. APA requested the following variables:

- Masked student IDs
- Demographic information
    - Race/ethnicity
    - Gender
    - Birth year and month (or birthdate)
    - ELL status
    - Primary Language Background
    - Free/reduced price lunch status
    - Grade level
    - School
    - Instructional Status (IEP)
- Test data
    - READ Act assessment score for fall [of relevant school year]
    - READ Act assessment score for spring [of relevant school year]

In order to maintain the security of the studentlevel data, the districts would de-identify the data before providing itto APA. This means thatstudents would be identified only by a student number, but the districtswould mask the student numberbefore providing the data to APA. APA did nothave access to any personal information for comparison students, including their names or unmasked student identification numbers. However, in order to match students across multiple school years, APA requested that the same masking formula

be applied to student identification numbers from year to year. This allowed APA to determine whether a student had previously participated in Reading Partners.

After receiving the data request, APA would work closely with the assigned data analyst at the district to ensure the receipt of relevant assessment data For example, some assessments are administered to students only once in the fall and once in the spring, while others are administered more frequently. APA worked with the analyst to ensure receipt of assessment data that had been administered at or around the same point in the school year, to ensure equivalence. APA also worked with analysts to determine which composite or subscales were needed.

After completing the data pull, the analyst would send multiple data files to APA, using secure data transfer procedures. Each year, APA would typically receive four data files from each district: demographics of comparison students; demographics of treatment students; assessment scores of comparison students; and assessment scores of treatment students. The file of demographics of treatment students would contain the program information originally provided by Reading Partners, including number of tutoring sessions received.

After receiving these multiple data files, APA would begin merging them to create a single dataset containing both assessment and demographic information for both treatment and comparison students. No weighting or recalibration was done during this process.

The only problem during this data process was a failure to obtain some data from APS for the 2014-15 students. As described in the body of the report, APA did receive a partial dataset from APS in April 2016, but it did not include fall assessment scores for students, which meant APA did not have pre-test data for APS students so could not perform propensity score matching with the APS students. This meant that no APS data could be included in the analysis for 2014-15. Despite continued contacts with APS since then, APA still has been unable to obtain a complete dataset. Therefore, the data for APS students in 2014-15 has not been included in this impact analysis.

Appendix C: Comparison Schools

Table C.1: Comparison schools

School	Comparison in 2013-14	Comparison in 2014-15	Comparison in 2015-16
Altura Elementary	yes		
Amesse Elementary	yes	yes	yes
Archuleta Elementary		yes	yes
Barnum Elementary		yes	yes
Barrett Elementary		yes	
Beach Court Elementary		yes	
Castro Elementary	yes	yes	
Columbian Elementary		yes	yes
Crawford Elementary	yes		
Doull Elementary		yes	yes
Ellis Elementary	yes	yes	yes
Fariview Elementary			yes
Farrell Howell Elementary		yes	
Godsman Elementary		yes	yes
Goldrick Elementary		yes	yes
Green Valley Elementary		yes	yes
Greenlee Elementary		yes	yes
Greenwood Elementary			yes
Harrington Elementary	yes	yes	yes
Holm Elementary		yes	yes
Johnson Elementary		yes	yes
Kaiser Elementary	yes		
Kenton Elementary	yes		
Marrama Elementary		yes	yes

School	Comparison in 2013-14	Comparison in 2014-15	Comparison in 2015-16
Maxwell Elementary	yes	yes	yes
McGlone Elementary		yes	yes
McMeen Elementary		yes	yes
Montclair Elementary	yes	yes	yes
Munroe Elementary		yes	yes
Newlon Elementary		yes	yes
Oakland Elementary		yes	yes
Palmer Elementary			yes
Schmitt Elementary	yes	yes	yes
Sixth Avenue Elementary	yes		
Stedman Elementary		yes	yes
Swansea Elementary	yes	yes	yes
University Park Elementary			yes
University Prep Elementary			yes
Valverde Elementary		yes	
Whittier Elementary		yes	
Wyatt Elementary			yes

Appendix D: Technical Details of the Propensity Score Match

For both assessments, the distribution of propensity scores across the comparison and treatment groups were very similar. Figure D.1, below, illustrates the distribution of propensity scores in the treatment and comparison groups for students who took the DRA2 assessment, while Figure D.2 illustrates the comparative distributions for students who took the iStation assessment.

Table D.1: Distribution of propensity scores for treatment and comparison students who took the DRA2 assessment
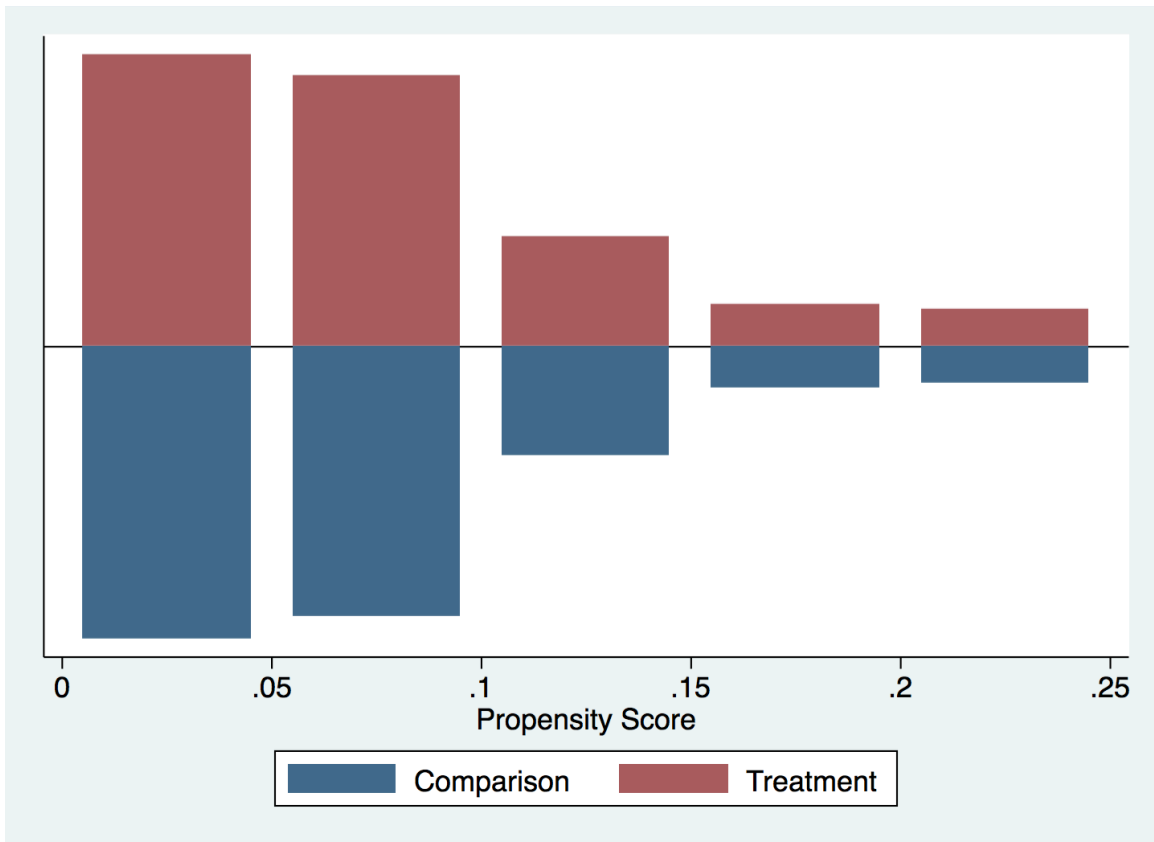
The distribution of propensity scores for comparison and treatment students who took the DRA2 assessment is nearly identical.

Table D.2: Distribution of propensity scores for comparison and treatment students who took the iStation assessment
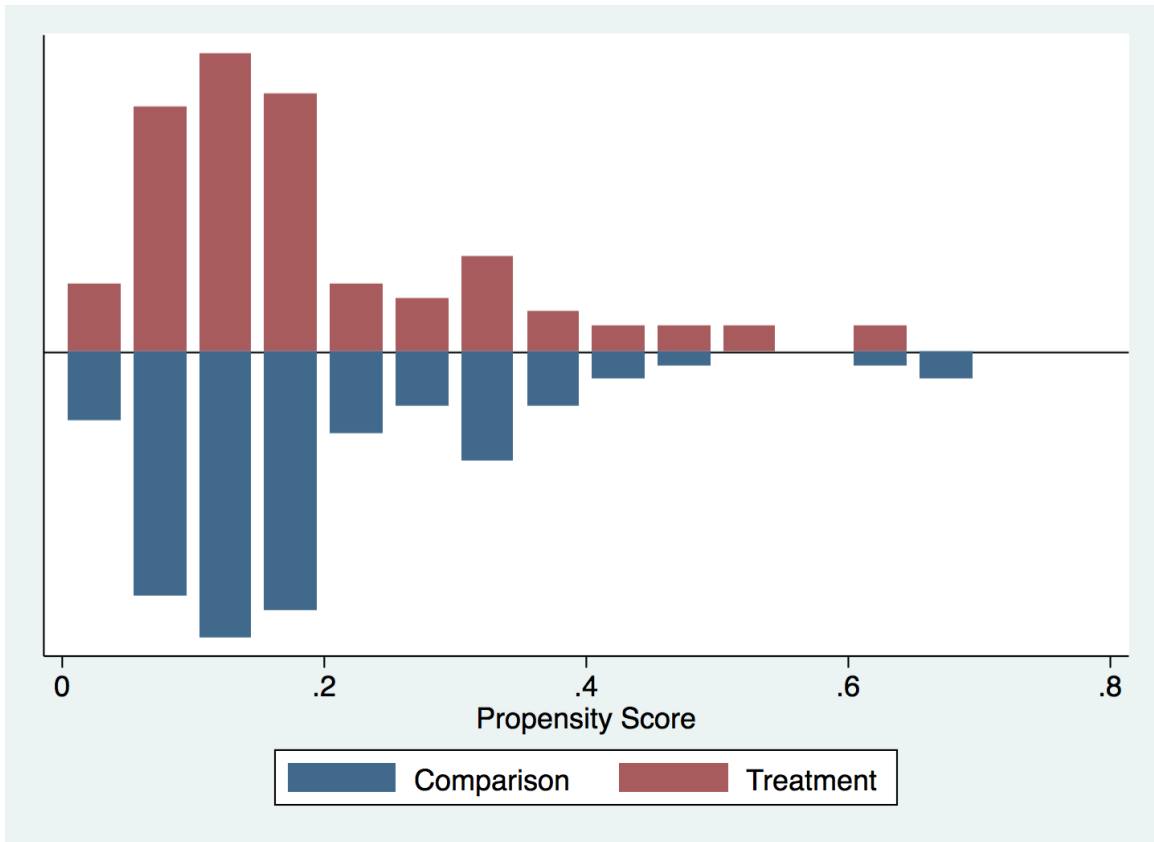
While there are some slight differences in the distribution of propensity scores between comparison and treatment students who took the iStation, they are very minor and affect few of the students.

These distributions of comparison and treatment students indicate that the groups were evenly matched after the propensity score match.

Appendix E: Model 1 Output

Table E.1, below, reports the full regression output for the first research question, evaluating the overall impact of participating in the Reading Partners program.

Table E1: Full Regression Output for Question 1 Analytic Model

	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP (first year)	4.42	1.44	3.07	0.002	1.60	7.24
FRL status	-0.57	1.39	-0.41	0.68	-3.30	2.15
ELL status	-1.00	0.91	-1.1	0.272	-2.78	0.79
SPED status	-0.97	1.14	-0.85	0.395	-3.21	1.27
Male	0.39	0.78	0.5	0.615	-1.13	1.91
Asian	-3.73	2.42	-1.54	0.124	-8.48	1.02
Other race	-1.66	2.49	-0.67	0.505	-6.55	3.22
Hispanic	-2.65	1.53	-1.73	0.084	-5.65	0.35
Black	-2.33	1.76	-1.32	0.186	-5.79	1.12
Grade 3	4.39	1.24	3.54	0	1.96	6.81
Grade 2	1.71	1.05	1.63	0.104	-0.35	3.77
iStation	33.01	2.01	16.45	0	29.08	36.94
DPS	-8.35	4.06	-2.06	0.04	-16.30	-0.39
Tested in 14-15	10.77	2.19	4.92	0	6.49	15.06
Tested in 15-16	-49.08	2.26	-21.68	0	-53.52	-44.64
Fall score NCE	0.73	0.02	29.78	0	0.68	0.78
Constant	12.72	4.00	3.18	0.001	4.87	20.57

The significant coefficients for the iStation and tested in 15-16 variables prompted APA to further investigate whether there were meaningful differences between the outcomes for students who took the iStation and those who took the DRA2. APA re-ran the model for question 1, excluding students who took the iStation and including only students who took the DRA2. Table E.2, below, reports the coefficients for the original model and the second model, which includes only DRA2 students:

Table E.2: Comparing model 1 with and without iStation scores

	Model including iStation		Model with DRA2 only	
	Coef.	P>z	Coef.	P>z
RP (first year)	4.42	0.002	5.17	0.000
FRL status	-0.57	0.680	-0.88	0.483
ELL status	-1.00	0.272	0.53	0.516
SPED status	-0.97	0.395	-3.23	0.002
Male	0.39	0.615	0.69	0.328
Asian	-3.73	0.124	-3.73	0.078
Other race	-1.66	0.505	-0.32	0.886
Hispanic	-2.65	0.084	-2.09	0.136
Black	-2.33	0.186	-2.05	0.200
Grade 3	4.39	0.000	4.42	0.000
Grade 2	1.71	0.104	2.33	0.016
iStation	33.01	0.000	-	-
DPS	-8.35	0.040	-7.93	0.004
Tested in 14-15	10.77	0.000	11.48	0.000
Tested in 15-16	-49.08	0.000	-51.93	0.000
Fall score NCE	0.73	0.000	0.73	0.000
Constant	12.72	0.001	11.12	0.000

As shown in Table E.2, the Reading Partners coefficient is still positive and significant in the DRA2-only model. In fact, the coefficient increases from 4.42 to 5.17. Otherwise, the coefficients and p-values do not change significantly between the two models. The only coefficients that are not significant in the model including iStation that become significant when excluding iStation students are special education status and Grade 2. Otherwise, both the coefficients and the significance levels are largely similar between the two models. This means that including iStation results presents a conservative estimate of the potential impact of the Reading Partners program.

Appendix F: Complete Model Output

Impact Question 1 (Confirmatory): Impact of Reading Partners on Student Reading Scores

Does Reading Partners' tutoring lead to improved near-term reading achievement for students in grades one through three when compared to similar students who do not receive tutoring?

Spring NCE Score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP 1st year	4.42	1.44	3.07	0.002	1.60	7.24
FRL	-0.57	1.39	-0.41	0.680	-3.30	2.15
ELL	-1.00	0.91	-1.10	0.272	-2.78	0.79
SPED	-0.97	1.14	-0.85	0.395	-3.21	1.27
Male	0.39	0.78	0.50	0.615	-1.13	1.91
Fall NCE score	0.73	0.02	29.78	0.000	0.68	0.78
Asian	-3.73	2.42	-1.54	0.124	-8.48	1.02
Other race	-1.66	2.49	-0.67	0.505	-6.55	3.22
Grad 3	4.39	1.24	3.54	0.000	1.96	6.81
Grad 2	1.71	1.05	1.63	0.104	-0.35	3.77
iStation	33.01	2.01	16.45	0.000	29.08	36.94
Latino	-2.65	1.53	-1.73	0.084	-5.65	0.35
Black	-2.33	1.76	-1.32	0.186	-5.79	1.12
DPS	-8.35	4.06	-2.06	0.040	-16.30	-0.39
Year 2014	10.77	2.19	4.92	0.000	6.49	15.06
Year 2015	-49.08	2.26	-21.68	0.000	-53.52	-44.64
Constant	12.72	4.00	3.18	0.001	4.87	20.57

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
schoolnumb~l: Identity				
var(_cons)	60.27	12.65	39.94	90.95
var(Residual)	158.99	6.88	146.05	173.07

Impact Question 2 (Exploratory): Impact of Reading Partners on Student Reading Scores as Dosage Increases

Do differences in reading achievement between students who receive Reading Partners tutoring and similar students who are not in Reading Partners increase as students receive more tutoring?

Model 2a: comparing Reading Partners students against themselves, excluding comparison students

Spring NCE Score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Number of sessions	-0.004	0.038	-0.09	0.926	-0.08	0.07
FRL	-0.209	2.462	-0.08	0.932	-5.03	4.62
ELL	2.179	1.493	1.46	0.144	-0.75	5.11
SPED	-0.073	1.840	-0.04	0.968	-3.68	3.53
Male	1.128	1.245	0.91	0.365	-1.31	3.57
Fall NCE score	0.669	0.053	12.71	0.000	0.57	0.77
Asian	-4.309	4.579	-0.94	0.347	-13.28	4.67
Other race	4.449	4.593	0.97	0.333	-4.55	13.45
Grad 3	6.176	2.475	2.5	0.013	1.33	11.03
Grad 2	4.411	1.826	2.42	0.016	0.83	7.99
iStation	-1.453	7.492	-0.19	0.846	-16.14	13.23
Latino	-0.992	2.911	-0.34	0.733	-6.70	4.71
Black	1.084	3.395	0.32	0.749	-5.57	7.74
DPS	-0.505	4.867	-0.1	0.917	-10.05	9.03
Year 2014	4.407	3.483	1.27	0.206	-2.42	11.23
Year 2015	-35.832	4.740	-7.56	0.000	-45.12	-26.54
Constant	11.176	5.531	2.02	0.043	0.34	22.02

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
schoolnumb~l: Identity				
var(_cons)	20.38826	8.688869	8.843485	47.00423
var(Residual)	139.64	10.1629	121.0765	161.0496

Model 2b: comparing Reading Partners students to comparison students

Spring NCE Score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Number of sessions	0.07	0.02	2.95	0.003	0.02	0.12
FRL	-0.47	1.39	-0.34	0.734	-3.20	2.25
ELL	-1.03	0.91	-1.13	0.257	-2.82	0.75
SPED	-0.61	1.16	-0.52	0.600	-2.88	1.66
Male	0.34	0.78	0.44	0.660	-1.18	1.87
Fall NCE score	0.73	0.02	29.44	0.000	0.68	0.77
Asian	-3.54	2.43	-1.46	0.145	-8.31	1.22
Other race	-1.78	2.50	-0.71	0.475	-6.68	3.11
Grad 3	4.14	1.24	3.35	0.001	1.72	6.57
Grad 2	1.75	1.05	1.66	0.097	-0.32	3.81
iStation	33.20	2.01	16.56	0.000	29.27	37.13
Latino	-2.92	1.54	-1.9	0.057	-5.94	0.09
Black	-2.56	1.77	-1.45	0.148	-6.03	0.91
DPS	-7.36	4.03	-1.82	0.068	-15.27	0.55
Year 2014	8.91	2.31	3.86	0.000	4.39	13.43
Year 2015	-50.88	2.32	-21.96	0.000	-55.42	-46.34
Constant	13.82	3.92	3.52	0.000	6.13	21.51

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
schoolnumb~l: Identity				
var(_cons)	57.73	12.29	38.04	87.62
var(Residual)	159.56	6.91	146.58	173.70

Impact Question 3 (Exploratory): Differences in Reading Partners Effects on Different Student Groups

Are there differential impacts of Reading Partners tutoring on different student groups including English-language learners (ELL) vs. non-ELL students, boys vs. girls, grade level, and different races?

Model 3a: Interaction of Reading Partners and ELL status

Spring NCE score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP & ELL	3.44	1.69	2.04	0.042	0.13	6.74
RP	2.07	1.72	1.2	0.231	-1.31	5.44
FRL	-0.59	1.39	-0.42	0.673	-3.31	2.14
ELL	-2.27	1.09	-2.08	0.037	-4.40	-0.14
SPED	-0.74	1.15	-0.65	0.518	-3.00	1.51
Male	0.40	0.78	0.52	0.605	-1.12	1.93
Fall NCE score	0.73	0.02	29.69	0.000	0.68	0.78
Asian	-3.48	2.43	-1.43	0.152	-8.24	1.28
Other race	-1.75	2.49	-0.7	0.484	-6.64	3.14
Grad 3	4.25	1.24	3.44	0.001	1.83	6.67
Grad 2	1.73	1.05	1.64	0.100	-0.33	3.79
iStation	33.35	2.02	16.54	0.000	29.40	37.30
Latino	-2.74	1.53	-1.79	0.074	-5.74	0.27
Black	-2.33	1.76	-1.32	0.188	-5.78	1.13
DPS	-8.67	4.04	-2.15	0.032	-16.58	-0.76
Year 2014	10.75	2.19	4.9	0.000	6.44	15.05
Year 2015	-49.14	2.26	-21.76	0.000	-53.57	-44.71
Constant	13.69	3.99	3.43	0.001	5.86	21.52

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
schoolnumb~l: Identity				
var(_cons)	59.28	12.61	39.07	89.95
var(Residual)	159.05	6.89	146.10	173.15

Model 3b: interaction of Reading Partners and Special Education status

Spring NCE score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP & SPED	1.28	2.33	0.55	0.583	-3.29	5.86
RP	3.51	1.57	2.23	0.026	0.43	6.60
FRL	-0.57	1.39	-0.41	0.685	-3.30	2.16
ELL	-1.05	0.91	-1.15	0.249	-2.84	0.74
SPED	-1.23	1.42	-0.87	0.384	-4.01	1.54
Male	0.36	0.78	0.46	0.647	-1.17	1.88
Fall NCE score	0.73	0.02	29.7	0.000	0.68	0.78
Asian	-3.66	2.43	-1.51	0.132	-8.43	1.11
Other race	-1.66	2.50	-0.66	0.506	-6.56	3.24
Grad 3	4.25	1.24	3.43	0.001	1.82	6.67
Grad 2	1.75	1.05	1.67	0.096	-0.31	3.82
iStation	33.10	2.01	16.43	0.000	29.15	37.05
Latino	-2.73	1.54	-1.78	0.076	-5.74	0.28
Black	-2.43	1.77	-1.37	0.170	-5.89	1.04
DPS	-8.52	4.04	-2.11	0.035	-16.43	-0.61
Year 2014	10.67	2.20	4.85	0.000	6.36	14.99
Year 2015	-49.07	2.26	-21.7	0.000	-53.50	-44.64
Constant	13.10	3.99	3.29	0.001	5.29	20.91

Model 3c: Interaction of Reading Partners and Gender

Spring NCE score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP & Male	2.13	1.61	1.32	0.186	-1.02	5.27
RP	2.70	1.71	1.58	0.113	-0.64	6.05
FRL	-0.58	1.39	-0.42	0.674	-3.31	2.14
ELL	-1.03	0.91	-1.13	0.260	-2.81	0.76
SPED	-0.74	1.15	-0.64	0.521	-3.00	1.52
Male	-0.39	0.96	-0.41	0.683	-2.27	1.49
Fall NCE score	0.73	0.02	29.69	0.000	0.68	0.78
Asian	-3.61	2.43	-1.48	0.138	-8.37	1.16
Other race	-1.59	2.50	-0.64	0.525	-6.48	3.31
Grad 3	4.33	1.24	3.5	0.000	1.91	6.76
Grad 2	1.85	1.05	1.75	0.080	-0.22	3.91
iStation	33.09	2.01	16.43	0.000	29.14	37.04
Latino	-2.71	1.53	-1.76	0.078	-5.71	0.30
Black	-2.36	1.77	-1.33	0.182	-5.82	1.11
DPS	-8.48	4.04	-2.1	0.036	-16.41	-0.55
Year 2014	10.52	2.20	4.79	0.000	6.21	14.83
Year 2015	-49.08	2.26	-21.7	0.000	-53.51	-44.64
Constant	13.39	4.00	3.35	0.001	5.55	21.23

Model 3d: interaction of Reading Partners and Grade 1

Spring NCE score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP & Grade 1	0.07	2.03	0.03	0.972	-3.90	4.04
RP	3.73	1.55	2.41	0.016	0.70	6.77
FRL	-0.53	1.39	-0.38	0.702	-3.26	2.20
ELL	-1.05	0.91	-1.15	0.249	-2.84	0.74
SPED	-0.78	1.15	-0.68	0.500	-3.04	1.48
Male	0.35	0.78	0.45	0.653	-1.18	1.88
Fall NCE score	0.73	0.02	29.65	0.000	0.68	0.78
Asian	-3.61	2.43	-1.48	0.139	-8.38	1.17
Other race	-1.68	2.50	-0.67	0.503	-6.58	3.22
Grad 3	4.27	1.36	3.15	0.002	1.61	6.93
Grad 2	1.78	1.23	1.45	0.148	-0.63	4.20
iStation	33.10	2.02	16.39	0.000	29.14	37.06
Latino	-2.73	1.54	-1.78	0.075	-5.74	0.28
Black	-2.40	1.77	-1.36	0.175	-5.87	1.07
DPS	-8.50	4.05	-2.1	0.036	-16.44	-0.57
Year 2014	10.57	2.21	4.79	0.000	6.25	14.90
Year 2015	-49.09	2.27	-21.62	0.000	-53.55	-44.64
Constant	13.02	4.03	3.23	0.001	5.12	20.93

Model 3e: interaction of Reading Partners and Grade 2

Spring NCE score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP & Grade 2	-0.15	1.72	-0.09	0.928	-3.52	3.21
RP	3.80	1.66	2.29	0.022	0.55	7.06
FRL	-0.53	1.39	-0.38	0.703	-3.26	2.20
ELL	-1.05	0.91	-1.15	0.249	-2.84	0.74
SPED	-0.78	1.15	-0.68	0.499	-3.04	1.48
Male	0.35	0.78	0.45	0.654	-1.18	1.88
Fall NCE score	0.73	0.02	29.69	0.000	0.68	0.78
Asian	-3.60	2.43	-1.48	0.139	-8.37	1.17
Other race	-1.67	2.50	-0.67	0.505	-6.57	3.24
Grad 3	4.25	1.24	3.43	0.001	1.83	6.68
Grad 2	1.82	1.21	1.5	0.134	-0.56	4.19
iStation	33.09	2.02	16.42	0.000	29.14	37.04
Latino	-2.73	1.54	-1.78	0.075	-5.74	0.28
Black	-2.40	1.77	-1.36	0.175	-5.87	1.07
DPS	-8.52	4.04	-2.11	0.035	-16.44	-0.60
Year 2014	10.59	2.20	4.82	0.000	6.28	14.90
Year 2015	-49.08	2.26	-21.68	0.000	-53.51	-44.64
Constant	13.02	4.00	3.26	0.001	5.19	20.85

Model 3f: interaction of Reading Partners and Grade 3

Spring NCE score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP & Grade 3	0.12	1.85	0.07	0.948	-3.51	3.75
RP	3.69	1.73	2.13	0.033	0.30	7.08
FRL	-0.53	1.39	-0.38	0.705	-3.26	2.20
ELL	-1.06	0.91	-1.16	0.247	-2.85	0.73
SPED	-0.78	1.15	-0.68	0.499	-3.04	1.48
Male	0.35	0.78	0.45	0.651	-1.17	1.88
Fall NCE score	0.73	0.02	29.61	0.000	0.68	0.78
Asian	-3.61	2.43	-1.48	0.138	-8.38	1.16
Other race	-1.67	2.50	-0.67	0.503	-6.57	3.22
Grad 3	4.22	1.33	3.17	0.002	1.61	6.83
Grad 2	1.77	1.06	1.67	0.094	-0.30	3.83
iStation	33.08	2.02	16.39	0.000	29.13	37.04
Latino	-2.73	1.54	-1.78	0.075	-5.74	0.28
Black	-2.41	1.77	-1.36	0.173	-5.87	1.06
DPS	-8.53	4.05	-2.11	0.035	-16.48	-0.59
Year 2014	10.60	2.22	4.78	0.000	6.26	14.95
Year 2015	-49.07	2.28	-21.53	0.000	-53.53	-44.60
Constant	13.06	4.00	3.27	0.001	5.22	20.90

Impact Question 4 (Exploratory): Impact of Modeling from Multiple Years of Reading Partners Treatment

How do the differences or similarities in the results using the impact and exploratory samples impact judgments about Reading Partners impact on near-term reading achievement?

Spring NCE score	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
RP 1 year	4.85	1.48	3.28	0.001	1.95	7.75
RP 2 years	2.89	2.05	1.41	0.159	-1.13	6.90
RP 3 years	20.42	5.85	3.49	0.000	8.95	31.88
FRL	-0.45	1.38	-0.33	0.745	-3.16	2.26
ELL	-0.90	0.91	-0.99	0.321	-2.68	0.88
SPED	-0.58	1.15	-0.5	0.614	-2.84	1.68
Male	0.24	0.78	0.31	0.753	-1.28	1.77
Fall NCE score	0.72	0.02	28.94	0.000	0.67	0.77
Asian	-3.49	2.42	-1.44	0.149	-8.23	1.25
Other race	-1.80	2.48	-0.72	0.470	-6.66	3.07
Grad 3	4.23	1.23	3.43	0.001	1.81	6.65
Grad 2	1.63	1.05	1.55	0.122	-0.43	3.69
iStation	33.87	2.00	16.96	0.000	29.95	37.78
Latino	-2.95	1.53	-1.93	0.054	-5.94	0.05
Black	-2.50	1.76	-1.42	0.155	-5.94	0.95
DPS	-5.82	4.01	-1.45	0.147	-13.68	2.04
Year 2014	7.60	2.43	3.13	0.002	2.84	12.37
Year 2015	-52.15	2.43	-21.44	0.000	-56.92	-47.39
Constant	13.43	3.89	3.46	0.001	5.81	21.04

Appendix G: Data Collection Activities and Respondents

This appendix provides additional detail on the data collection instruments and processes. This information is provided by year.

As part of the planning for the evaluation APA identified the key components of developing a new organization in Colorado (e.g. staffing and school site selection) and implementing the for the first time in Colorado schools (e.g. identify students and tutors, provide tutoring and adjust tutoring to meet student needs). The process model served to organize both the 2012-13 evaluation questions and the development of the instruments. During this first-year APA developed interview protocols, tutor surveys and an observation protocol. The development process included reviews by Reading Partners National staff and the instruments were reviewed by an IRB. Our data collection activities were:

- Review of program documentation including Reading Center and Site Coordinator checklists, student folder materials, and Program Manager training materials.
- Analysis of Reading Partners administrative data on tutors and students
- Five interviews with Reading Partners' National staff
- Two interviews with Reading Partners Colorado staff
- Site visits to all seven Reading Partners sites
- Observations of 14 Reading Partners tutoring sessions
- Interviews with seven Site Coordinators
- Interviews with four points of contact (school staff) at Reading Partners Sites
- An electronic survey of tutors (134 respondents)

The 2013-14 implementation monitoring was limited to seven study sites with limited data collection. No additional instruments were developed. Instead, the existing instruments were used. The data collection activities during 2013-14 were:

- Analysis of Reading Partners administrative data on tutors and students
- Site visits to seven Reading Partners sites
- Observation of a tutoring session in each Reading Partners site
- Interviews with Site Coordinators in seven schools
- Interviews with literacy leaders at five Reading Partners schools and four comparison schools

In 2014-15 the implementation study was expanded to meet Reading Partners' needs. New folder review protocols were developed drawing upon protocols used by MDRC. New site coordinator, and regional staff protocols were developed in consultation with Reading Partners.

The folder reviews and site coordinator interviews were conducted in both the fall and repeated in the spring semesters. APA slightly revised the folder review and interview protocols for the spring data collection after reviewing initial findings with Reading Partner staff. Folder data was collected from 10 study sites in the fall. A school was added mid-year so folder data was collected from 11 sites in the spring. In the spring and the fall, approximately 20 folders were randomly selected at each school for review of the four most

recent tutoring sessions. Reading Partner also allowed APA to insert several questions into their regular tutor survey to learn about experiences with Reading Partners and their plans for future participation. The data collection activities in 2014-15 where:

- Analysis of Reading Partners administrative data on tutors and students
- Fall review of 228 randomly selected folders
- Spring review of 224 folders
- Fall and spring interviews with all site coordinators, the executive director, program managers, and program associate.
- Fall interviews with eight school literacy leaders and spring interviews with nine school literacy leaders
- Tutor survey data (484 respondents)

The 2015-6 implementation study was modified based on information gathered from the prior year. Folder review protocols were slightly revised. The sampling strategy was refined to all tutoring sessions from 15 students randomly selected at each school. An on-line survey protocol for school literacy leaders was developed to collect data similar to the interview data collected in prior years. As was done in 2014-15 Reading Partners provided tutor survey data.

APA, in coordination with Reading Partners, developed new survey and interview protocols to describe practices at the sites. APA used a two-step process of collecting data from Program Managers and Associates about the sites they supervised. Detailed survey data was collected from the Program Managers/Associates about each of the schools they supervise. Then follow-up interviews were conducted to add detail to information collected through those surveys. Reading Center Coordinators also completed on-line surveys.

APA also interviewed the Executive Director and Community engagement manager using protocols developed in the prior year with questions added to explore themes from the survey data. Data collection activities during 2015-16 were:

- Review of 210 student folders
- School literacy leader surveys (86% response rate)
- Program Manager and Associate on-line surveys (100% response rate)
- Site Coordinator Surveys (100% response rate)
- Interviews with each Program Manager and Associate
- Tutor survey data (446 respondents)

The 2015-16 data collection was very limited. In consultation with Reading Partners APA developed an interview protocol for Reading Partners Colorado and National staff as well as MHUW staff.

- Interviews with six Reading Partners staff
- An interview with two MHUW staff