

FINAL REPORT

Outcome Evaluation of the Florida Reading Corps PreK Program

JANUARY 2019

PRESENTED TO:
Reading & Math, Inc.
2400 Park Ave. S
Minneapolis, MN 55404

PRESENTED BY:
NORC at the University of Chicago
1155 E 60th Street
Chicago, IL 60637

AUTHORS:
Carrie E. Markovitz
Marc W. Hernandez
Eric C. Hedberg
Jennifer L. Satorius
Julie Kubelka

Acknowledgement

Many individuals and organizations contributed to the design and implementation of the *Outcome Evaluation of the Florida Reading Corps PreK Program*. While it is not possible to name everyone, we would like to acknowledge some of the individuals and organizations who played a significant role in the successful execution of this study. We would like to begin by extending a special thank you to the select evaluation sites, their staff, and the Reading Corps tutors serving in the Florida Reading Corps program for agreeing to participate in evaluation. The comparison sites deserve extra recognition; this evaluation was possible only because their administrators and teachers were willing to trust our field assessors to enter their classrooms and collect assessment data from their students several times during the school year. We would like to acknowledge the contributions of the study site liaisons, Kelly Holt, Julie Kubelka, and Jennifer Scolese, who coordinated and managed the data collection at individual PreK sites, as well as Heidi Whitmore who managed the data collection. We would also like to recognize the contributions of the field assessors who collected assessment data from the students at the PreK sites. Finally, the outcome evaluation would not have been possible without the assistance of the Master Coaches, Internal Coaches, and staff of the Florida Reading Corps program, and without the strong support of Jerry Calvert and Sadie O'Connor from Reading & Math, Inc.; Karin Charles, David Parker, and Patrick Kaiser from ServeMinnesota; and Lindsay Dolce from Reading & Math Foundation.

Table of Contents

Acknowledgement	i
Executive Summary	iv
About Reading Corps.....	iv
Study Methodology	vii
Findings and Conclusions	viii
Concluding Thoughts	xii
Introduction	1
The Reading Corps Program Model	2
Response to Intervention Framework.....	2
PreK Program Literacy Focus	3
Use of Data in Program Implementation	5
Coaching and Support	6
Training.....	7
Florida Reading Corps	7
About the Study	8
Evaluation Logic Model	8
Research Questions	9
Selection and Matching of Schools	9
Data Sources	12
Analysis Methods.....	13
Study Limitations.....	14
Site Visits	15
Study Findings	17
What is the effect of Florida Reading Corps on PreK students?	17
Does program effect vary by participant characteristics, such as demographics?	22
Conclusions and Discussion	23
Discussion	23
Appendix A: Logic Model	27
Appendix B: Comparison Site Matching Process	28
Appendix C: Detailed Analysis Methods	29
Appendix D: Results Tables	33
Appendix E: Baseline Balance	34

List of Figures/Exhibits

Figure i.1.	Mean Letter Name Fluency Scores for PreK program and comparison students ...	ix
Figure i.2.	Mean Letter Sound Fluency Scores for PreK program and comparison students....	x
Figure i.3.	Mean Rhyming Fluency Scores for PreK program and comparison students	x
Figure i.4.	Mean Alliteration Fluency Scores for PreK program and comparison students	xi
Figure i.5.	Mean Picture Name Fluency Scores for PreK program and comparison students .	xi
Figure 1.	Response to Intervention – Three Tier Model	3
Figure 2.	Mean Letter Name Fluency Scores for PreK program and comparison students ..	18
Figure 3.	Mean Letter Sound Fluency Scores for PreK program and comparison students..	19
Figure 4.	Mean Rhyming Fluency Scores for PreK program and comparison students	20
Figure 5.	Mean Alliteration Fluency Scores for PreK program and comparison students	21
Figure 6.	Mean Picture Name Fluency Scores for PreK program and comparison students	22
Exhibit A.1.	Reading Corps Logic Model (for Outcome Evaluation).....	27
Figure C.1.	The difference between the Latent Ability and Measured distributions (hypothetical data)	31

List of Tables

Table 1.	Spring PreK Benchmarks (4/5 Year-Old Children)	6
Table 2.	PreK QED Matching Variables	10
Table 3.	Characteristics of Schools Participating in the Florida Reading Corps Outcome Evaluation (Fall 2017)	11
Table 4.	Student participants for the Florida Reading Corps Outcome Evaluation (Fall 2017)	12
Table C.1.	Results of Simulation	32
Table D.1.	Florida PreK Reading Competency Results	33
Table E.1.	Florida PreK Fall Baseline Equivalence of Analysis Sample: Reading Competency Results.....	35

Executive Summary

Reading Corps is a multi-state AmeriCorps program that supports reading development for students in Pre-Kindergarten (Pre-K) through grade 3, with the ultimate goal of ensuring that students become successful readers and meet reading proficiency targets by the end of the third grade. The program's core activities consist of training, placing, and supporting literacy tutors in school-based settings to implement research-based, early-literacy interventions. Reading Corps provides multiple layers of supervisory support to ensure integrity of program implementation. Site-specific "Internal Coaches" serve as supervisors, mentors, and advocates for tutors and external "Master Coaches" are literacy experts who provide expert consultation on literacy instruction for tutors and Internal Coaches.

As part of an Innovative Approaches to Literacy (IAL) Grant through the U.S. Department of Education (ED), a multi-site impact evaluation of the Reading Corps program was conducted in three states during the 2017-2018 school year. This evaluation builds on the results from prior independent evaluations of the *original* Minnesota Reading Corps program by assessing the effect of the Reading Corps program model in three sites: Minnesota, as well as two replication sites, Florida and Wisconsin. This report presents the findings from the evaluation of the Florida Reading Corps PreK Program, describing its effect on 630 four- and five-year old PreK students who attended either a Florida Reading Corps PreK program site or a matched-comparison site during the 2017-2018 school year.

As the focus of this report, the Florida Reading Corps program is a replication of the successful Minnesota Reading Corps PreK model. At the time this evaluation was conducted, the Florida Reading Corps program was in its third year of operation with 35 Reading Corps tutors in 12 schools in Miami, Florida.

About Reading Corps

Reading Corps is the largest AmeriCorps tutoring program in the country, providing trained literacy tutors to implement evidence-based literacy instruction and assessment protocols for students in PreK through grade 3 in school-based settings. Beginning in 2003 with just 250 students in Minnesota, Reading Corps has since expanded to 12 states and Washington D.C. Reading Corps aims to broadly impact literacy outcomes for children. A primary goal for the program is that all children, age 3 to grade 3, who qualify for Reading Corps, will meet reading proficiency standards by third grade.

Literacy Focused Response to Intervention Framework

Reading Corps uses a Response to Intervention (RtI) framework that employs a multi-tier approach to the early identification and provision of support to struggling readers. Key features of the Reading Corps RtI framework include:

- Clear literacy targets at each age level from PreK through grade 3
- Benchmark assessment data to identify students eligible for one-on-one interventions

- Evidence-based interventions
- Frequent progress monitoring during intervention delivery
- High-quality training, coaching, and observations to support fidelity of implementation

The Reading Corps PreK program includes an immersive component, where tutors provide whole-class emergent literacy enrichment for all students using evidence-based practices assessed in the ELLCO (Early Language & Literacy Classroom Observation) tool, as well as a targeted small group and one-to-one component, where tutors provide more individualized interventions to students struggling with emergent literacy skills. During these small-group and one-to-one sessions, tutors implement 5-to 15-minute scripted interventions focused on specific emergent literacy skills tailored to students' current proficiency levels.

PreK Program Literacy Focus

The Reading Corps PreK program focuses on integrating the “Big Five” Early Literacy Predictors outlined by the SEEDS of Early Learning curriculum. These five categories summarize the work of the National Early Literacy Panel,¹ and are incorporated into all aspects of the daily classroom routine. The “Big Five” for preschool students include: 1) oral language, conversation, and comprehension, 2) vocabulary and meaning, 3) concepts about books and print, 4) phonological memory and awareness (i.e., rhyming and alliteration), and 5) alphabet knowledge (i.e., letter name recognition and letter sound correspondence). Reading Corps tutors are tasked with creating a Literacy Rich Classroom using evidence-based practices assessed in the ELLCO (Early Language & Literacy Classroom Observation) tool, and implementing a Literacy and Math Rich Schedule. The Reading Corps Literacy and Math Rich Schedule is designed to provide children with daily routines that embed early literacy predictors into fun and meaningful learning.

Reading Corps tutors serving in the PreK program are also responsible for enacting the SEEDS of Early Learning,² which provides both tutors and teachers with specific strategies to enhance literacy instruction for *all children* in the classroom. SEEDS is a relationship-based instructional approach that maps out for teachers five ways to *intentionally* interact with children in order to promote academic growth and social-emotional development. SEEDS high quality teacher-student interactions include the following five elements: sensitivity, encouragement, education, development of skills through doing, and self-image support.

Use of Data in Program Implementation

Assessment data play a key role in the implementation of Reading Corps. In the Reading Corps PreK program, tutors collect two types of assessment data on all PreK students in a classroom in the Fall, Winter, and Spring of each school year. Specifically, three age-appropriate Individual Growth and

¹ National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.

² Horst, K., & Passe, A. (2004). *Creating Literacy Rich Classrooms for Preschool Children (Ages 0-5)*. Presented at the 2004 CEED Symposium. Minneapolis, MN.; <https://www.seeds-learning.com/>

Development Indicators (IGDIs) and two Fastbridge (FAST) assessments are used to measure the five critical emergent literacy skills listed below:

1. **Letter Name Fluency** evaluates students' alphabet knowledge by assessing how many upper case letter names a student can identify within one minute.
2. **Letter Sound Fluency** assesses how many letter sounds a student can identify and say (e.g., l, f, m) within one minute and is considered a more difficult/advanced measure of a child's alphabet knowledge.
3. **Rhyming Fluency** tests students' phonological awareness by assessing how many rhyming matches a student can correctly identify (e.g., bat and cat, car and jar, ham and jam) within two minutes.
4. **Alliteration Fluency** tests students' phonological awareness by assessing how many alliteration matches (e.g., hill and heart, cat and can, bike and ball) a child can correctly identify within two minutes.
5. **Picture Name Fluency** tests students' vocabulary by assessing how many pictures a child can name correctly within one minute.

Students' scores on these five measures are used to identify students eligible for receiving small-group and one-to-one interventions with Reading Corps tutors, tracking their progress toward achieving academic goals related to critical literacy skills, and informing instruction of the whole class, immersive component of the program. The data also provide tutors and coaches with objective information about the efficacy of the interventions with individual students, allowing them to tailor their instruction so that it is most effective for the student's skill level.

Coaching, Support, and Training

Reading Corps provides multiple layers of supervision, training, and support to ensure integrity of program implementation. Internal Coaches serve as a front-line supervisor and a source of support for tutors in the implementation of the Reading Corps' PreK program in participating classrooms. Master Coaches are literacy experts who provide Internal Coaches and Reading Corps tutors with expert consultation on literacy instruction, while also ensuring that the Reading Corps program elements are implemented with integrity. Along with these two coaching layers, a third layer consisting of Reading Corps program support staff (i.e., Reading Corps employees) provides administrative oversight for program implementation to sites participating in Reading Corps. Lastly, each Reading Corps site utilizes an on-site administrative liaison, who provides general oversight to tutors at the school (e.g., attendance tracking).

Prior to the start of each school year, Reading Corps hosts a three-day Institute to train returning and new Master Coaches, Internal Coaches, and Reading Corps tutors in the assessments and research-based emergent literacy interventions employed by Reading Corps. This intensive, information-rich training provides tutors with the skills, knowledge, and tools needed to implement the program in PreK classrooms. Ongoing training and coaching sessions are also provided throughout the tutors' year of

service. Additionally, classroom teachers are provided training on the Reading Corps program in order to assist Reading Corps tutors with program implementation in their classroom.

Study Methodology

The primary goal of the Florida Reading Corps PreK program outcome evaluation was to independently and quasi-experimentally assess the effect of the Florida Reading Corps PreK program on 4- and 5-year old students' emergent literacy skills. Thus, the primary research questions for the evaluation are as follows:

1. What is the effect of the Florida Reading Corps PreK program on program participants compared to students at similar schools who did not receive Reading Corps?
2. Does program effect vary by participant characteristics, such as demographics?

To answer these research questions, the evaluation team analyzed age-appropriate and semester-specific emergent literacy assessment scores collected from 630 PreK students enrolled at 24 preschools in Miami-Dade County Public Schools (M-DCPS) and YWCA PreK Centers in Florida during the 2017-2018 school year.

Selection and Matching of Schools

As the Florida Reading Corps PreK program is a whole-class/whole-site intervention, the evaluation had to employ a quasi-experimental design that used a multi-staged matching process to identify a set of comparison schools. All 12 Florida Reading Corps PreK program sites operating across M-DCPS and YWCA in 2016-2017 were selected to participate in the PreK evaluation during the 2017-2018 school year (i.e., program sites).³ The remaining 12 sites (i.e., comparison sites) were selected for participation in the study based on a prescribed set of educationally significant school-level characteristics, which matched those of the Florida Reading Corps program sites: 1) location/neighborhood; 2) program type (public PreK, community-based PreK, Head Start); 3) ages served (4 and 5 year olds); 4) Socio-Economic Status (SES) as measured by percentage of students eligible for Free or Reduced Price Lunch; 5) percentage of students who are Dual Language Learners (DLL); 6) total site enrollment; and 7) racial/ethnic distribution of PreK students. The final sample comprised 24 sites (12 Florida Reading Corps and 12 comparison sites) and included 630 students.

Data Sources

The evaluation's primary sources of data for measuring PreK student emergent literacy outcomes were three IGDI and two FAST assessments collected from PreK students at the 12 Florida Reading Corps sites and 12 comparison sites. While Reading Corps tutors collected the Fall benchmark assessment data at the 12 Reading Corps sites, trained independent assessors from the evaluation team collected the Spring benchmark assessment data from PreK students at these sites to address any potential bias at posttest. The

³ Two additional M-DCPS Florida Reading Corps sites were just starting their first year of implementation during the 2017-2018 school year and, therefore, did not participate in the study.

independent assessors collected both the Fall and Spring benchmark assessment data from PreK students at the 12 comparison sites using the exact same assessments and tools as the Florida Reading Corps PreK program staff.

Lastly, program staff provided the evaluation team with other administrative data on Florida Reading Corps program students, including tutoring attendance, average length and number of sessions, and demographic characteristics. The evaluation team obtained similar demographic data on comparison students from M-DCPS and the YWCA.

Analysis Methods

To estimate program effects, the average spring score of students enrolled at sites that participated in the Florida Reading Corps PreK program was statistically compared with the average spring score of a comparison group of students at similar sites that did not participate. The calculated difference in group averages is the plausible effect of program participation.

The outcome evaluation analyses were conducted using mixed-regression models, also known as hierarchical linear models. This type of model uses maximum likelihood to estimate differences in outcomes associated with predictor variables while also modeling the multi-level structure of the data, which in this case are students nested within sites. The analytic approach accounted for the effect of program status (i.e., Reading Corps program or comparison group) in isolation as well as when relevant control variables (i.e., student demographic characteristics and site characteristics) were included.

Findings and Conclusions

Below, the evaluation team presents the study findings, followed by final thoughts on the implications of these findings for the future of the Florida Reading Corps program.

What is the effect of Florida Reading Corps on PreK students?

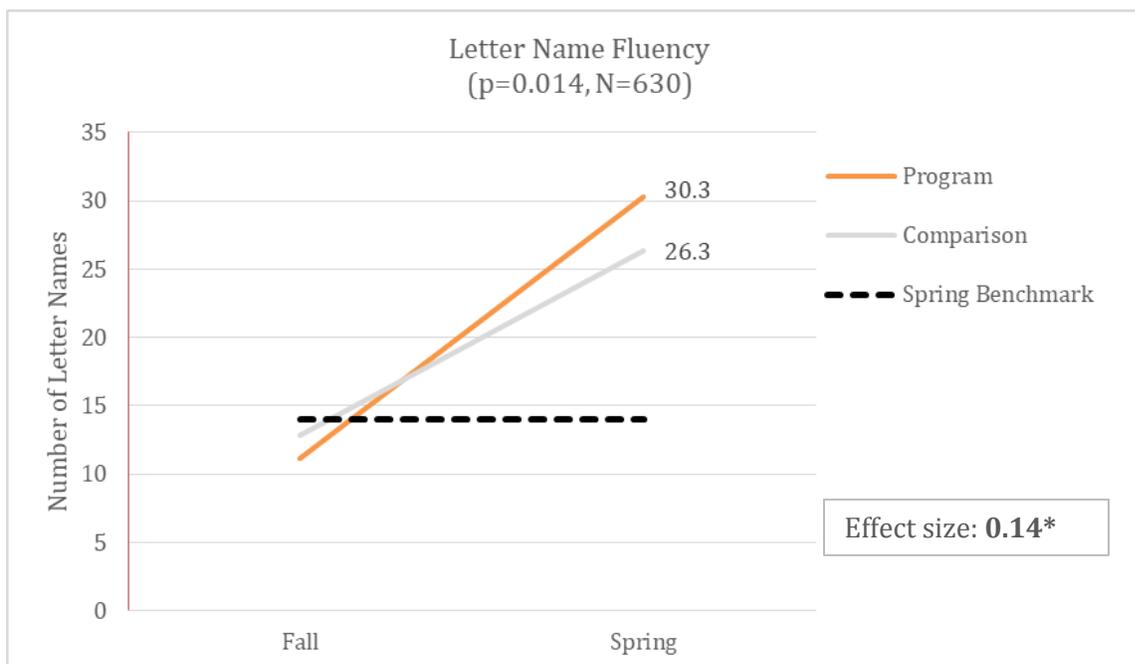
To answer this question, the analysis examined spring 2018 differences between students who participated in the Florida Reading Corps PreK program and comparable students at comparison schools who did not, on the following five emergent literacy skills: letter name identification, letter sound correspondence, rhyming, alliteration, and picture naming. Statistically significant differences between the Florida Reading Corps and comparison groups were observed for four of the five measures, and the students' average score exceeded the Spring benchmark on two of the five measures listed below.

1. **Letter Name Fluency:** Program students correctly identified 4.0 more letter names on average than students in the comparison group ($p=0.014$, $N=630$). This represents an effect size of 0.14, which is small considering the typical year-over-year growth of students this age (**Figure i.1**). Both groups on average exceeded the end-of-year benchmark for this measure.
2. **Letter Sound Fluency:** On average, program students correctly identified 3.2 more letter sounds than students in the comparison group ($p=0.027$, $N=614$). This represents an effect size of 0.11,

which is small considering the typical year-over-year growth of students at this age (**Figure i.2**). Both groups on average exceeded the end-of-year benchmark for this measure.

3. **Rhyming Fluency:** On average, program students correctly identified 1.2 more rhyming words than students in the comparison group ($p=0.001$, $N=630$).⁴ This represents an effect size of 0.30, which is small considering the typical year-over-year growth of students at this age (**Figure i.3**).
4. **Alliteration Fluency:** On average, students in the program group correctly identified 1.1 more alliterative words than students in the comparison group ($p < 0.001$, $N=615$). This difference represents an effect size of 0.22, which is small given the typical year-over-year growth of students this age (**Figure i.4**).
5. **Picture Name Fluency:** There were no statistically significant differences between study groups for picture name fluency (**Figure i.5**).

Figure i.1. Mean Letter Name Fluency Scores for PreK program and comparison students



⁴ Readers will note that there is a similar difference between Program and Comparison students at the Fall baseline in Figure 4. However, our test of statistical significance includes the Fall baseline score as a covariate (in addition to demographic characteristics) to remove this source of variation.

Figure i.2. Mean Letter Sound Fluency Scores for PreK program and comparison students

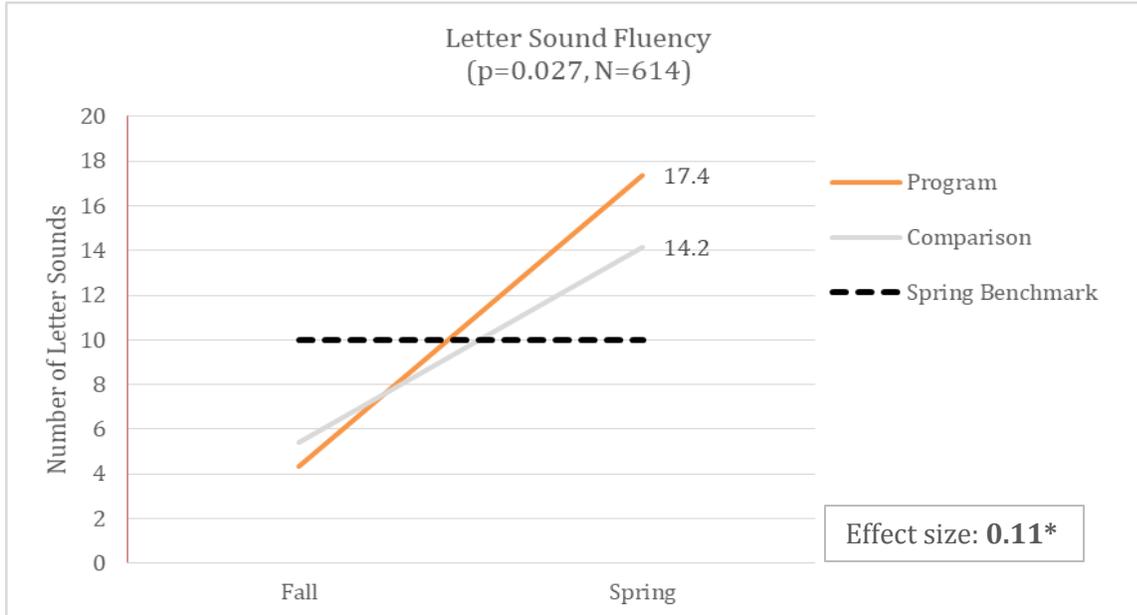


Figure i.3. Mean Rhyming Fluency Scores for PreK program and comparison students

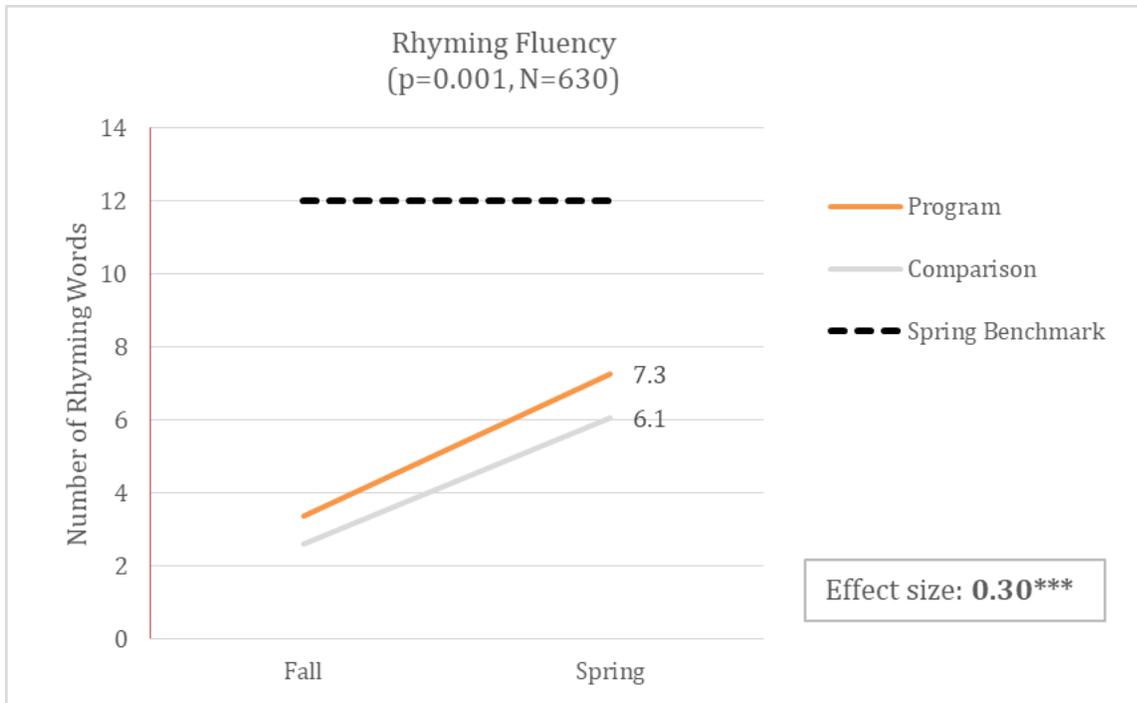


Figure i.4. Mean Alliteration Fluency Scores for PreK program and comparison students

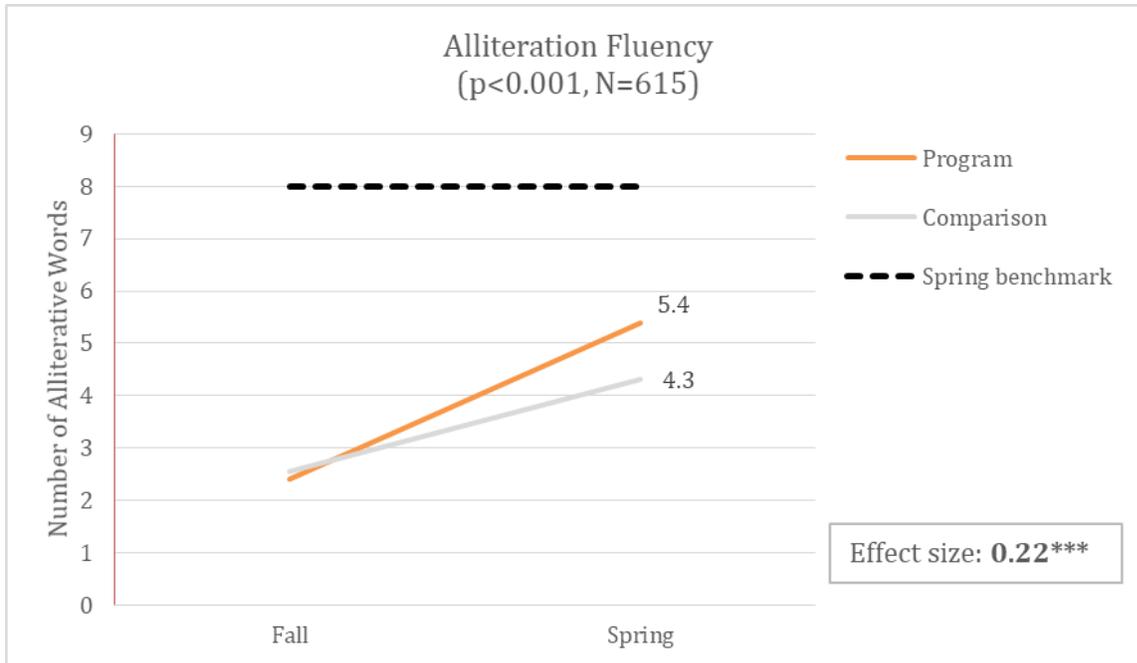
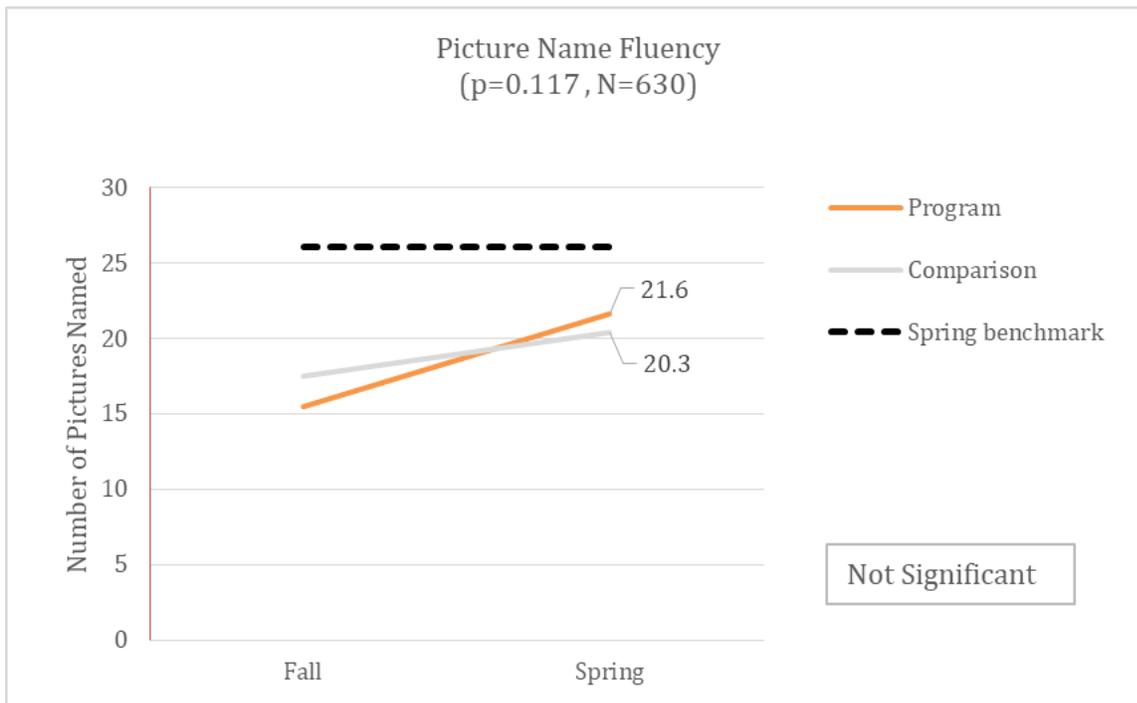


Figure i.5. Mean Picture Name Fluency Scores for PreK program and comparison students



Does program effect vary by participant characteristics, such as demographics?

A set of subgroup analyses of PreK student outcomes was conducted to examine whether program effects vary by participant demographics, including gender, race/ethnicity, and DLL status. For those subgroups where the sample sizes were sufficient to conduct analyses, either the magnitude of the effects were small or not statistically significant.

Concluding Thoughts

PreK students who received the Florida Reading Corps PreK program’s whole class and targeted interventions achieved significantly higher assessment scores for four of five tested emergent literacy skills by the Spring benchmark than did comparison students who did not participate in the program. The program had a net positive effect on both measures of students’ alphabet knowledge (i.e., *letter name fluency* and *letter sound fluency*), even though both groups in the study on average exceeded end-of-year benchmarks for these two measures. The program’s largest, positive effects were on both measures of phonological awareness (i.e., *rhyming fluency* and *alliteration fluency*). Overall, the results, based on effect sizes, for all four statistically significant measures were relatively small compared to the average score of students at comparison schools.

We contextualize these results given the early implementation of the Reading Corps PreK program in Florida, as well as some of the unique structural features of the program’s replication/implementation in M-DCPS and other unanticipated challenging circumstances that occurred during the 2017-2018 school year.

- **Program maturity.** The evaluation of the Florida Reading Corps program took place during its third year of implementation. This is a relatively early period in the program’s initial implementation phase,⁵ especially for a program that has an ambitious goal of whole class quality instructional improvement. As part of the outcomes study, the NORC research team completed a set of in-person site visits to a diverse cross-section of Florida Reading Corps and comparison group schools to assess the general classroom environment and language and literacy instructional practices. Observations from evaluator site visits confirmed that schools were continuing to work at implementing the program with fidelity. As such, it is recommended that the results of this evaluation be interpreted with the understanding that an *early* implementation of the Reading Corps model was assessed, and therefore, potentially more robust results may be observed after the program has matured to a full implementation stage.
- **Replication/Implementation.** Although the Florida Reading Corps PreK program is a replication site for the Minnesota Reading Corp PreK program, during the site visits conducted to select schools in the study, the evaluation team observed some important differences in program implementation of key program model elements. For example, in contrast to the broad representation of sites included in the state-wide Minnesota program, all of the sites in the Florida Reading Corps program were within

⁵ Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core Implementation Components. *Research on Social Work Practice*, 19(5), 531–540. <https://doi.org/10.1177/1049731509335549>

an urban setting, and also, Reading Corps and M-DCPS partners had agreed to focus initial implementation on schools with the greatest need, known as "Preschools on Probation." As a result, the Florida Reading Corps PreK program occurred in sites with a higher proportion of students with greater need. Second, initial implementation of the Florida Reading Corps proceeded with M-DCPS district personnel selecting sites to participate, whereas in the Minnesota model statewide expansion had allowed many sites to self-select into the program. Given the quasi-experimental nature of the evaluations, the difference in self-selection between the two evaluations could have impacted the findings, because self-selection can be a proxy for many other variables including teacher and staff support, knowledge of program elements, and other factors that can impact student outcomes.

- In addition to the structural differences in the program replication in Florida, it is important to note that several unfortunate events may have influenced the implementation of the Florida Reading Corps PreK program that were specific to the 2017-2018 school year, during which the study was conducted. These events included a Category 4 hurricane (Irma) causing several weeks of serious disruption to the school schedule; unusually high tutor turnover throughout the school year; and an emotionally difficult loss of a tutor on the first day of service. The Florida Reading Corps PreK program experienced 100% staff turnover in the spring of 2018, possibly due in part to these unfortunate events. Taken together, these unique and emotional challenges may have had a negative effect on the effectiveness of the Florida Reading Corps PreK program during the 2017-2018 school year.

In spite of the contextual factors described above that may have influenced the outcomes of the study, the Florida Reading Corps PreK program was able to produce statistically significant, positive results on four of five outcomes. Overall, the results of the evaluation show that the Florida Reading Corps PreK program had a positive effect on preschool students' alphabet knowledge and phonological awareness. These findings in combination with the strong results shown by the original model in Minnesota suggest that with more experience implementing the program, increased integration of the program in whole class/school instruction, and ongoing improvement in terms of fidelity to program implementation, the Florida Reading Corps PreK program may produce even larger effects in the future.

Introduction

Reading Corps is an AmeriCorps program that provides trained literacy tutors to support reading development for students in PreKindergarten (PreK) through grade 3, with the ultimate goal of ensuring that students become successful readers and meet reading proficiency targets by the end of the third grade. The core activities of the program are training, placing, and supporting literacy tutors in school-based settings to implement research-based, early-literacy strategies. Tutors are supported by a multi-level coaching model that includes site-specific (“Internal”) and external (“Master”) coaches. Since 2012, the PreK and the K-3 Reading Corps model has been replicated in several locations, including Colorado, Washington, D.C., Milwaukee, and Miami.

As part of an Innovative Approaches to Literacy (IAL) Grant through the U.S. Department of Education (ED), a multi-site impact evaluation of the Reading Corps program was conducted in three states during the 2017-2018 school year. While prior independent evaluations have been conducted of the original *Minnesota* Reading Corps program, this multi-state evaluation assessed the effect of the Reading Corps program model on students in Minnesota, as well as two replication sites, Florida and Wisconsin.

This report describes the findings from the evaluation of the Florida Reading Corps PreK Program and its effect on 630 four- and five-year old PreK students who attended either a Florida Reading Corps PreK program site or a matched-comparison site during the 2017-2018 school year. Students were enrolled at 12 Florida Reading Corps program sites and 12 comparison sites matched to each Reading Corps site based on specific site-level characteristics (e.g., location, school type, ages served, percentage of students eligible for Free or Reduced Price Lunch (FRPL), percentage of students who are Dual Language Learners (DLL), total site enrollment; and racial/ethnic distribution of students). The purpose of the evaluation was: 1) to evaluate the effect of the Florida Reading Corps PreK program on program participants compared to similar students who did not receive the Reading Corps program; and 2) to assess the degree to which the effect varies by participant characteristics (e.g., demographics).

The Reading Corps Program Model

Reading Corps is the largest AmeriCorps tutoring program in the country, providing trained literacy tutors to implement evidence-based literacy instruction and assessment protocols for students in PreK through grade 3 in school-based settings. Beginning in 2003 with just 250 students in Minnesota, Reading Corps has since expanded to 12 states and Washington D.C. Reading Corps aims to broadly impact literacy outcomes for children. A primary goal for the program is that all children, age 3 to grade 3, who qualify for Reading Corps, will meet reading proficiency standards by third grade.

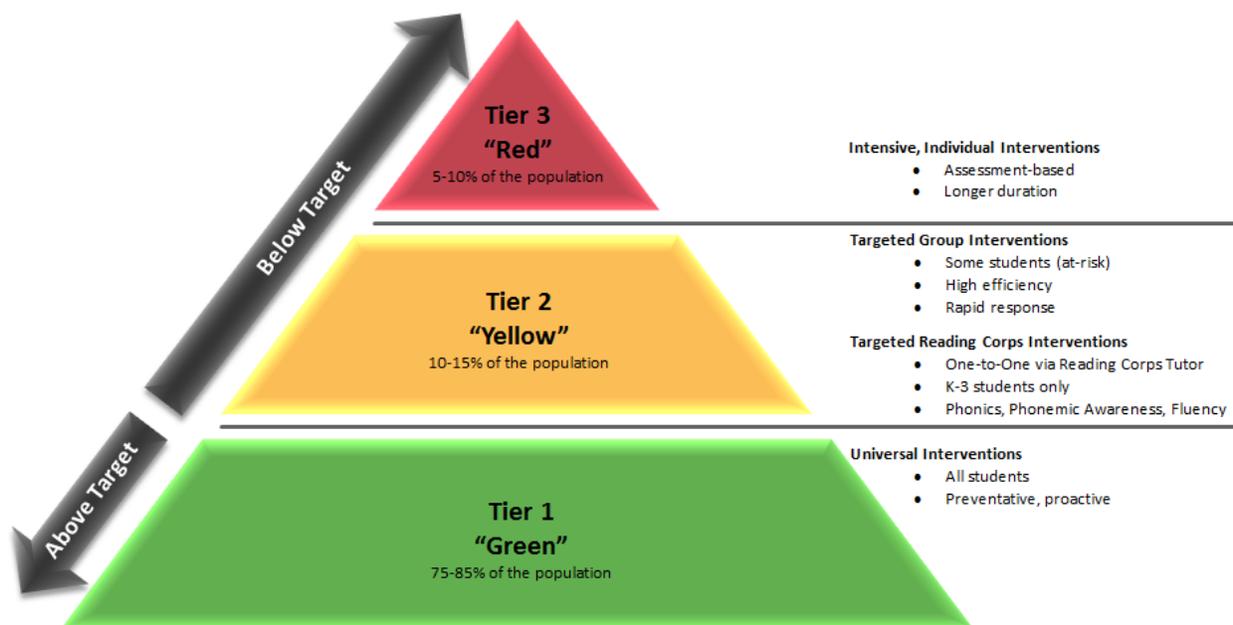
Response to Intervention Framework

Reading Corps is based on a Response to Intervention (RtI) framework which uses a multi-tier approach to the early identification and support of struggling learners. The key aspects of the Reading Corps RtI framework are:

- Clear literacy targets at each age level from PreKindergarten through grade 3
- Benchmark assessment to identify students eligible for one-on-one or small group interventions
- Evidence-based interventions
- Regular progress monitoring during intervention delivery
- High-quality training, coaching, and observations to support fidelity of implementation

In the RtI framework, data play the key roles of determining students' eligibility for additional services and then monitoring students' progress towards achieving academic expectations (i.e., benchmarks). Reading Corps screens students for program eligibility three times a year (i.e., fall, winter, spring) using grade-specific, literacy-focused measures that possess criterion-referenced grade- and content-specific performance benchmarks. These assessments are collectively called curriculum-based measures (CBM), because they correspond closely with curricular expectations for literacy skills at each developmental level. Program staff use scores from these general outcome measures to categorize students into one of three possible tiers (i.e., proficiency levels; see **Figure 1**): Tier 1 students score at or above benchmark and benefit from typical classroom instruction (75-80% of students score in this category); Tier 2 students score below benchmark and require specific supplemental interventions until they meet benchmarks (15-20% of students fall into this category); and Tier 3 students require intensive intervention provided by a special education teacher or literacy specialist and often have individualized educational plans (5-10% of students qualify for this category).

Figure 1. Response to Intervention – Three Tier Model



The Reading Corps PreK program includes both an immersive component, where tutors provide whole-class emergent literacy enrichment for all students (i.e., Tier 1), as well as a targeted small group and one-to-one component, where tutors provide more individualized interventions to students struggling with emergent literacy skills (i.e., Tiers 2 and 3). Although Reading Corps provides both PreK and K-3 interventions to students, the focus of this report is on the Florida Reading Corps PreK program. Therefore, the remainder of this report will focus on describing the PreK program and evaluation.

PreK Program Literacy Focus

The Reading Corps PreK program focuses on integrating the “Big Five” Early Literacy Predictors outlined by the SEEDS of Early Learning curriculum. These five categories summarize the work of the National Early Literacy Panel,⁶ and are incorporated into all aspects of the daily classroom routine. The “Big Five” for preschool students include: 1) oral language, conversation, and comprehension, 2) vocabulary and meaning, 3) concepts about books and print, 4) phonological memory and awareness (i.e., rhyming and alliteration), and 5) alphabet knowledge (i.e., letter name recognition and letter sound correspondence). Reading Corps tutors are tasked with creating a Literacy Rich Classroom using evidence-based practices assessed in the ELLCO (Early Language & Literacy Classroom Observation) tool, and implementing a Literacy and Math Rich Schedule. Along with implementing classroom-based strategies, tutors provide targeted individual or small group literacy tutoring for Tier 2 and Tier 3 students until they meet program-specified targets that predict end-of-year proficiency. During these small-group

⁶ National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.

and one-to-one sessions, tutors implement 5-to 15-minute scripted interventions focused on specific emergent literacy skills tailored to students' current proficiency levels.

The Reading Corps Literacy and Math Rich Schedule is designed to provide children with daily routines that embed early literacy predictors into fun and meaningful learning. As such, each classroom's day is structured around a Literacy and Math Rich Schedule that is integrated with the school's curriculum. The schedule includes activities, which include: Arrival, Sign-in, Meal Time, Large Group, Daily Message, Repeated Read Aloud, Tier 1 Small Group, Journal (weekly), Choice Time/Active Learning, Tier 2 or Tier 3 Targeted Interventions, "Big Five" Transitions, and Family Engagement (through Talk, Read, and Write with Me!). Within and between each scheduled activity, teachers and tutors strive to integrate Reading Corps expected routines, including "Strive for Five" conversations, following an overarching monthly instructional theme (e.g., weather, ecosystems, sports), functional vocabulary, and "Big Five" Transition Songs. The "Big Five" Transitions occur as children move from one part of the daily schedule to another. Tutors and teachers maximize instructional time by engaging children in an activity, such as rhyming games or letter sound songs, focused on one of the "Big Five" early literacy predictors.

Reading Corps tutors serving in the PreK program are also responsible for enacting the SEEDS of Early Learning,⁷ which provides both tutors and teachers with specific strategies to enhance literacy instruction for *all children* in the classroom. The SEEDS model is interactive, skills-focused and based on current research in early childhood education, child development, emergent literacy, social-emotional development, and effective teaching. SEEDS is a relationship-based instructional approach that maps out for teachers five ways to *intentionally* interact with children in order to promote academic growth and social-emotional development. SEEDS high quality teacher-student interactions include the following five elements:

- **Sensitivity:** Look, listen, and ask questions to become aware of each child's needs, thoughts, abilities and feelings;
- **Encouragement:** Use intentional affirmations and positive non-verbal communication to create a shared positive learning environment;
- **Education:** Embed the "Big Five" early literacy predictors in daily routines (vocabulary, conversation, phonological awareness, book and print rules, and letter knowledge);
- **Development of Skills Through Doing:** Help children explore their world through hands-on learning; and
- **Self-Image Support:** Balance the SEEDS quality interactions to support a child's feeling of being respected and capable.

⁷ Horst, K., & Passe, A. (2004). *Creating Literacy Rich Classrooms for Preschool Children (Ages 0-5)*. Presented at the 2004 CEED Symposium. Minneapolis, MN.; <https://www.seeds-learning.com/>

Use of Data in Program Implementation

In the Reading Corps PreK program, tutors collect general outcome measure data on all PreK students in a classroom in the Fall, Winter, and Spring of each school year using three age-appropriate Individual Growth and Development Indicators (IGDIs; i.e., *rhyming fluency*, *picture name fluency*, *alliteration fluency*)⁸ and two Fastbridge (FAST)⁹ assessments (i.e., *letter name fluency* and *letter sound fluency*). The assessments measure the following five critical emergent literacy skills:

1. **Letter Name Fluency (LNF)** evaluates students' alphabet knowledge by assessing how many upper case letter names a student can identify within one minute. Students receive one point for every letter name correctly stated within one minute. The total number of letters named correctly is the *letter name fluency* score.
2. **Letter Sound Fluency (LSF)** assesses how many letter sounds a student can identify and say (e.g., l, f, m) within one minute. Students receive one point for every correct letter sound provided within one minute. The total number of correct letter sounds is the *letter sound fluency* score. *Letter sound fluency* is considered a more difficult/advanced measure of a child's alphabet knowledge.
3. **Rhyming Fluency (RF)** tests students' phonological awareness by assessing how many rhyming matches a student can correctly identify (e.g., bat and cat, car and jar, ham and jam) within two minutes. Students receive one point for every correct rhyme identified within two minutes. The total number of correct rhymes is the *rhyming fluency* score.
4. **Alliteration Fluency (AF)** tests students' phonological awareness by assessing how many alliteration matches (e.g., hill and heart, cat and can, bike and ball) a child can correctly identify within two minutes. Students receive one point for every correct alliteration identified within two minutes. The total number of correct alliterations is the *alliteration fluency* score.
5. **Picture Name Fluency (PNF)** tests students' vocabulary by assessing how many pictures a child can name correctly within one minute. Students receive one point for every picture named correctly within one minute. The total number of pictures named correctly is the *picture name fluency* score.

Students' scores on these five measures are used to identify PreK students who are eligible for Tier 2 or 3 interventions. Tutors then collect monthly progress monitoring data on those Tier 2 and 3 students to whom they provide individualized or small-group interventions. **Table 1** lists the five outcome measures and corresponding benchmark scores used to identify eligible 4- and 5-year old PreK students for Tier 2 and 3 interventions. Students who score far from target are prioritized for small group or one-to-one interventions. The benchmarks for Winter and Spring are the same for 4- and 5-year old PreK students.

⁸ <https://www.myigdis.com/>

⁹ <http://www.fastbridge.org/assessments/>

Table 1. Spring PreK Benchmarks (4/5 Year-Old Children)

Assessment	On or Above Target	Close to Target	Far from Target
Letter Name Fluency	14+	8 – 13	0 – 7
Letter Sound Fluency	10+	5 – 9	0 – 4
Rhyming Fluency	12+	7 – 11	0 – 6
Alliteration Fluency	8+	5 – 7	0 – 4
Picture Name Fluency	26+	16 – 25	0 – 15

The data intensive nature of the Reading Corps program provides a consistent, objective means of identifying students to receive small-group and one-on-one interventions, tracking their progress toward achieving academic goals related to critical literacy skills, and informing Tier 1 instruction. The data also provide tutors and coaches with objective information about the efficacy of the interventions with individual students, which can in turn be used to tailor the most effective instruction for the student’s skill level.

Coaching and Support

Reading Corps provides multiple layers of supervision to ensure integrity of program implementation. Site-specific Internal Coaches serve as supervisors, mentors, and advocates for tutors. The Internal Coach’s role is to monitor tutors and provide guidance in the implementation of the Reading Corps PreK program in participating classrooms (e.g., Literacy and Math Rich Schedule, assessments, “Big Five” Transitions, “Strive for Five”, SEEDS quality interactions, and small group and one-on-one interventions). As the front-line supervisor, the Internal Coach is a critical component of the supervisory structure. In Florida, district-level literacy curriculum specialists, who were already assigned to participating M-DCPS elementary schools, tended to serve as Internal Coaches for the program.¹⁰ These coaches observed tutors delivering interventions once or twice per month. Coaches were also available for additional support and guidance, as needed.

The external, or Master Coach, is a literacy expert who provides site staff (e.g., Internal Coaches and Reading Corps tutors) with expert consultation on literacy instruction and ensures implementation integrity of Reading Corps program elements. In addition to these two coaching layers, a third layer consisting of Reading Corps program support staff helps ensure a successful year of AmeriCorps service. Program support staff are Reading Corps employees who provide administrative oversight for program implementation to sites participating in Reading Corps. Finally, each Reading Corps site utilized an on-site administrative liaison, who provided limited general oversight to tutors at the school (e.g., monitoring attendance).

The classroom teachers were provided training on the Reading Corps program in order to assist Reading Corps tutors with program implementation in their classroom. Tutors were expected to conduct monthly

¹⁰ Internal Coaches at two YWCA schools were on-site staff (i.e., a teacher and a sub/floater) rather than district-level employees.

literacy assessments to increase the number of children on target with early reading predictors. Tutors periodically reviewed their students' data and progress with the teachers and Internal Coaches to make determinations about students' progress. Tutors also watched videos of themselves and received feedback on improving their delivery of tutoring sessions and classroom-level interventions, to make sure that students were effectively engaged in literacy activities.

Training

Prior to the start of each school year, Florida Reading Corps hosts a three-day Summer Institute to train returning and new Master Coaches, Internal Coaches, and Reading Corps tutors. This intensive, information-filled training provides foundational training in the assessments and research-based emergent literacy interventions the Reading Corps uses. During several sessions at the Institute, tutors learn the skills, knowledge, and tools needed to implement the program in PreK classrooms. These sessions introduce tutors to the Florida Reading Corps PreK program model, the Literacy and Math Rich Schedule, the SEEDS of Early Learning approach to high quality adult-child interaction, the interventions used in the Tier 2 and 3 small group and one-to-one tutoring sessions, instruction in administering and scoring student assessments, as well as the underlying research and theories supporting the interventions and program model. The SEEDS of Early Learning approach serves as the pedagogical framework within which tutors create a literacy rich classroom environment.

Tutors are provided with detailed manuals as well as online resources that mirror and supplement the contents of the manual (e.g., videos of model interventions and best practices). Both the manuals and online resources are intended to provide tutors with just-in-time support and opportunities for continued professional development and skill refinement. Additional training and coaching sessions are provided throughout the tutors' year of service.

Florida Reading Corps

As the focus of this report, the Florida Reading Corps program is a replication of the successful Minnesota Reading Corps PreK model. In the 2017-2018 school year when this evaluation was conducted, the Florida Reading Corps was in its third year of operation with 35 Reading Corps tutors in 12 schools in Miami: 10 Miami-Dade County Public School District (M-DCPS) elementary schools and two YWCA PreK Centers. Florida Reading Corps also receives local support through key partnerships with Volunteer Florida, M-DCPS, The Children's Trust of Miami-Dade, and the YWCA of Miami. The Children's Trust of Miami-Dade, which is a dedicated source of revenue established by voter referendum, funds initiatives to benefit children and families in Miami-Dade County, to better their lives and outcomes, and ensure a successful future for the county.

About the Study

This section provides a detailed description of the methodology used to implement the Florida Reading Corps PreK program outcome evaluation. The methodology was largely informed by the previous *Outcome Evaluation of the Minnesota Reading Corps PreK Program*.¹¹

Evaluation Logic Model

A logic model for the Reading Corps program illustrating key program and school inputs and activities, as well as the program's desired short-term outcomes and long-term goals for PreK and K-3 students is provided in **Appendix A**. The Reading Corps logic model serves as the conceptual framework for the design of the Florida Reading Corps outcome evaluation. The logic model presents a comprehensive illustration of the complete Reading Corps program and includes inputs, activities, short-term outcomes, and long-term goals for PreK and K-3 students. The focus of the Florida Reading Corps PreK program outcome evaluation was to assess the effect of program participation on PreK students' emergent literacy skills. As such, the evaluation focused on only those components of the logic model relevant to PreK students.

Four key program- and school-based inputs and resources are essential to successful implementation of the Florida PreK program: 1) district-level selection of schools based on the degree of student and school need; 2) school identification of at-risk (Tier 2 and Tier 3) PreK students within the classroom based on benchmark assessments of students' emergent literacy skills; 3) web-based data management systems to track and monitor student progress with literacy interventions; and 4) school use of research-based core literacy curriculum.

Three additional Reading Corps PreK program inputs related to Reading Corps tutors include: 1) recruitment, screening and placement of tutors in schools; 2) comprehensive Reading Corps training of tutors and Internal Coaches in literacy interventions, assessment, data-driven decision-making and program policies; and 3) district-level¹² identification and assignment of dedicated Internal Coaches to support tutors. The logic model also illustrates the multiple layers of supervision and coaching the Reading Corps program provides to its Internal Coaches and tutors.

As shown in the logic model, the Reading Corps PreK program's primary activities include: 1) enriching the literacy environment; 2) supporting classroom instruction via whole class, small group, and one-on-one interventions (i.e., tutoring); 3) conducting benchmark assessments on emergent literacy skills three times per year (fall, winter and spring) to identify students in need of small group or one-on-one intervention (i.e., Tier 2 or Tier 3 students, respectively); 4) delivering one-on-one and small group emergent literacy interventions to students needing extra assistance; 5) assessing and charting monthly student progress on age-specific emergent literacy skills for targeted children (i.e., monthly progress

¹¹ Markovitz, C.; Hernandez, M.; Hedberg, E.; Silbergliitt, B. (2014). Outcome Evaluation of the Minnesota Reading Corps PreK Program. NORC at the University of Chicago: Chicago, IL.

¹² The Head Start Program Director identified Internal Coaches for the community-based YWCA preschools.

monitoring); 6) “exiting” students from Tier 2 or Tier 3 interventions once they achieve assessment scores putting them on track to meet or exceed the spring benchmark; and 7) identifying new students eligible for small group or one-on-one interventions. The intended short-term outcomes of these activities are demonstrated improvement on the three Individual Growth and Developmental Indicator (IGDI) and two FastBridge (FAST) measures at the subsequent benchmarking period (i.e., Winter, Spring). The desired long-term outcome of the Reading Corps program is for all PreK students to achieve “kindergarten ready” targets on IGDI and FAST measures by the end of the school year.

Research Questions

The primary goal of the Florida Reading Corps PreK program outcome evaluation was to independently and quasi-experimentally assess the effect of the Florida Reading Corps PreK program on 4- and 5-year old students’ emergent literacy skills.

Thus, the primary research questions for the evaluation are as follows:

1. What is the effect of the Florida Reading Corps PreK program on program participants compared to students at similar schools who did not receive Reading Corps?
2. Does program effect vary by participant characteristics, such as demographics?

To answer these research questions, the evaluation team analyzed age-appropriate and semester-specific emergent literacy assessment scores collected from 630 PreK students enrolled at 24 public preschools and YWCA PreK Centers in Florida during the 2017-2018 school year. Half of the sites (12) had implemented the Florida Reading Corps program for at least one year and served as the evaluation’s “program” sites. The remaining 12 sites (i.e., comparison sites) were selected for participation in the study based on a prescribed set of educationally significant school-level characteristics (described in detail in the Selection and Matching of Schools section below), which matched those of the Florida Reading Corps program sites. Thus, 12 pairs of matched sites (i.e., 12 program sites and 12 comparison sites) participated in the evaluation. Fall and Spring student benchmark data were used to answer the evaluation’s primary research questions, which estimate the effect of the Florida Reading Corps PreK program on students’ emergent literacy outcomes.

Selection and Matching of Schools

All 12 of the Florida Reading Corps PreK program sites operating across M-DCPS in 2016-2017 were selected to participate in the PreK evaluation during the 2017-2018 school year.¹³ Power analyses determined that the students enrolled in PreK classrooms at the 12 Reading Corps and 12 matched comparison schools would provide a sufficient sample size for the evaluation.¹⁴ Due to the nature of the

¹³ Two additional M-DCPS Florida Reading Corps sites were just starting their first year of implementation during the 2017-2018 school year and, therefore, did not participate in the study.

¹⁴ The study is designed to have the power to detect an effect size of 0.40 (assuming a two tailed test with 0.8 power, ICC= 0.2, and 5% error rate), which is equivalent to the effect size found among PreK students in the previous Minnesota Reading Corps study.

Florida Reading Corps PreK program as a whole-class/whole-site intervention, the study had to employ a quasi-experimental design and identify a set of comparison schools matched closely with Florida Reading Corps PreK program schools on characteristics that education researchers and evaluators know to be critical to student achievement (presented in **Table 2** below in no particular order): 1) location/neighborhood; 2) program type (public PreK, community-based PreK, Head Start); 3) ages served (4 and 5 year olds); 4) Socio-Economic Status (SES) as measured by percentage of students eligible for Free or Reduced Price Lunch; 5) percentage of students who are Dual Language Learners (DLL); 6) total site enrollment; and 7) racial/ethnic distribution of PreK students.

Table 2. PreK QED Matching Variables

Matching Variable	Description
Location/Neighborhood	Proximity to Florida Reading Corps School
Program Type	Public PreK, YWCA, Head Start
Composition of Student Ages Served	4 and 5 year olds
Socio-Economic Status (SES)	Percent of students eligible for Free or Reduced Price Lunch (FRPL) or equivalent
Percent Students Who are Dual Language Learners (DLL)	Percent of students who do not speak English as a first language
Site Enrollment	Number of students enrolled in PreK program
Race/Ethnicity	Racial/ethnic distribution of PreK students

Matching is a multi-staged process, which starts with finding sites that match on location/neighborhood and/or are in proximity to the Florida Reading Corps PreK schools. Comparison sites were located in close enough proximity to the Florida Reading Corps PreK programs to have similar neighborhood and population characteristics. Information on additional site-level variables was then used to select sites that best matched the Florida Reading Corps PreK sites. At the end of the matching process, a single “primary” site and a second “backup” site were identified that matched best with each Florida Reading Corps PreK site across the seven matching variables. After the initial sample of 12 matched comparison sites was selected to be paired with the 12 Reading Corps sites, the sample was reviewed for completeness and representativeness. Florida Reading Corps and M-DCPS generously offered to assist with the recruitment process, reaching out to individual principals and directors to explain the study. For this reason we had a very high acceptance rate and only a few instances where we had to rely upon our second selection.

The final sample comprised 24 sites (12 Florida Reading Corps and 12 comparison sites) and included 630 students. **Table 3** lists all 24 PreK sites that participated in the evaluation. The total student sample size was determined by the number of students enrolled in a participating PreK classroom at each of the 24 selected sites. For the Reading Corps schools, participating PreK classrooms are defined as those with an assigned Reading Corps tutor, and all students who were enrolled in these classrooms were asked to participate. At the comparison schools, all PreK students were asked to participate.

Table 3. Characteristics of Schools Participating in the Florida Reading Corps Outcome Evaluation (Fall 2017)

School and Study Group	Number of Reading Corps Tutors ^{a,b,c}	Number of participating classes ^{a,c}	% FRPL ^{d,e}	% School Non-White ^{d,e}	PreK Enrollment ^d	Total School Enrollment ^{d,e}	Study Participants ^c
Reading Corps Schools							
Biscayne Gardens	1	2	94.2%	98.8%	64	602	23
Brentwood	3	3	91.8%	99.4%	56	625	46
Campbell Drive K-8	2	2	97.8%	98.3%	50	928	27
Charles Drew	2	2	F	f	f	f	26
Comstock	1	2	99.3%	99.1%	53	564	25
Florida City	1	1	96.6%	97.4%	39	805	18
Lillie C. Evans	2	2	98.4%	99.8%	54	444	30
Nathan B. Young	3	3	99.5%	99.2%	73	358	40
Paul L. Dunbar K-8	2	2	98.2%	97.7%	56	395	23
Treasure Island	2	2	80.8%	83.1%	72	562	35
YWCA Colonel Zubkoff	5	5	F	f	f	f	42
YWCA Gerry Sweet	2	3	F	f	f	f	27
Comparison Schools							
Biscayne Elementary	—	2	88.6%	93.1%	63	740	32
Bunche Park	—	2	98.7%	98.2%	38	379	26
Carol City	—	2	93.8%	98.9%	73	448	36
Leisure City K-8	—	3	97.5%	97.5%	89	1061	41
Liberty City	—	5	97.0%	99.4%	132	471	35
Maya Angelou	—	3	99.3%	98.8%	47	759	27
Oak Grove	—	2	94.0%	98.6%	35	581	19
Orchard Villa	—	1	95.4%	99.5%	42	434	18
Phyllis Wheatley	—	1	98.8%	99.2%	17	251	11
West Homestead K-8	—	1	99.0%	97.5%	19	773	12
Y Intergenerational (Windows)	—	1	f	f	f	f	8
YWCA Cain Tower	—	1	f	f	f	f	3

a: 2017-2018 data reported directly from Reading Corps and confirmed with individual schools as of Spring 2018; changes may have occurred throughout the year.

b: Tutors for participating classrooms.

c: Numbers are reported for largest analytic sample N=630. Analytic sample sizes range from 614 to 630 by outcome.

d: 2015-2016 data from CCD (NCES)

e: All students in all grades, excluding Alternate Education

f: CCD (NCES) data not available for any year 2001-2016.

Table 4 presents descriptive statistics for all PreK students included in the evaluation. Demographics include gender, race/ethnicity, and Dual Language Learner (DLL) status. The analytic sample sizes for the evaluation range from N=614 to 630 depending on outcome. Assessments were administered over multiple days and because some children were absent or unable to complete all five assessments, the analytic sample sizes for each measure differ slightly.

Table 4. Student participants for the Florida Reading Corps Outcome Evaluation (Fall 2017)

	Program N=362	Comparison N=268
<i>Gender^a</i>	---	---
Male	47%	50%
Female	53%	50%
<i>Race/Ethnicity^a</i>	---	---
Black or African American	67%	65%
Hispanic/Latino	28%	32%
Multi-racial	3%	0%
White	1%	3%
American Indian or AK Native	—	—
Asian ^b	<1%	0%
Native HI or Other Pacific Islander	—	—
Other	—	—
<i>Dual Language Learner (DLL) Status^c</i>	---	---
DLL	31%	28%
Non-DLL	69%	72%

a: Data not available for 15 comparison students.

b: Absorbed into White category for analysis.

c: Data not available for 74 program students and 43 comparison students.

As shown in **Table 4** above, at least 314 PreK students within classrooms across the 24 sites were included in the evaluation, resulting in adequate power to detect the minimally expected effect sizes (>0.40 standard deviation units).

Data Sources

As discussed above, three IGDI and two FAST assessments are the primary sources of data for PreK student emergent literacy outcomes at Florida Reading Corps sites. The evaluation team obtained Fall benchmark assessment data collected on all PreK students enrolled in classrooms with a Reading Corps tutor at the 12 Florida Reading Corps sites. As mentioned previously, Reading Corps tutors collect benchmark data three times a year on all students in their Reading Corps PreK classrooms to determine eligibility for Tier 2 or 3 small group or one-to-one interventions. Given that tutors already collect this data, and so as not to duplicate assessment efforts, the evaluation team requested access to the 2017 Fall benchmark data for our analysis of student outcomes. Also, bias in the collection of pretest Fall scores by program staff was highly unlikely because the assessments were collected at the start of the school year before much instruction occurred, and prior to any Tier 2 or Tier 3 small group or one-on-one intervention began. In order to address potential bias at posttest, independent assessors from the NORC evaluation

team collected Spring benchmark assessment data from students for estimating program effects over a single school year.

Because the 12 comparison sites did not collect IGDI and FAST data, NORC independent assessors collected this data from PreK students at these sites twice during the year (Fall and Spring). The independent assessors completed the exact same assessments using the same tools as the Florida Reading Corps PreK program staff and participated in an assessment training that is similar to the training provided to Reading Corps tutors. This process ensured that the field testers: 1) used the same data collection instruments for collecting student level benchmarks on comparison site students as those used to collect data from students at the Florida Reading Corps PreK program sites; and 2) received the same standardized training in assessment administration and scoring as the Reading Corps tutors. This common use of instrumentation and training ensured commonality of assessment practices across both program and comparison sites. Informed written consent from program and comparison students' parents was obtained prior to inclusion in the evaluation.

Finally, program staff provided the evaluation team with other administrative data on Florida Reading Corps program students, including tutoring attendance, average length and number of sessions, and demographic characteristics. The evaluation team obtained similar demographic data on comparison students from the Miami-Dade County Public School District. Upon completion of data collection, the student-level IGDI and FAST data collected from comparison sites were combined for analysis with de-identified IGDI and FAST data from the Florida Reading Corps sites described earlier. The resulting dataset contained all the assessment and demographic data collected or extracted for Florida Reading Corps and comparison sites. The final dataset was processed, cleaned, and coded in preparation for data analysis.

Analysis Methods

The study analysis is designed to determine if there is an effect of the Florida Reading Corps PreK program on student emergent literacy outcomes, and if so, to estimate a plausible average treatment effect of the program. We estimate program effects by statistically comparing the average spring score of students enrolled at sites who participated in the Florida Reading Corps PreK program with the average spring score of a comparison group of students at similar sites who did not participate. The calculated difference in group averages is the plausible effect of program participation.

The outcome evaluation analyses were conducted using mixed-regression models, also known as hierarchical linear models (HLM).¹⁵ This model uses maximum likelihood to estimate differences in outcomes associated with predictor variables while also modeling the multi-level structure of the data, which in this case are students nested within sites. Two models are estimated for each outcome. The first model uses program status and age in months as the only predictors. From this first model, we calculate an effect size based on typical standard deviation based effect sizes such as the Cohen's *d* or Hedges' *g*

¹⁵ Also note that our models include weights to adjust for student non-response and school generalizability.

parameter for multilevel designs.¹⁶ The second model includes control variables to remove any spurious effects of observed demographic and site characteristics on outcome values. While the program and control groups are well balanced, they are not perfectly equal, and thus, the treatment effect is slightly different. From this second model, we then predict the scores for program and control groups based on the program variable and holding all other controls as their sample means. Information from the second model was used to produce the plots and tables provided in the report. More detailed information about the specific statistical models employed is available in **Appendix C**.

Study Limitations

The primary objective of the Florida Reading Corps PreK evaluation was to assess the effect of the Florida Reading Corps PreK program on preschool students' emergent literacy outcomes. In order to achieve this objective, the evaluation was designed to measure Florida Reading Corps PreK program effects using a quasi-experimental design. However, unavoidable limitations inherent in the study design and in working with schools and students constrained some aspects of the evaluation's design, implementation, and analysis.

The single most important limitation of the study design is its inability to control for unobserved differences between the Florida Reading Corps PreK program and comparison sites. Despite implementing a thoughtful approach to the matching process using educationally significant baseline measures, it is not possible to obtain or match on all possible site-level factors which may bias student outcomes. Since comparison sites were matched by the research team to the pre-determined sites that participated in the program, we are only able to control for observed characteristics and not for all possible characteristics that may influence the outcomes. We attempt to control for some unobserved characteristics that may be correlated with the observed school variables used for matching through the use of site-pair blocking variables. However, given that treatment selection is not a randomized mechanism, our ability to make causal interpretations is limited. For example, the strongest predictor of student performance on outcome measures is their performance at baseline (i.e., Fall benchmark [pretest] score). Our matching process was conducted without access to any baseline data because it had yet to be collected, thereby, possibly resulting in bias within our site-level pairs. To examine these differences and test the integrity of our matching process, we conducted statistical tests of average Fall scores between sites (simple t-tests) and found the majority of our site-level pairs did not result in significant differences at baseline across the five outcome measures. Moreover, we also estimated the difference in means using the Hedges *g* effect size, and found that most of the differences in the baseline scores were represented by small effect sizes. For further information on our balance tests (school matching validity analysis), please see **Appendix E**.

Another major constraint is that site participation in the evaluation was voluntary for comparison schools. For example, a small number of the original sites chosen for the comparison group declined to participate.

¹⁶ Hedges, L. V., & Rhoads, C. (2010). *Statistical Power Analysis in Education Research*. NCSER 2010-3006. *National Center for Special Education Research*; Hedberg, E. C. (2017). *Introduction to Power Analysis: Two-group Studies* (Vol. 176). SAGE Publications.

In anticipation of this common challenge in quasi-experimental study site recruitment, we initially identified two well-matched sites for each Reading Corps site, and were successfully able to recruit the second site for those few instances with the first match declined.

A common concern in studies is student attrition. Some students inevitably moved away, were chronically absent (i.e., attrition), or declined to participate in the spring benchmark assessment. When this occurred, the affected student was dropped from the study's analytic sample. The deletion of cases was moderate, but different for each of the five outcomes due to the timing of the spring assessments, which occurred on different days. The sample started with 669 eligible students (380 in the program and 289 in the comparison group). The total number of sample members removed for *picture name fluency*, *rhyming fluency*, and *letter name fluency* was 39 students, resulting in an analytic sample of 630 students.¹⁷ For *alliteration fluency* and *letter sound fluency*, the analytic sample was slightly smaller at 615 and 614 students, respectively.¹⁸ As noted above, the evaluation team conducted a power analysis prior to developing the sampling plan, to ensure an adequate number of students in order to detect a difference between the treatment and control groups. While the sample is adequately powered to detect effects, it is still limited by the availability of students. This situation raises the possibility that the effects represented in this report only reflect the effects for a specific population of students and schools (i.e., those who stay at the school and are not chronically absent). However, the numbers of deleted cases were small, so this was not a serious concern for the study.

Another common limitation related to interpretation of outcome evaluation results is concern about whether alternative interventions similar to the treatment (in this case, the Florida Reading Corps' Literacy Rich Schedule and SEEDS approach to instruction) could similarly affect student outcomes at comparison sites. No other similar alternative interventions were identified at the PreK comparison sites, including during site visits to select comparison schools.

Site Visits

The NORC research team chose a diverse cross-section of 12 M-DCPS and YWCA Reading Corps and comparison group schools to visit in-person between March and April 2018 based on descriptive information provided by Florida Reading Corps, as well as administrative data from M-DCPS and YWCA. Schools were selected with consideration for the following criteria: school/program type (public, private, Head Start); student race/ethnicity; and percentage of Dual Language Learners (DLLs). NORC visited a select group of Florida Reading Corps schools and their matched comparison schools to better understand differences between Reading Corps and non-Reading Corps sites. Two-person teams visited each school. Qualitative information was collected from discrete observations of the general classroom

¹⁷ 37 students withdrew (17 program and 20 comparison) from the schools; 1 student was identified as special needs and was removed from the program group; and 1 student in the comparison group declined the assessment. The overall attrition rate was 6% (5% program and 7% comparison).

¹⁸ For *alliteration fluency*, six additional students were absent from the Fall benchmark assessment day (all comparison), five were absent from the spring testing (three program and two comparison), and another four withdrew from school. This yielded an analysis sample of 615 students. For *letter sound fluency*, an additional one student in the program group was simply not tested by mistake, yielding an analysis sample of 614. This produced an overall attrition rate of 8% (6% program and 11% comparison) for both *alliteration fluency* and *letter sound fluency*.

environment and language and literacy instructional practices, using a tool specifically developed to assess fidelity of implementation of the activities of the Literacy Rich Schedule and their alignment with the corresponding evidence-based elements of the ELLCO. The ELLCO PreK tool is a validated observational measure used by the Reading Corps to assess and monitor the quality of the classroom environment and literacy and language practices in PreK classrooms. The ELLCO observation instrument is used to gather information about the quality of the language and literacy environment in the classroom by observing literacy activities, teacher-child interactions, and use of materials and classroom space. The tool provides descriptive indicators to guide the assessor's observation and provide evidence for the evaluative ratings.

Each visit began by collecting information about the physical classroom environment, as well as a 2-hour observation of a typical day and classroom routines. During this time observers described the language and literacy instruction and activities in which the teacher and children were engaged. Following the observation, interviews were conducted with the school principal, classroom teacher, and Reading Corps tutor(s) at the Florida Reading Corps schools, to get a fuller picture of how Reading Corps operated in that school. These observations and interviews were used to contextualize the quantitative student findings in the outcomes study.

Study Findings

In this section, we present the major findings from the analysis of PreK study participants. Detailed results are found in **Table D.1** in **Appendix D**.

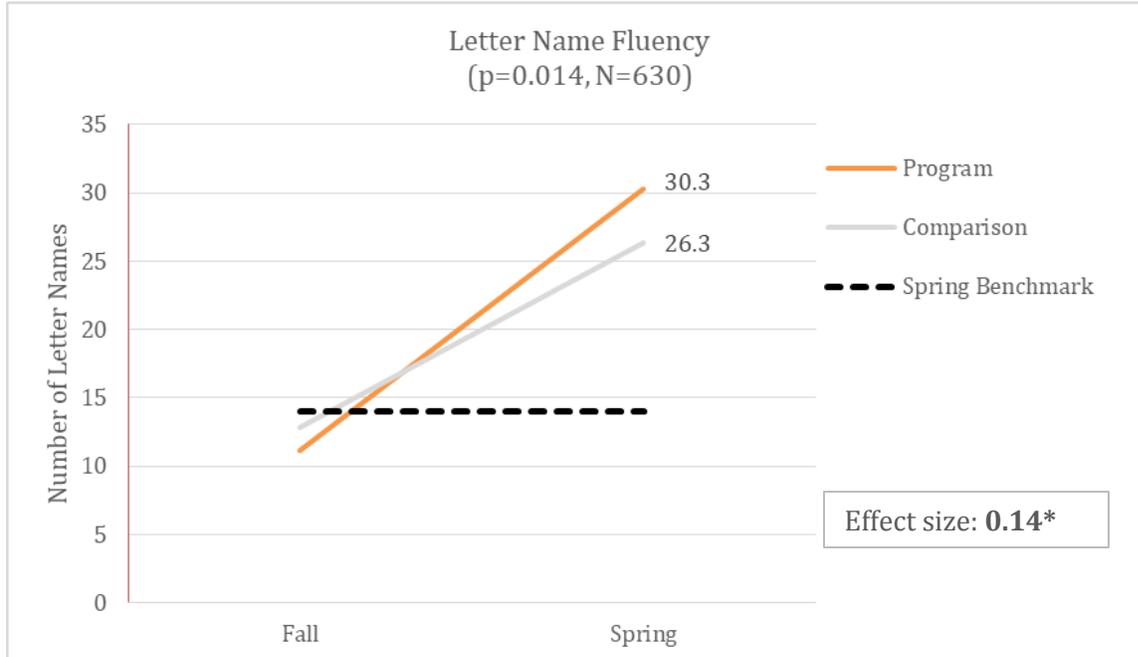
What is the effect of Florida Reading Corps on PreK students?

As discussed above, the Florida Reading Corps PreK outcome evaluation of 630 students from 24 PreK programs included 12 schools implementing the Florida Reading Corps PreK program and 12 demographically comparable schools, which did not participate in the Florida Reading Corps program. NORC's analysis examined spring 2018 differences between children who participated in the Florida Reading Corps PreK program and comparable students at comparison schools who did not, on the following five emergent literacy skills: letter name identification, letter sound correspondence, rhyming, alliteration, and picture naming. Our analysis revealed statistically significant differences between the Florida Reading Corps and comparison groups for four of the five measures: *letter name fluency*, *letter sound fluency*, *rhyming fluency*, and *alliteration fluency*. There were no significant differences between study groups for *picture name fluency*. Below we contextualize our results as effect sizes and compare the achieved effect sizes to the typical effect sizes that represent age appropriate, year-over-year growth of students in reading skills.

Letter Names

The Florida Reading Corps PreK program had a statistically significant effect on PreK students' *letter name fluency* scores between the Fall and Spring benchmark. Program students correctly identified 4.0 more letter names on average than students in the comparison group ($p=0.014$, $N=630$). This represents an effect size of 0.14, which is small considering the typical year-over-year growth of students this age. Program students' average score of 30.3 was higher than the comparison group's average score of 26.3, as presented in **Figure 2**, though both groups surpassed the Spring benchmark score of 14 correctly identified letter names by a substantial amount.

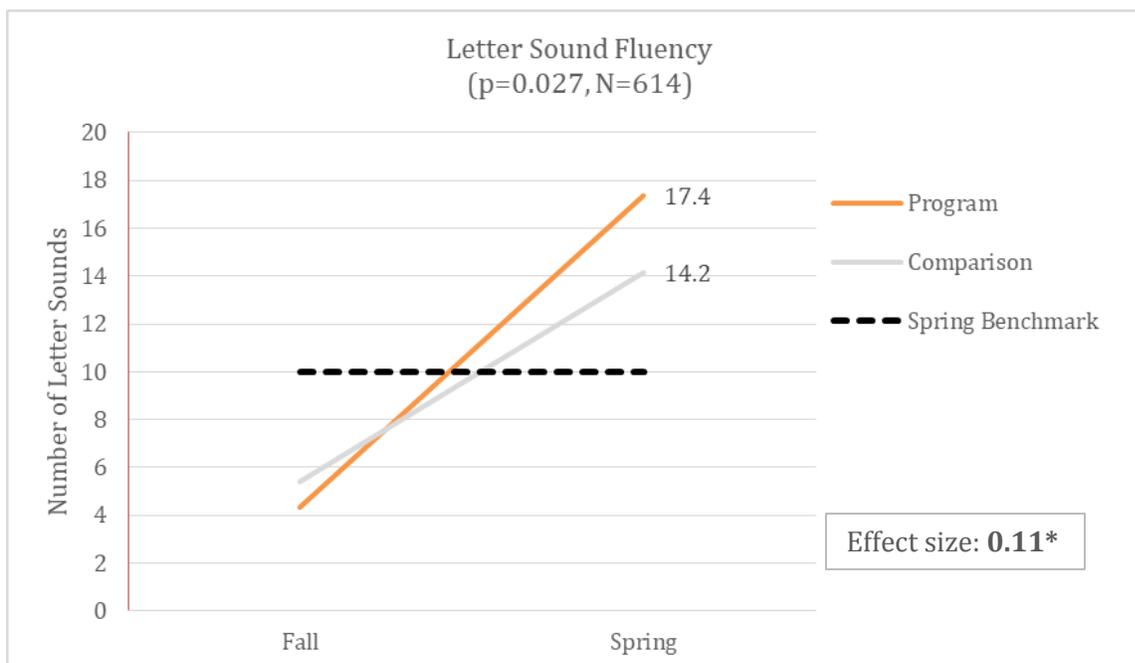
Figure 2. Mean Letter Name Fluency Scores for PreK program and comparison students



Letter Sounds

The Florida Reading Corps PreK program had a statistically significant effect on PreK students' *letter sound fluency* scores between the Fall and Spring benchmark. On average, program students correctly identified 3.2 more letter sounds than students in the comparison group ($p=0.027$, $N=614$). This represents an effect size of 0.11, which is small considering the typical year-over-year growth of students at this age. Program students' average score of 17.4 was significantly higher than the comparison group's average score of 14.2, as shown in **Figure 3**. Also, both the program and comparison groups surpassed the Spring benchmark score of 10 letter sounds.

Figure 3. Mean Letter Sound Fluency Scores for PreK program and comparison students

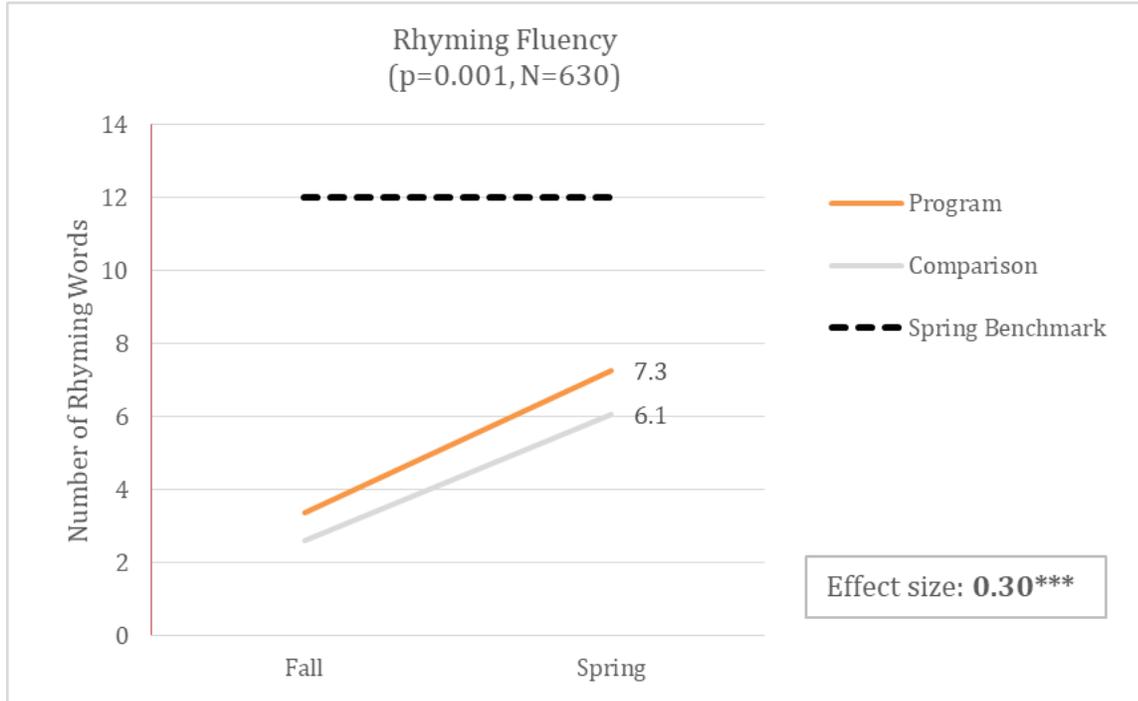


Rhyming

The Florida Reading Corps PreK program had a statistically significant effect on PreK students' *rhyming fluency* scores between the Fall and Spring benchmark. On average, program students correctly identified 1.2 more rhyming words than students in the comparison group ($p=0.001$, $N=630$).¹⁹ This represents an effect size of 0.30, which is small considering the typical year-over-year growth of students at this age. The program students' average score of 7.3 was significantly higher than the comparison group's average score of 6.1, as shown in **Figure 4**. However, neither group met the Spring benchmark score of 12 rhyming words.

¹⁹ Readers will note that there is a similar difference between Program and Comparison students at the Fall baseline in Figure 4. However, our test of statistical significance includes the Fall baseline score as a covariate (in addition to demographic characteristics) to remove this source of variation.

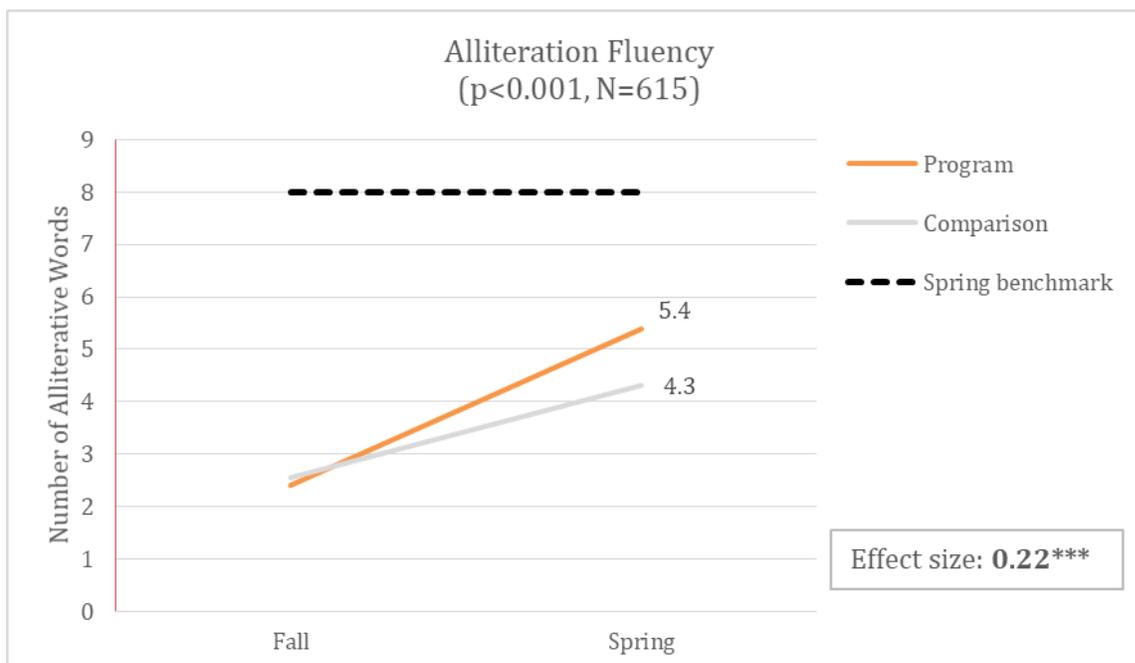
Figure 4. Mean Rhyming Fluency Scores for PreK program and comparison students



Alliteration

The Florida Reading Corps PreK program had a significant effect on PreK students' *alliteration fluency* scores between the Fall and Spring benchmark. On average, students in the program group correctly identified 1.1 more alliterative words than students in the comparison group ($p < 0.001$, $N=615$). This difference represents an effect size of .22, which is small given the typical year-over-year growth of students this age (which is typically an effect size of 1.5 or more for early reading skills). While the program students' average score of 5.4 was higher than the comparison group's average score of 4.3, as shown in **Figure 5**, neither group met the Spring benchmark of eight alliterative words correctly identified.

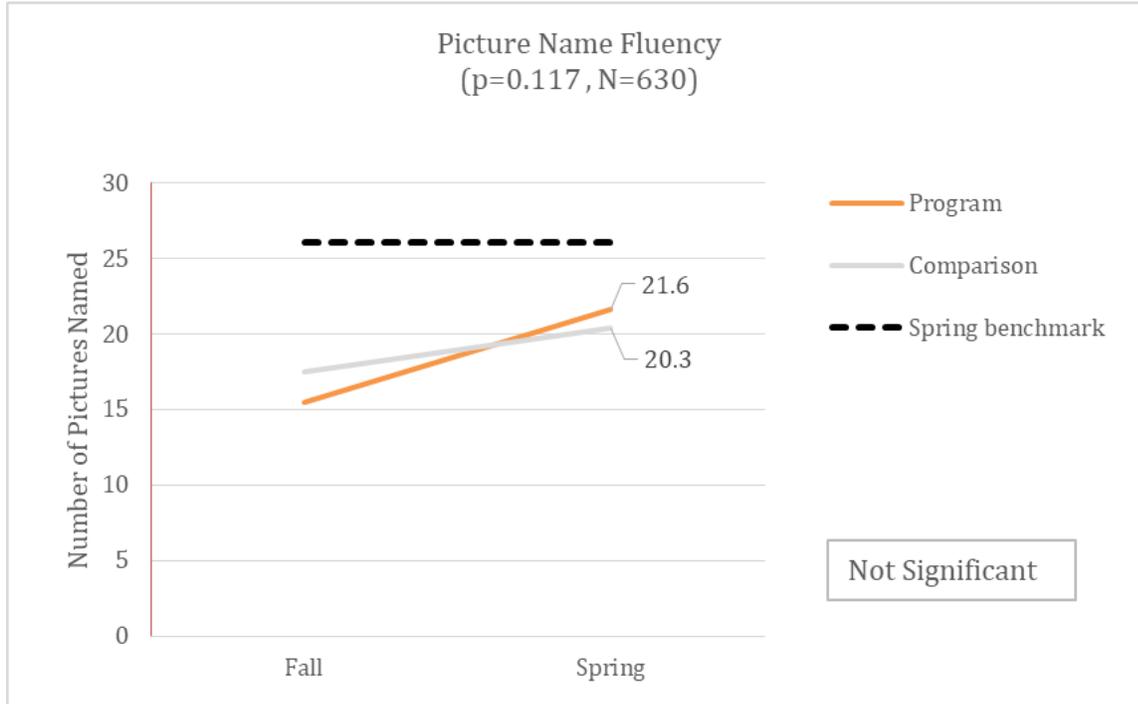
Figure 5. Mean Alliteration Fluency Scores for PreK program and comparison students



Picture Name Fluency

The Florida Reading Corps PreK program did not have a statistically significant effect on PreK students’ *picture name fluency* scores between the Fall and Spring benchmark ($p=0.117$, $N=630$). **Figure 6** presents the findings visually, showing that the program students’ average score of 21.6 was not significantly higher than the comparison group’s average score of 20.3. Also, neither group met the Spring benchmark of 26 correctly identified picture names. Therefore, the study observed no meaningful difference between the two study groups in vocabulary as measured by the *picture name fluency* assessment.

Figure 6. Mean Picture Name Fluency Scores for PreK program and comparison students



Does program effect vary by participant characteristics, such as demographics?

To address the study’s second research question focused on whether program effects vary by participant demographics, including gender, race/ethnicity, and DLL status, we conducted a set of subgroup analyses of PreK student outcomes. For those subgroups where the sample sizes were sufficient to conduct analyses, either the magnitude of the effects were small or not statistically significant. Given that they did not substantially differ from the full-group analyses, we direct the reader to the overall program effects described above.

Conclusions and Discussion

The findings from the Florida Reading Corps PreK outcome evaluation provide important evidence for addressing the study's research questions. Below, the evaluation team offers our conclusions on the effect of the Florida Reading Corps PreK program on PreK students' emergent literacy skills. Following our assessment of the questions is a discussion on the implications of our findings for the Florida Reading Corps PreK program.

1. What is the effect of Florida Reading Corps on program participants compared to students who did not receive Reading Corps?

The results of the study show that PreK students who received the Florida Reading Corps PreK program achieved significantly higher assessment scores for four of five tested emergent literacy skills by the Spring benchmark than did comparison students who did not participate in the program, specifically *letter name fluency*, *letter sound fluency*, *rhyming fluency*, and *alliteration fluency*. The program's combination of whole class and targeted interventions produced gains in these four outcome measures, including achieving the Spring benchmark expectation on average for *letter sound fluency* and *letter name fluency*. There were no significant differences between study groups for *picture name fluency* (i.e., vocabulary). We also contextualized our results as effect sizes and concluded that the results for the statistically significant measures were relatively small compared to the average score for students at comparison schools. These findings indicate that the Florida Reading Corps PreK program had a net positive effect on students' alphabet knowledge (i.e., letter name and letter sound correspondence) proficiency, even though both groups in the study on average exceeded end-of-year benchmarks for these two measures. The program had its largest, positive effects on both measures of phonological awareness (i.e., rhyming and alliteration), showing that it improves PreK students' phonological awareness development.

2. Does the effect vary by participant characteristics, such as demographics?

When sample sizes were sufficient, we conducted analyses to examine whether differential effects of the program existed for specific subgroups of students. However, the findings by subgroup were either not statistically significant or effect sizes were small. As a result, we were unable to draw conclusions about the effect of the program on demographic subgroups.

Discussion

Overall, the results of the evaluation show that the Florida Reading Corps PreK program had a positive effect on preschool students' alphabet knowledge and phonological awareness. Below, we contextualize these findings given the early implementation of this program in Florida (i.e., 3rd year), unique structural features of the program's replication/implementation in M-DCPS, and extraordinary, challenging events that occurred during the 2017-2018 school year. In spite of such challenges, the program was able to produce statistically significant results on four of five outcomes. These findings in combination with the strong results shown by the original model in Minnesota suggest that with more experience implementing

the program, increased integration of the program in whole class/school instruction, and ongoing improvement in terms of fidelity to program implementation, the Florida Reading Corps PreK program may produce even larger effects in the future.

In our previous outcome evaluation of the original Minnesota Reading Corps PreK program, we concluded that the program model is both effective and highly replicable in multiple settings for students often considered at-risk. This conclusion was based upon statistically significant, substantial gains across all five emergent literacy skills assessed,²⁰ as well as strong structural and organizational supports provided by the Reading Corps program and the schools participating in the program, which implemented the intervention model with fidelity. Importantly, when we evaluated the Minnesota Reading Corps program during the 2013-2014 school year, it was a mature program that had been operating across the state for eight years. In contrast, the evaluation of the Florida Reading Corps program took place during its third year of implementation. For the Reading Corps PreK program, this would be considered a relatively early period in their initial implementation phase,²¹ especially for a program that has an ambitious goal of whole class quality instructional improvement. Observations from the site visits confirmed that schools were continuing to work at implementing the program with fidelity. As such, it is recommended that the results of this evaluation be interpreted with the understanding that we assessed the *early* implementation of the Reading Corps model, and therefore, potentially more robust results may be observed after the program has matured to a full implementation stage. With four out of five measures showing a statistically significant effect, these findings are promising while at the same time highlight opportunities for program growth.

As noted previously, the Florida Reading Corps PreK program is a replication site for the Minnesota Reading Corp PreK program. However, it is important to note key differences in how the model was implemented in Florida. In considering program implementation factors, it may be helpful to review the findings from the previous process evaluation of the original Minnesota Reading Corps PreK Program, which identified several key elements of the program that were determined to be contributors to the large effects found on PreK students in the latter outcome evaluation of the program.²² Differences both in district-level infrastructure and in school-level environment observed during the Florida site visits conducted to select programs in the study, may require creative solutions in future programming.

For example, for the M-DCPS schools, the Florida Reading Corps Internal Coaches were district-level, literacy coaches who visited schools periodically to provide support to individual teachers whose classrooms were participating in the program rather than on-site, school staff appointed by the school principal, as was the case for most schools in the original Minnesota program. The previous study found that having members of the individual school staff serve as Internal Coaches was particularly effective, since these individuals would typically be more familiar and in regular contact with the principal and

²⁰ We assessed the same five skills in Minnesota as we did in Florida, using the same student assessments.

²¹ Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core Implementation Components. *Research on Social Work Practice*, 19(5), 531–540. <https://doi.org/10.1177/1049731509335549>

²² Hafford, C., Markovitz, C., Hernandez, M.W., Diaconis, A., Estrera, E., & Muyskens, P. (2014). *Process Assessment of the Minnesota Reading Corps PreK Program: Phase II Extension*. (Prepared under contract to the Corporation for National and Community Service). Chicago, IL: NORC at the University of Chicago.

teachers than an off-site coach. Familiarity and proximity aided in the development of critical relationships between the program staff (tutor and Master Coach) and the school (principal, teachers, and support staff). This is essential for the program to successfully change classroom practice and implement all components of the Reading Corps model, not just those components that can be implemented independently by Reading Corps tutors. In addition, it may be easier for an Internal Coach to provide the consistent, routine support and coaching needed by tutors if they are continuously on-site in the same school as the tutor. While the Internal Coaches were highly qualified and experienced coaches, the simple fact that they are external to the schools is a variation of the original program model that was identified as a critical role in facilitating teacher “buy-in” of the Reading Corps model, mentoring and coaching tutors, and facilitating positive relationships between AmeriCorps tutors and school staff.

Two additional differences in the site selection processes between the Minnesota and Florida Reading Corps evaluations warrant mentioning. First, in Minnesota state-wide implementation of the program permitted a site sampling procedure that included a broader representation of sites, and by extension, student backgrounds and educational need. By contrast, in Florida all sites were within an urban setting, and further, Reading Corps and M-DCPS partners had agreed to focus initial implementation on schools with the greatest need, known as "Preschools on Probation." As a result, the Florida Reading Corps occurred in sites with a higher proportion of students with greater need. Second, initial implementation of the Florida Reading Corps proceeded with M-DCPS district personnel selecting sites to participate, whereas in the Minnesota model statewide expansion had allowed many sites to self-select into the program. Given the quasi-experimental nature of the evaluations, the difference in self-selection between the two evaluations could have impacted the findings, because self-selection can be a proxy for many other variables including teacher and staff support, knowledge of program elements, and other factors that can impact student outcomes.

In addition to the structural differences in the program replication in Florida, it is important to note that several unfortunate events may have influenced the implementation of the Florida Reading Corps PreK program that were specific to the 2017-2018 school year, during which the study was conducted. In late August 2017, Florida experienced a Category 4 hurricane, Irma, less than two weeks into the school year, causing several weeks of serious disruption to the school schedule. In the days leading up to its arrival, the school district was closed and many Miami-Dade County residents were evacuated during the storm. The school district remained closed for a week following the hurricane, and many students, staff, and schools remained without power for some time. In our assessment, the hurricane appears to have disrupted at least a month – or over 10% of the school year – of learning.

Lastly, several tutors left their positions in the aftermath of the hurricane, which may have resulted in a cascading effect of tutor turnover throughout the school year. Tutor turnover rates were unusually high during the 2017-2018 school year, which may also have affected the Florida Reading Corps’ effect on students. In addition, the death of a tutor on the first day of service contributed to an emotionally difficult year for both program and school staff. The Florida Reading Corps PreK program experienced 100% staff turnover in the spring of 2018, possibly due in part to these unfortunate events. Taken together, these unique and emotional challenges may have had a negative effect on the effectiveness of the Florida Reading Corps PreK program during the 2017-2018 school year.

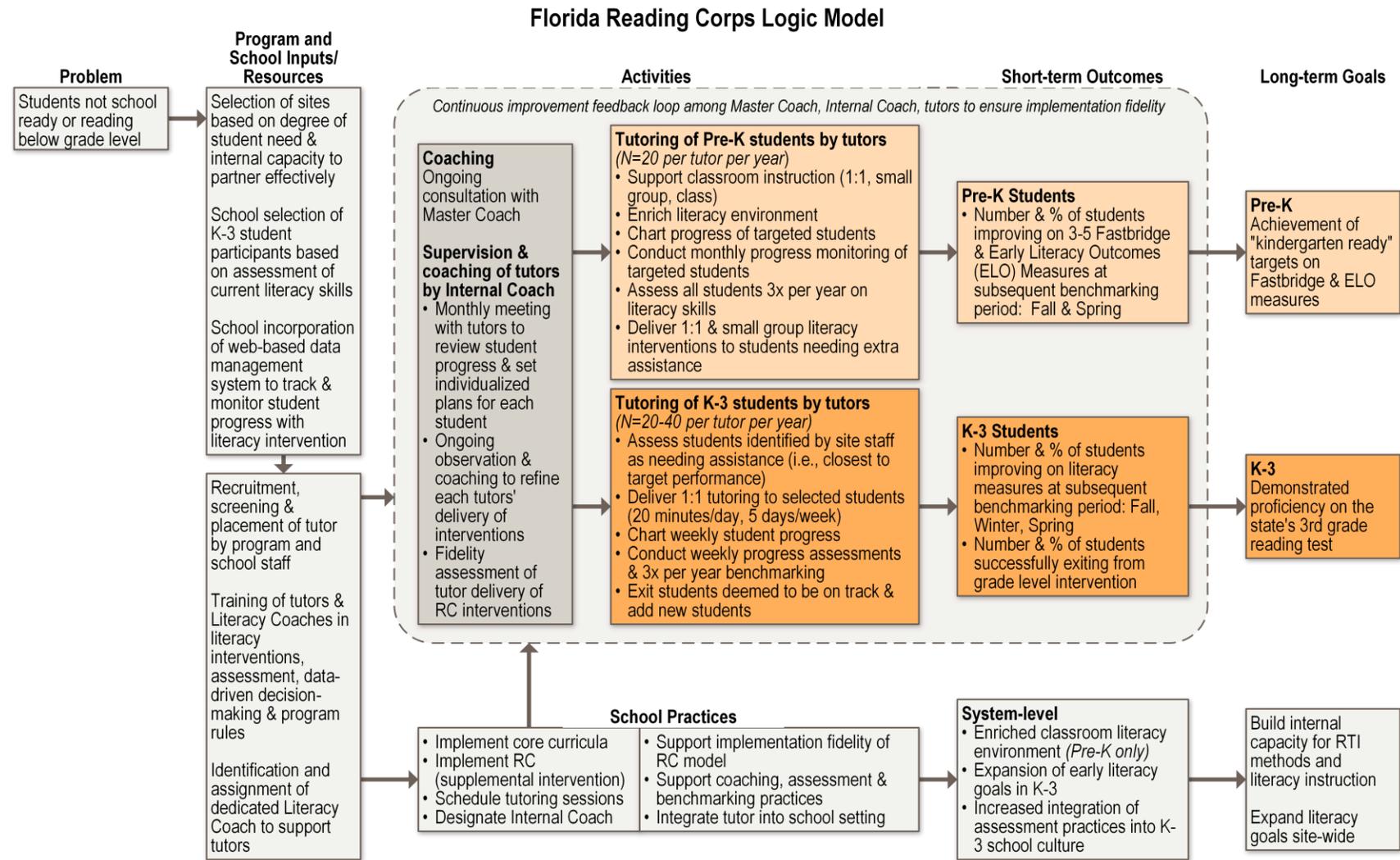
Program Implications

In conclusion, while the results of the study showed that the Florida Reading Corps PreK program produced statistically significant effects among PreK students on multiple measures of alphabet knowledge and phonological awareness, through ongoing improvement and increased implementation, Florida Reading Corps may have the potential to produce larger effects on students. Even if the contextual factors described above (hurricane, staff turnover, etc.) may have influenced the outcomes of the study, this evaluation provides the Florida Reading Corps Program several useful directions to consider for improving program implementation and ultimately its impact on student emergent literacy scores.

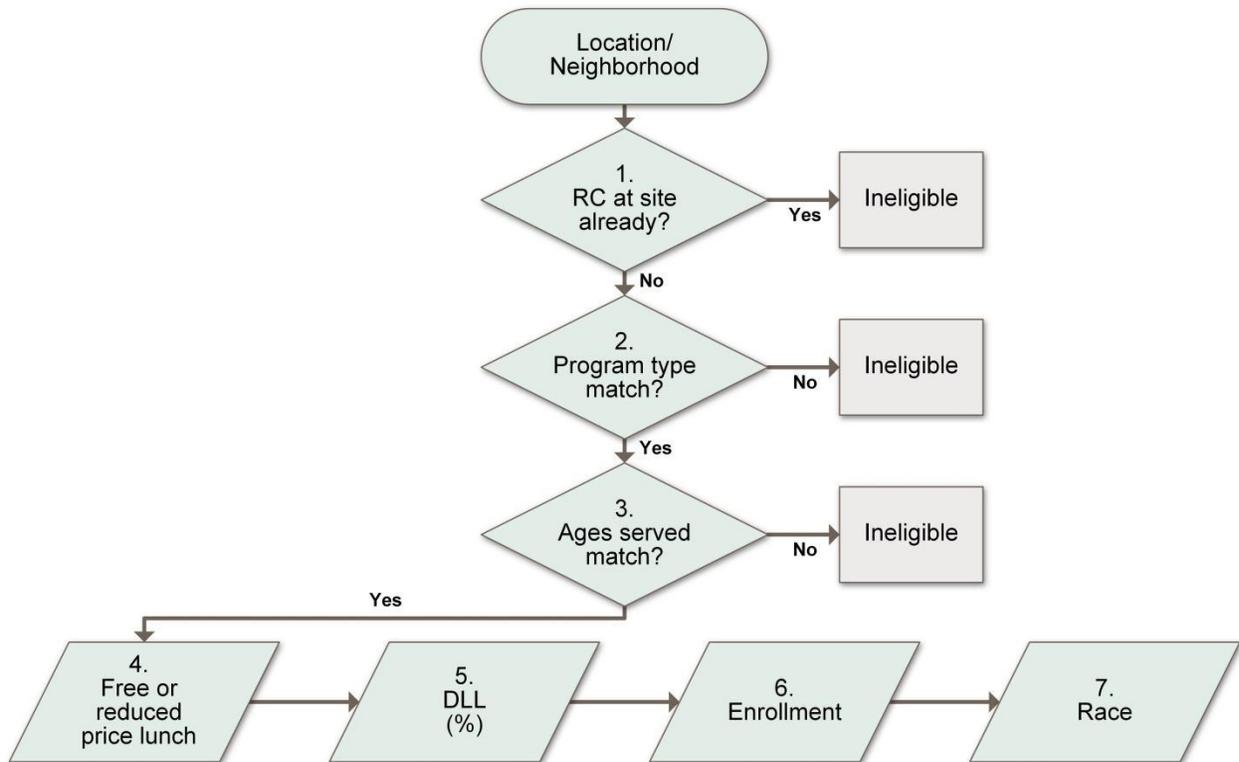
The evaluation team, having examined the Reading Corps program in Florida and other locations in multiple studies, in particular, a key mechanism for continued improvement in student outcomes would be additional emphasis on implementing the whole class components of the Reading Corps model. Tutors alone are capable of successfully implementing small group and one-on-one interventions with students they objectively identify as in need of extra instruction. However, to generate optimal effects on more complex skills, particularly vocabulary and phonological awareness, all adults in the classroom need to collaboratively implement strategies we believed were responsible for the effects we measured in the original Minnesota program, specifically the SEEDS quality adult-child interactions, “Strive for Five” conversations, transition activities, Literacy Rich Schedule fidelity implementation, organizing themes, and high quality Repeated Read Aloud. These elements are challenging for tutors to implement on their own and achieve the desired effect. They require complete buy-in from the lead teacher and effective collaboration with the tutor. When this occurs, children in the class experience significantly more high-quality instruction than they could obtain from the tutor themselves, or in the absence of the program. With strong principal leadership, informed and committed teachers, and talented tutors working together to implement the Reading Corps model, the Florida Reading Corps can experience the maximal effect of the program.

Appendix A: Logic Model

Figure A.1. Reading Corps Logic Model (for Outcome Evaluation)



Appendix B: Comparison Site Matching Process



Appendix C: Detailed Analysis Methods

Outcome Analysis. The analysis for the outcome evaluation employed a mixed model (defined below) to estimate the treatment effect for each outcome using data from PreK students. The statistical model uses site participation in the program as the primary predictor in a linear equation to estimate the difference between the averages of those students in participating sites to students in comparison sites. This predictor is coded as

$$P = \begin{cases} -.5 & \text{if assigned control} \\ .5 & \text{if assigned treatment} \end{cases}$$

Specifically, we fit two models. The first model for an outcome y for student i in site j is

$$(1) y_{ij} = \alpha_0 + \alpha_1 P_j + u_{0j} + e_{ij},$$

where α_0 is the overall average of site means of student outcomes, α_1 is the difference in the student averages of program sites and control sites, u_{0j} is the site effect (i.e., the difference between the overall average and the site-specific average), and e_{ij} is the student residual.

From this first model (1), we estimate an effect size, δ , defined as

$$\delta = \frac{\alpha_1}{\sigma} = \frac{\alpha_1}{\sqrt{\text{var}(u_{0j}) + \text{var}(e_{ij})}},$$

which defines the difference in standard deviation units; the standard deviation is estimated with

$$\sigma = \sqrt{\text{var}(u_{0j}) + \text{var}(e_{ij})}.$$

The second model includes controls, which were used for three reasons. First, when there is balance between program and comparison groups (i.e., they have similar averages of control variables), including controls in the model improves the statistical significance by reducing the variances of the random effects (u_{0j} , and e_{ij}) which comprise (in part) the estimated sample variance of the treatment effect. The second reason controls were employed was to account for any imbalances that exist between program and control groups. When small differences exist, the concern with estimated treatment effects is that some of the difference between treatment and control groups may be attributed to differences in sample demographics. By entering these variables into the model, these factors are held “constant” and thus estimate a plausible treatment effect net of the influences of the entered variables. The third reason is that the controls include the site-pair indicator which may also control for some unobserved characteristics of the sites that are common within pairs. For example, we do not have data on staff/student ratios, but if these are similar between sites within pairs, the pair indicators serve as a proxy control. Given that this is a quasi-experimental study, we make all efforts to control for any confounding factors.

The variables entered as controls include site pair indicators, English Language Learner status, gender, race, and most importantly, the Fall benchmark score. In addition to entering the student values into the equation, the study team also calculated the school averages of these variables and entered them into the model as well. The model which includes controls can be noted as

$$(2) y_{ij} = \beta_0 + \beta_1 P_j + \sum_c (\gamma_c X_{cij} + \lambda_c \bar{X}_{cj}) + q_{0j} + \sum_s^{S-1} \Delta_s + r_{ij},$$

where β_0 is the overall average of school means of student outcomes net of controls, β_1 is the average within-pair difference in the student means of program and control sites net of controls, q_{0j} is the site mean effect (i.e., the difference between the overall average and the school-specific average) net of controls, Δ_s is the blocking effect of a site pair indicator, r_{ij} is the student residual net of controls, and $\sum_c (\gamma_c X_{cij} + \lambda_c \bar{X}_{cj})$ represents all the effects of control variables at both the student and school levels. Note that, for example, $var(u_{0j}) > var(q_{0j})$, which is why we do not derive effect sizes from the second model (2).

In this report, we present estimates and statistical tests for the β_1 parameter for each outcome and grade. We also present marginal predictions of y_{ij} based on the specific values of the treatment predictor (-.5 for control and .5 for treatment) and setting all controls to sample average values. These represent the expected control and treatment averages for typical students. Finally, we present the effect size parameter estimates, δ , defined above.

Outcome Distribution. The distribution our outcomes is complicated by the two-step process in the assessment. The first step is a procedure that determines if a numerical score for the student can be assessed. If this hurdle is passed, the assessment continues and the student is scored with a continuous numeric score.

The challenge with this analysis is the meaning of 0 for the outcomes: it is less a score and more of an indicator of not being able to be scored. We explored if this needed to be handled with a more complicated (hurdle) model that separately considered the likelihood of 0s, and then produced a treatment effect given being scored.

After some simulation analyses, the conclusion is that a typical (linear) analysis is appropriate since the 0s are essentially a collapsed “clump” of what would otherwise be lower values of a latent ability variable that the instrument cannot detect. The simulation analysis concluded that a linear model on the data that are observed detects the correct effect with the same precision as a model on the full (unobserved) latent scale.

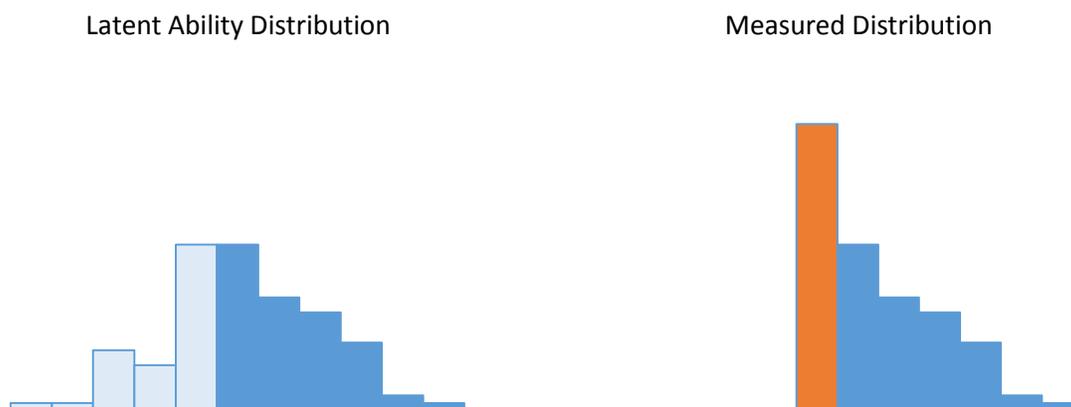
Below, we present details on our model of the measured outcomes and the simulations.

Latent and Measured Abilities

The idea behind most psychological tests is that some latent trait exists and varies from student to student. In our PreK outcomes, this trait is early reading ability, which we assume is roughly normally distributed in the population, as it height, weight, etc.

Suppose we had a tool that could measure the entire range of abilities, we would be able to draw a histogram like the one represented on the left side of Figure C.1. However, suppose that a portion of this distribution cannot be measured by our tool, such as the lighter blue portion of the left side in **Figure C.1**. Instead, we must collapse this portion of the distribution into one category, such as the orange portion of the right side of **Figure C.1**. Note that once we pass this threshold, we observe the latent ability distribution.

Figure C.1. The difference between the Latent Ability and Measured distributions (hypothetical data)



Simulation Analysis

We conducted a simulation study to determine the most appropriate method of analysis of data produced with tools such as the PreK measures, making the assumption that our measurement tool was collapsing a portion of the ability distribution as described above.

Our procedure was to create a normally distributed outcome for a treatment and control group (each with 300 cases) where the difference in means was 1 standard deviation. That is, the correct effect size for the analysis is 1. Our first analysis was to analyze the normally distributed outcome as if we were analyzing the latent ability.

Next, we collapsed the lower portion of the outcome’s distribution and analyzed it the same way we analyzed the latent ability. We then analyzed the data in two other ways: first, we eliminated the collapsed data, analyzing only the truncated portion of the distribution that reflects latent ability, and second, using a more complicated “hurdle” model that first models the collapsed portion, then models the ability portion

of the distribution. For each of the four analyses (latent, collapsed, truncated, and hurdle) we recorded the effect and standard error. We also calculated the squared error of each replications’ effect from the true effect of 1.

We replicated this exercise 500 times. We then examined several statistics from the simulations. First, we examined the average effect of each analysis. Second, we examined the mean squared error of the analysis, which is a measure of the extent to which the analysis misses the true effect. Next, we examined the standard deviation of the simulated effects from each analysis to estimate the sampling distribution of the analyses. Next, we examined the average standard error that was calculated from each analysis, which should reflect the standard deviation of the simulations if the analysis is correct.

The results of the simulation study are presented in **Table C.1**. The latent analysis, which analyzed the outcome before any collapsing or truncation is the gold standard and presented in the first row. We see that, on average, it detected the correct effect of 1 and did so with little error on average (mean squared error is very small: .007). We also see that the standard deviation of the effects, which reflect the sampling distribution is adequately measured by the analyses standard errors.

Table C.1. Results of Simulation

Analysis	Average Effect	Mean Squared Error from True Effect (1)	Standard Deviation of Effect (sampling distribution)	Average Modeled Standard Error of Effect (estimate of sampling distribution)
Latent	1.003	0.007	0.083	0.081
Collapsed	1.028	0.008	0.088	0.086
Truncated	0.488	0.267	0.071	0.075
Hurdle	0.998	0.032	0.180	0.177

The results were very similar for the collapsed analysis, which simply examined a linear model on the data where the lower end of the distribution is collapsed by the tool. We see that the collapsed analysis performs much like the ideal analysis on the latent ability: it measures and effect of 1 with little error and also estimates the correct standard error to reflect the sampling distribution.

The other two analysis did not perform as well. Truncating the data (removing those who could not be assessed) produced an effect that was much smaller, on average, then the true effect (.488 vs 1). It also had a much larger mean squared error and slightly smaller standard error.

The hurdle model produced the correct effect, on average, but had a higher mean squared error and much larger standard errors. Based on the simulation analyses, we conclude that the best approach to the analysis of the FL PreK data is to code those not able to be assessed as 0 and proceed with a linear analysis.

Appendix D: Results Tables

Within **Table D.1** below, the column labeled “Treatment effect” represents the β_1 parameter detailed in the analysis section, which represents the difference between the program and control group predicted means. The columns labeled “p-value” represent the probability of observing the data, or more extreme, if we assume the null hypothesis that the control groups have the same average outcomes. The “Effect size” column presents the results program and from a model without controls that is converted into standard deviation units. The final columns are the predicted program and control means in the outcome metric and the total sample size. Below we summarize the study’s key findings.

Table D.1. Florida PreK Reading Competency Results

Outcome	Treatment effect ^{a,b}	p-value ^a	Effect size ^c	Prog. mean ^d	Comparison mean ^d	Total N
Picture Name Fluency	1.287 (0.822)	0.117	0.061	21.63	20.34	630
Alliteration Fluency	1.082 (0.288)	0.000	0.220	5.40	4.32	615
Rhyming Fluency	1.176 (0.346)	0.001	0.301	7.25	6.07	630
Letter Sound Fluency	3.213 (1.449)	0.027	0.107	17.38	14.17	614
Letter Name Fluency	3.952 (1.600)	0.014	0.142	30.27	26.32	630

a: Reported statistics from model that includes the Fall assessment, age in months, gender, race, DLL status, and school pair indicators. The model includes students for whom demographic data is missing; missing value indicators included as covariates.

b: Cluster-robust standard errors in parentheses

c: Effect size reported from separate model without demographic control or site pair indicator variables, but including age in months

d: Reported statistics represent marginal predictions based on the model controlling for the Fall assessment, age in months, gender, race, DLL status, and school pair indicators.

Appendix E: Baseline Balance

NORC at the University of Chicago conducted a baseline analysis of the fall data collected from the Florida Reading Corps PreK program, and from PreK students at 10 matched comparison public elementary schools and two matched comparison YWCA early learning centers within the Miami metropolitan area who did not receive the Florida Reading Corps PreK program.

Results of our analysis of the fall assessment data provide snapshots of the typical student at Florida Reading Corps schools and at their matched comparison students. Fall assessment scores are the most important measures of baseline equivalence between the treatment and comparison groups because they are most correlated with the posttest outcomes of interest. NORC conducted baseline equivalence analyses, including chi-square tests of independence for assessment score validity²³, t-test analyses of two populations²⁴, and Hedges' g ²⁵ to determine effect size, to test whether the two groups were similar at baseline (i.e., prior to program intervention). When two groups are similar at baseline (i.e., no statistical differences and low effect size differences), we are able to make more confident conclusions about the program's effects on student outcomes post-intervention. In addition to statistical tests for differences, a key metric we examined was the Hedges g effect size, a measure of differences between groups that ideally should be below 0.15 standard deviations. A Hedges g effect size between 0.15 and 0.25 indicates that the analysis model should control for that variable. A Hedges g larger than 0.25 precludes the use of control variables to statistically control for baseline imbalance.

While some differences were found between sample groups on some of the baseline assessments, the overall balance between the program and comparison groups for PreK students in Florida is reasonable. These results are not surprising given that this is a QED study of the program with limited or imperfect information about schools' PreK assessment scores prior to matching. Because the sample is not randomized, but rather comprised of students from schools matched on school-level demographic characteristics, there is potential for unmeasured variables to influence the analysis in unpredictable ways. Note that the largest difference was in *picture name fluency*, which did not yield a significant effect in the final outcome analysis.

²³ Conducted as appropriate based on variable type (e.g. categorical, interval, ratio, etc.) A chi-square test of independence determines whether any significant relationship exists between two nominal (categorical) variables.

²⁴ Conducted as appropriate based on degrees of freedom, related to sample size and number of response categories. A two-sample t- or z-test for independence determines whether there is a statistically significant difference between means in two unrelated populations, or in this analysis, between two sample groups. This helps to establish whether the samples are significantly different in terms of any specific characteristics, as measured by variables.

²⁵ Hedges' g is a statistical measure of effect size. That is, Hedges' g describes how much one sample group differs from another in standard deviation units.

Table E.1. Florida PreK Fall Baseline Equivalence of Analysis Sample: Reading Competency Results

Outcome	Group Difference in Fall score ^{a,b}	p-value ^a	Effect size ^c	Prog. mean ^d	Comparison mean ^d	Total N
Picture Name Fluency	-2.043 (0.719)	0.005	-0.325	15.46	17.50	630
Alliteration Fluency	-0.188 (0.377)	0.618	-0.044	2.40	2.59	615
Rhyming Fluency	0.788 (0.436)	0.071	0.167	3.39	2.61	630
Letter Sound Fluency	-1.048 (0.220)	0.000	-0.150	4.34	5.39	614
Letter Name Fluency	-1.661 (0.673)	0.014	-0.114	11.15	12.81	630

a: Reported statistics from model that includes the age in months, gender, race, DLL status, and school pair indicators. The model includes students for whom demographic data is missing; missing value indicators included as covariates.

b: Cluster-robust standard errors in parentheses

c: Effect size reported from separate model without demographic control or site pair indicator variables, but including age in months

d: Reported statistics represent marginal predictions based on the model controlling for the age in months, gender, race, DLL status, and school pair indicators.